7-1-1994

Nanoelectronic Functional Devices

Pratheep Balasingam
*Purdue University School of Electrical Engineering*

Vwani P. Roychowdhury
*Purdue University School of Electrical Engineering*

Follow this and additional works at: [http://docs.lib.purdue.edu/ecetr](http://docs.lib.purdue.edu/ecetr)


This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.
NANOELECTRONIC FUNCTIONAL DEVICES

PRATHEEP BALASINGAM
VWANI P. ROYCHOWDHURRY

TR-EE 94-24
JULY 1994

SCHOOL OF ELECTRICAL ENGINEERING
PURDUE UNIVERSITY
WEST LAFAYETTE, INDIANA 47907-1285
Nanoelectronic Functional Devices

Pratheep Balasingam\textsuperscript{1} and Vwani P. Roychowdhury\textsuperscript{2}
School of Electrical Engineering
Purdue University
West Lafayette, IN 47907, USA

\textsuperscript{1} theep@ecn.purdue.edu
\textsuperscript{2} vwani@ecn.purdue.edu
Summary

Continuing advances in the miniaturization of electron devices have made possible the fabrication of nanoelectronic devices with feature sizes in the $1 \sim 100\,\text{nm}$ range. However, it is widely believed that conventional integrated circuit design techniques will become impractical, due to the small size and the low current carrying capacity of nanostructured devices. Proposals which envision novel ways to circumvent this problem have begun to appear in earnest. The majority of these proposals use globally coherent quantum systems to generate computational abilities. We differ from these proposals in that we use semiclassical global models, as the basis from which to conceive nanoelectronic functional devices. An additional point of departure, is our assumption that currently perceived limitations to realizing interconnects amongst the nanostructured devices, will in time be overcome. Finally, we restrict attention to niche applications, in which the collective activity of a large number of nanostructured devices give rise to useful computational functions. The special purpose functional device concept adopted here can be contrasted with other approaches which envision the design of general purpose computers on the basis of quantum mechanical logic gates.

We adopt a research methodology in which computational tasks which are naturally suited to a collective solution strategy are first identified, and then mapped to nanoelectronic physical systems. In making these associations we make well-defined assumptions concerning the properties of interconnection networks. The justification for this approach comes from the extensive experimental activity on novel wiring technologies tailored specifically for nanoelectronics. An aspect of these technologies is the fact that electronic transport along these wires can themselves introduce nonlinearities which can influence the global behavior of networks.

The work discussed in this report has been unified under a particular technology based on the creation of arrays of nanometer-sized metallic islands. We then consider different types of network mechanisms for the transfer of electrons between islands. Depending on the types of transport nonlinearities permitted by the network links, we show that it is possible to generate different kinds of global activity in these networks. We show, in addition that it is possible to impart a computational interpretation to these global activities. In particular, we show that within a classical circuit theoretic model, non-monotone nonlinearities in the local transport can yield global associative memory effects. We then show that this interpretation will remain valid even when single-electron effects come into play, provided that the effective capacitance of the nanometallic islands is not too small. We then investigate networks of islands in which the sole nonlinearity arises from single-electronics. These networks are shown to be capable of associative memory effects, as well as yield approximate solutions to certain NP-complete optimization problems, provided that there is sufficient flexibility in the choice of inter-island capacitances.
## Contents

1 Introduction 1

2 Primitive and Functional Devices 2
   2.1 Nanostructures and Integrated Electronics 2
       2.1.1 Mesoscopic Device Physics 2
       2.1.2 Semiclassical Functional Devices 4
       2.1.3 Quantum Coherent Computing 5
   2.2 An Approach to Functional Devices 5
       2.2.1 Methodology 6
       2.2.2 Nanoscale Interconnection Networks 7
       2.2.3 Input and Output 7

3 Networks of Nanoscale Metallic Islands 8
   3.1 Neuromorphic Continuous-Charge Networks 8
       3.1.1 Liapunov Stability Analysis 10
       3.1.2 System with Capacitive Networking 15
       3.1.3 Illustrative Examples 16
       3.1.4 Technological Issues 20
   3.2 Limits to Continuous Models: Single-Electron Effects 20
       3.2.1 Modeling Single-Electron Effects 21
       3.2.2 Monte Carlo Simulation Technique 22
       3.2.3 Illustrative Examples 22
       3.2.4 Technological Issues 24
   3.3 Ising-type Single-Electron Networks 24
       3.3.1 A Single-Electron Computational Module 25
       3.3.2 Networks with Local Capacitive Coupling 27
       3.3.3 Programming Stable States 28
   3.4 Single-Electron Networks for Combinatorial Optimization 31
       3.4.1 A General Methodology 31
       3.4.2 Examples 35
   3.5 Alternative Networks 45
       3.5.1 Electron Dynamics in Networks with Quantum Links 45
       3.5.2 Hybrid Functional Devices 46

4 Concluding Remarks 46
1 Introduction

There is at present a worldwide effort to overcome technological barriers to nanoelectronics, and recently there has been a significant increase in experimental activity involving electron devices with feature sizes in the 1 – 100nm range [43, 31, 54]. While nanoelectronics offers the promise of unprecedented increase in computational power, it is well known that conventional strategies for the integration of devices to form complex circuits will be impractical due to the small size of nanoelectronic devices. Consequently, many novel proposals have recently been advanced, articulating primarily, visions of future nanoelectronic systems [40, 8, 35, 42]. The majority of these proposals limit consideration to global quantum coherence as the basis from which complex computational abilities are generated. We differ fundamentally from these proposals in that we use semiclassical models, as the basis from which to conceive nanoelectronic functional devices. An additional point of departure, is our assumption that currently perceived limitations to realizing interconnects amongst the nanostructured devices, will in time be overcome. Finally, we restrict attention to niche applications, in which the collective activity of a large number of nanostructured devices gives rise to useful computational functions. The special purpose functional device concept adopted here can be contrasted with other approaches which envision the design of general purpose computers on the basis of quantum mechanical logic gates [40, 8].

<table>
<thead>
<tr>
<th>§3.1</th>
<th>nonlinear resistive</th>
<th>spatially variable resistances, access to interior nodes</th>
<th>Liapunov stability theory</th>
<th>image processing, associative memory</th>
</tr>
</thead>
<tbody>
<tr>
<td>§3.2</td>
<td>single-electron tunneling</td>
<td>same as above, and small capacitance C</td>
<td>Liapunov theory and monte carlo simulations</td>
<td>associative memory (breaks down for very small C)</td>
</tr>
<tr>
<td>§3.3</td>
<td>single-electron tunneling</td>
<td>access to interior nodes, and arbitrary choice of capacitances C&lt;sub&gt;ij&lt;/sub&gt;</td>
<td>linear programming</td>
<td>associative memory</td>
</tr>
<tr>
<td>§3.4</td>
<td>single-electron tunneling</td>
<td>same as above</td>
<td>matrix theory and discrete math</td>
<td>combinatorial optimization</td>
</tr>
<tr>
<td>§3.5</td>
<td>phase-coherent</td>
<td>molecular wires</td>
<td>nonlinear dynamics</td>
<td>effort, in progress</td>
</tr>
</tbody>
</table>

Table 1: Summary of key results discussed in §3, for systems of large arrays of nanometallic islands with different types of interconnection mechanisms.

In this report, we will describe a general approach to the investigation of nanoelectronic functional devices, and study in detail the information processing capabilities of several specific networks built from arrays of nanostructured metallic islands. In section 2 we will outline the basic principles of nanoelectronic device physics. We will then propose to use these device concepts to create collective systems, by making certain assumptions concerning interconnection and input/output technologies suitable for nanoelectronic integration. Section 3 contains the key results of our work which are summarized in table 1. In that section we will adopt a research methodology which consists in exploring the characteristics of nanoelectronic networks which bear strong correspondences to well-known collective computational models. Several types of networks based on both the continuous-charge, and the single-electron picture are shown to exhibit collective computer...
tional characteristics. The networks are investigated both theoretically and numerically. We will also outline the behavior of networks in which local transport occurs phase-coherently, while the global dynamics is described semiclassically. The report will be concluded with a summary in section 4.

2 Primitive and Functional Devices

The primitive building block in microelectronics is the transistor, whose switching property is exploited in the design of complex computers. The design problem leading up from the switch to the computer is divided into a hierarchy of smaller design problems, each of which addresses the optimization of a subsystem [22]. In this scheme, behavioral optimization at a given design level can proceed with characteristics passed along from the previous level, without much regard to any other levels. This neat demarcation of design levels is likely to be threatened, over the next decade, or two, by the rapid advances in the miniaturization of electron devices. Over time, we can expect that several lowermost design levels may need to be merged together, to form a single design level which optimizes 'functional devices', which will then constitute the building blocks from which computing machinery will be fashioned. The degree of functionality achieved by a device of this kind will depend both on the physical phenomenon being exploited in its design, and on the economic resources demanded by the design effort. The notion of a 'functional device' is in fact of long vintage [48], and has in recent years been revived within the context of quantum electron devices [15].

In section 2.1 we will review basic concepts of nanoelectronic device physics, and present a scenario in which these concepts can be applied to the creation of functional devices. The research methodology and the key assumptions we make about enabling technologies are summarized in section 2.2.

2.1 Nanostructures and Integrated Electronics

Integrated circuits with field-effect transistor (FET) gate lengths as small as 250 nm can at present be manufactured. There is also a considerable amount of research work dedicated to discrete FETs with gate length smaller than 100 nm. The issues addressed in scaling electron devices into this size regime are an assortment of short-channel effects, all of which can be understood within the Boltzman transport equations of conventional device physics [25, 62, 57]. There are indications, however, that adroit use of technology cannot continue to overcome more fundamental effects which will begin to influence device behavior below the 100 nm gate length regime [70, 53, 68]. Once this bottleneck is encountered at some point within the next two decades, novel functional device concepts based on nanoelectronics will experience intensive development.

2.1.1 Mesoscopic Device Physics

An electron device which is scaled down well below the 100 nm mark, in more than one of its three dimensions, may reveal quite new physical effects, arising from the wave, as well as the particle aspects of electrons. These effects have been observed in a class of nanostructured laboratory devices, which are frequently called mesoscopic devices. These devices are very similar to FETs of conventional electronics [9], in that they are usually three terminal devices with clearly identifiable source, drain and gate terminals. In addition to requiring very low operating temperatures, these devices also differ from their microelectronic counterparts in that the gate width, the dimension transverse to the direction of current flow, is made very small to accentuate the wave, and particle
Figure 1: A schematic reduction of typical semiconductor realizations of nanoelectronic concepts to collective functional devices based on arrays of metallic islands. Arrays of very small metallic islands can be created by techniques which are not limited by lithography. (a) and (b) MODFET-type semiconductor devices which exhibit single- and phase-coherent electronics at $T \sim 1K$. (c) and (d) Much smaller but similar metallic devices, which have higher operating temperatures either due to low capacitance, or due to the need for phase-coherence only over short distances. (e) and (f) Collective semiclassical systems which can be analyzed, by including quantum effects locally over the tunnel barriers, or on the quantum links. Each island is approximated as an electron reservoir, in which energy and phase information is dissipated.

Aspects of the electrons. In conventional FETs, when gate lengths are scaled down into the deep-submicron regime, the gate widths are maintained at relatively large sizes ($\sim 10\mu m$), so that while the switching times are scaled down, the switched currents remain large. Thus, for example, a single electron transistor (SET) whose active region might be confined in all dimensions to within 100 nm length scales, is a substantially smaller device than a FET with a gate length of 100 nm, and can switch only a very small current. The very low capacity for currents, as well as the small overall
device dimensions are aspects which prohibit the use of conventional interconnect technologies to integrate nanoelectronic devices.

The wave nature of electrons has been demonstrated in experiments which involve a waveguide-like device [9], whose transport properties can be understood using the Landauer formalism [14, 17, 15], which can be stated concisely in terms of the conductance formula,

\[ G = \frac{e^2}{\pi h} \left( \frac{T}{1 - T} \right) \]

where \( T \) is the quantum mechanical transmission probability for electrons issuing from the source and entering the drain electrode in figure 1(b). This conductance formula is usually used to describe the behavior of individual devices, which have been configured in the manner shown in figure 1(b). In our work, we will not invoke conductance formulae of this kind to model devices, but rather to describe the behavior of nanoscale wires which connect devices. A key requirement for the observability of this transport mechanism is \( L < L_\phi \), where \( L_\phi \) is a characteristic size of the wire and \( L_\phi \) is the phase coherence length of electrons. The phase-coherence length is a sharply decreasing function of temperature, which implies that smaller wire sizes will favor higher operating temperatures.

The particle aspect of electrons can also be demonstrated in a waveguide-like device, as shown in figure 1(a), provided that the electronic capacity of the device is sufficiently small that the addition of a single electron to the cavity region of the device can cause a substantial change in its electrostatic potential [29]. A key manifestation of this effect is the fact that the discrete transfer of electrons through an insulating junction of small area can give rise to a gap in the current-voltage characteristics of that junction [33]. This can be understood by examining the tunnel rate through the junction,

\[ \Gamma = \frac{J(\Delta E/q)}{q(1 - \exp(-\Delta E/kT))} \]

where \( J(V) \) is the current-voltage characteristics of the junction under a large area assumption, and \( \Delta E \) is the energy dissipated as a result of the transfer of a single electron with charge \( q \). We can deduce from the above that the dissipated energy \( \Delta E \) must be much greater than the thermal energy \( kT \) in order to clearly observe the Coulomb gap. The dissipated energy, on the hand is inversely related to the junction capacitance \( C_\phi \), which shows that smaller device dimensions will favor higher operating temperatures.

It is apparent from the foregoing that it is desirable to make devices and wires as small as possible, in order to elicit single- or phase-coherent electronic effects, at relatively high operating temperatures. This issue will play a significant role in our choice of material systems for the design of functional devices, since room temperature operation is a key target of this effort. The possibility of observing single-electron effects at room temperature has already been experimentally demonstrated, in systems involving metallic [56], and polysilicon particles [69]. Phase-coherent effects can in principle be observed at room temperature in short molecular wires.

2.1.2 Semiclassical Functional Devices

We propose to investigate the collective properties of networks of devices which derive their primitive attributes from the mesoscopic devices outlined in section 2.1.1. Despite the dual character of the physics exemplified by the single- and phase-coherent electron devices, these two classes of devices share several common design features, as might be deduced from figure 1(a) & (b). We seek therefore a design strategy which might allow the investigation of both types of effects in an integrated fashion. Now, the pioneering experiments which demonstrated the controllability of wave and particle [29]
effect devices were carried out on lithographically defined semiconductor nanostructures. These conventional designs, however, are not suitable for our study, due to bottlenecks associated with very low operating temperatures, and the lack of tractable ways to create nanoscale interconnection networks.

We envision, instead the creation of systems based on arrays of nanostructured metallic islands. There is experimental evidence that these metallic island arrays can be created by deposition techniques which are not lithographically constrained, and that it is practical to establish inter-island conductive links with molecular wires [51]. Although, all of the work described in this report will be anchored around physical realizations of approximately the same kind, we show that very different network dynamics can be generated if different properties are ascribed to the interconnection network linking the islands. For example, two distinct types of network behavior can be elicited from the two collective systems shown in figures 1(e)&(f). The global dynamics of each of these networks is described semiclassically, while quantum effects are confined to local transport effects between islands. In particular, we will assume that each metallic island, though nanostructured, is large enough to be treated as an electron reservoir in which energy and phase information is dissipated, permitting a semiclassical global description. Equations 1 and 2, and their generalizations will in effect be used as terminal characteristics of the discrete elements, from which global network equations are then derived. In section 3 we will also show that it is possible to use the same metallic island array, together with a resistive network to generate nonlinear dynamical equations which treat electronic charge neither as waves, nor as particles, but rather as a fluid in the manner used in microelectronics. It is also possible to conceive hybrid networks in which different network mechanisms coexist to varying degrees.

2.1.3 Quantum Coherent Computing

In recent years, there has been a marked increase in the number of proposals which advance perspectives on how one might design computing machinery starting from nanoelectronic primitives [40, 8, 35, 42]. These proposals, however, are very different from the semiclassical approach outlined in section 2.1.2. In particular, most of these proposals involve quantum mechanics at a global level, and therefore, require phase coherence over large distances which in turn requires very low operating temperatures. Furthermore, in some cases the behavior of the functional system is not conclusively understood due to the lack of adequate computer simulation tools to solve demanding time-dependent many electron quantum mechanical problems. In addition, several of these approaches develop quantum mechanical logic gates, and build integrated circuits by wiring these gates together following the fashion of conventional VLSI circuits. Issues concerning directionality, and low gain remain moot within this scheme. Finally, Landauer has pointed out a set of fundamental problems, such as localization, and reversal of computation, which are likely to be encountered by computing systems, which invoke quantum mechanical coherence in an essential way [39].

2.2 An Approach to Functional Devices

Our approach is based on the observation, that complex spatial and temporal patterns of the electronic charge distributions, resulting from the nonlinear interactions amongst thousands of devices, which are collectively driven far from equilibrium by applied signals of high amplitude may reveal a fundamental kind of computational effect [49]. Our methodology for the exploration of these ideas, and assumptions concerning nanoscale interconnect, and input/output technologies are described in the following paragraphs. Certain technological assumptions are speculative in nature, and are buttressed only by pointing to available technologies which are in various stages of development.
This work is oriented mainly towards the demonstration of nontrivial collective activity in arrays of mesoscopic structures. In consequence, conclusive solutions to the technological problems associated with input and output are not attempted. However, this issue will be approached by positing 'what-if?' type questions in our discussion of specific systems in section 3.

<table>
<thead>
<tr>
<th>Foundations</th>
<th>Technology</th>
<th>Mechanisms</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collective Computing</td>
<td>Neural Networks</td>
<td>Novel Deposition Techniques</td>
<td>Arrays of Metallic Islands</td>
</tr>
<tr>
<td>Nonlinear Dynamics</td>
<td>Nanoelectronic Device Physics</td>
<td>Molecular Wires</td>
<td></td>
</tr>
<tr>
<td>Nanotechnology</td>
<td>Neuronal Networks</td>
<td>Charge models</td>
<td></td>
</tr>
<tr>
<td>Device Physics</td>
<td></td>
<td>Quantum links</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hybrid concepts</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Functional Devices</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Numerical Modeling</td>
<td></td>
</tr>
</tbody>
</table>

Figure 2: Proposed research methodology. Pictorial representations of the array, and the various network mechanisms are given in figures 1 and 3.

### 2.2.1 Methodology

In view of the speculative nature of this work, we shall briefly set forth our method in this paragraph. As was indicated in section 2.1.2, we restrict attention to electron dynamics in arrays of metallic islands, and introduce design flexibility by invoking different types of interconnect technologies. Depending on the character of electron transport in the network links, it is possible to generate quite different network dynamics. An advantage of this unified approach is that system behavior determined under the assumption of a particular kind of network mechanism can be used to clarify the global behavior of systems whose network mechanisms render them less amenable to theoretical analysis. An example of this is illustrated in sections 3.1 and 3.2, where we use the analytical results obtained using a continuous-charge model to examine the numerical results obtained for a corresponding single-electron network.

Our broader strategy consists in finding nanoelectronic realizations which bear strong correspondence to well-known computational models which are suited to a collective solution method. Once such strong correspondences are established, extensive exploration of the parameter space is carried out using numerical simulators which capture nanoelectronic phenomena in considerable detail. Since the principal focus of this work is on the collective activity of large arrays of mesoscopic devices, it is essential to be able to simulate very large numerical problems. This computational challenge has been met by implementing our simulation tools on several parallel supercomputing systems.
platforms.

In creating a functional device, we can choose from among logic, memory and communication applications, the three main branches of information handling technologies [32]. In this report we will consider, primarily, memory-like devices which operate on the basis of neuromorphic principles. We show in section 3 that it is possible in general, to establish a connection between neuromorphic computational models and arrays of metallic islands, in the presence of each of the different network mechanisms indicated in figure 2. For example, when the network dynamics is assumed to be governed by single-electron effects, the choice of a neuromorphic approach is motivated by a straightforward comparison between the dynamics of electrons in an array of Coulomb islands, and the dynamics of discrete Hopfield networks [26]. Additional impetus for restricting attention to memory applications comes from the realization that single-electron dynamics is stochastic in nature, which hinders the conceptualization of logic applications. Associative memory applications, on the other hand, can benefit from an appropriate kind of stochasticity. We show also that the stochastic dynamics of single electrons can be used to obtain approximate solutions to several types of optimization problems, which are conventionally solved by simulated annealing. Additional examples following this fashion are discussed in section 3.

2.2.2 Nanoscale Interconnection Networks

In this report we will postulate the availability of different types of interconnection networks. This will require that we shift out of the context of lithographically defined metallic wires, and into a different type of interconnect technology. Recent research on molecular networks [51], and nanotubes [11, 61], are indicative of approaches which may lead to viable nanoscale links. In microelectronic integrated circuits, the interconnects are essentially linear wires which convey signals from one device to another, without operating on the signals [45, 32]. We believe, that nanoscale interconnection networks will depart from this picture in that the nonlinearities inherent in the transport across molecular links can be made to contribute constructively to the global functionality of the network. Experimental work already indicates that it may be practical to establish conduction between quantum dots via molecular wires. There is also evidence that phase-coherent links may prevail between metallic spheres which have been embedded in a porous dielectric matrix [12]. In addition to providing nonlinear conductive links between devices, nanoscale networks of this kind might also make it possible to engineer the dielectric region linking non-adjacent devices, so that the capacitive interaction between them is strengthened. The technology of nanoscale interconnects is in a rudimentary state, and at present lacks the capability to arrange detailed interactions of the kind, which might allow the emergence of complexity. However, this technology is evolving rapidly at present, and may in time lead to the realization of self-assembling networks which might provide an adequate degree of design flexibility. The theoretical and numerical results obtained from our work, can help visualize the potential of experimental interconnection networks which are currently under investigation.

2.2.3 Input and Output

A key problem in the design of nanoelectronic functional devices concerns the placement of input and output ports. Several of the recent proposals for achieving nanoelectronic computing, adopt edge-driven schemes where the signals are applied, or retrieved at the periphery of a network of devices [8, 40]. In addition, these schemes disallow the large-scale transport of electrons across the networks. Instead the effects of an electrostatic [40], or magnetic disturbance [8] imposed at the periphery of the network are assumed to propagate into the interior, when each device influences
the electrostatic or, magnetic polarization states of its nearest neighbors. In our work, we have found that the edge-driven concept, which disallows a steady flow of electrons is limited in scope, due to the high probability that the system will get trapped in a large set of spurious metastable states. We will in general assume that at least a subset of the interior nodes of the array can be directly contacted and supplied with charge and energy. In addition, we will consider networks with and without the steady flow of electrons in section 3, and discuss the merits of each approach. An approach which has been discussed in the literature [35] consists in using optical pumping signals which can stimulate, and probe each device individually. In our work we focus primarily on electrical signal processing, and a plausible approach which might permit the introduction of charge and energy to a subset of the interior nodes of a large array is suggested by recent work on arrays of scanning tunneling microscope tips, in which each tip has been individually addressed [55].

3 Networks of Nanoscale Metallic Islands

In this section we will study the behavior of systems whose building blocks are metallic islands, whose diameter d is in the range $1 - 100 \text{nm}$. At the upper end of this range the pertinent device characteristics can be adequately modeled using microelectronic considerations. Consequently, the properties of networks consisting of relatively large islands can be analyzed using classical circuit theory by regarding electronic charge as a continuous variable. For smaller islands with $d \sim 10 \text{nm}$, the discreteness of electronic charge will play a key role, which will necessitate the use of Monte Carlo simulation techniques. However, even for an island with $d \sim 10 \text{nm}$, quantum mechanics need not be used in an essential way, since in contrast to a semiconductor quantum dot of similar size, a nanoscale metallic island can contain anywhere from $10^2$ to $10^5$ conduction electrons [3], which will make a bulk-like treatment accurate. Phase-coherent effects may also be introduced into this system by connecting the islands with a network of molecular wires. It is clear therefore, that continuous charge, single-electronic and phase coherent network equations can all be associated with generic systems of the kind depicted in figure 3.

The systems we shall study will consist of arrays of islands, which have been deposited on a layer of material with non-ohmic electrical characteristics, which in turn has been grown on a conductive substrate, as shown in figure 3(a). In section 3.1 we shall investigate the neuromorphic behavior of a nonlinear network using classical circuit theory, and then in section 3.2 we shall investigate the same network when single-electron effects become pertinent. In section 3.3 we will study a network in which single-electron effects provide the sole nonlinearity, and establish a strong correspondence between that network and discrete Hopfield neural networks. In section 3.4 we show that it is possible to obtain approximate solutions to several optimization problems, using directly the stochastic dynamics of single electrons. Finally, in section 3.5.1 we will develop a model for the inclusion of quantum links between metallic islands.

3.1 Neuromorphic Continuous-Charge Networks

In this section we will examine networks of the kind shown in figure 3(a), with continuous flow of charge as shown in figure 3(b), and will make precise the general observation that nonlinear networks can be used to generate collective computational effects [71]. A one-dimensional version of the particular network realization under investigation is shown in figure 4, in which we assume that the current between islands i and j is $J_{ij}$, and that the current between island i and the
Figure 3: (a) A generic array of metallic islands deposited on a non-ohmic layer which has been grown on a conductive substrate. It is assumed that all islands have direct conductive links to nearest neighbors. A subset of the islands serve as program nodes which are driven by current sources. Another subset of islands serve as input/output ports, and the remaining islands introduce complexity to the system through nonlinear conductive links to the substrate. The substrate nonlinearity could arise for example from resonant tunneling. (b),(c)&(d) Cross-sectional view showing island $i$, whose potential $v_i$ is influenced by the charges $Q_j$ in the rest of the network through a capacitance matrix. The continuous-charge dynamics depicted in (b) will be investigated in section 3.1. An analogous system governed by single-electron dynamics is depicted in (c), and will be discussed in sections 3.2, 3.3 and 3.4. Dynamics of networks with quantum links, such as shown in (d) will be discussed in section 3.5.1.

grounded substrate is $J_{is}$. We can then write from Kirchoff's current balance condition that,

$$\frac{d}{dt}q_i = C_i \frac{d}{dt}v_i = - \left[ J_{is}(v_i) + \sum J_{ij}(v_i - v_j) \right] + I_i(t) \quad (3)$$

where $I_i(t)$ is the driving current, $q_i$ is the charge, $v_i$ is the potential, and $C_i$ is the capacitance, with the subscript $i$ indicating the relevant island. This equation bears strong similarities to the equations representing the additive short term memory (STM) models of neural networks \[23\]. However, the circuit realizations of the additive STM model involve amplifiers as well as a massively interconnected network of resistors \[27, 44, 47\]. We will show in section 3.1.1 that, even without amplifiers, and massive connectivity it is possible to generate nontrivial computational abilities.
using equations 3, if certain reasonable nonlinearities are permitted to enter the network equations. In particular, we show that non-monotonic nonlinearities such as that shown in figure 4(b), are necessary to obtain nontrivial collective activity in these networks. This result can be contrasted with the general requirement of monotonic nonlinearities in STM systems [23]. Our result, also has immediate consequences to nanoelectronics, in that staircase nonlinearities, which have been produced in quantum constriction \[64, 66\] and assymmetric double junction devices \[1\], can be seen to be inadequate for the realization of collective activity. However, there are several other electronic devices which can produce non-monotonic nonlinearities \[15\]. For the purposes of this section we will be interested chiefly in semiconductor heterostructure devices which can in principle be integrated vertically beneath the array of islands. There has been some prior research involving resonant tunneling diodes (RTD), for the realization of neuromorphic systems \[41\]. There has also been some early research activity on networks of tunnel diodes \[59, 67\], which also exhibit current-voltage characteristics of the kind shown in 4(b). However, our work will focus on nanoelectronic realizations, in which complex circuitry cannot be integrated with the basic elements, due to the small size of the islands. In addition, while the examples discussed in section 3.1.3, are limited to substrate nonlinearities of the type depicted in figure 4(b), the Liapunov analysis in 3.1.1 is much more general. In fact, these results will be used to guide our study of related single-electron networks in section 3.2, as well as our on-going investigation of networks with quantum links, which will be briefly considered in section 3.5.1. In the latter case more elaborate nonlinearities with multiple peaks and valleys could become relevant.

3.1.1 Liapunov Stability Analysis

Qualitatively, the multiple solutions to the nonlinear system of equations which results from equations 3 under steady-state conditions \((d\nu_i/dt = 0\) for all \(i)\) will be taken as a set, of memory states which can be programmed by properly choosing the current biases \(I_i\). The current biases \(I_i\) will be assumed to be either time independent, or slowly varying on time scales over which the network relaxes into its memory states. If the network is begun at time \(t = 0\), with a certain initial condition \(\nu_i = q_i/C_i\) for all \(i\), arranged, for instance, by an initial impulse of charges \(q_i\) at each of

![Figure 4: (a) A one-dimensional array of islands, whose dynamics is described by equation 3 (b) A non-monotone substrate nonlinearity of the kind shown here, is the minimal condition for the realization of collective effects.](image)
the islands, then the network will evolve towards the closest memory state, as programmed by the current biases \( I_i \). In the remainder of this section, we will rigorously derive conditions on the forms of the nonlinearities in \( J_{i*} \) and \( J_{ij} \), which will allow the emergence of associative memory effects.

We rewrite the system of equations 3 under the assumption that the current biases \( I_i \) are time independent:

\[
C_i \frac{d v_i}{dt} = I_i - \sum_{j \neq i} J_{ij}(v_j - v_i) - \sum_{j > i} J_{ij}(v_i - v_j),
\]

where \( J_{ij}(\cdot) \), for all \( i = 1, \ldots, n \), and \( J_{i*}(\cdot) \), for all \( i = 1, \ldots, n \), are piecewise continuous functions, and \( J_{ij}(\cdot) = J_{ji}(\cdot) \). In order to establish the applicability of the above system as a content addressable memory (CAM) or associative memory, we need to prove the following properties:

1. The system is globally stable, and that it does not have any limit cycles. This will ensure that every trajectory of the system will converge to one of its equilibrium points.

2. The system has the capability of possessing multiple stable equilibrium points. Each stable equilibrium point can then be considered as one of the stored 'patterns' of the system.

3. Sufficient conditions to guarantee stability of individual equilibrium points. By meeting these conditions, one can in principle 'program' a desired set of stored patterns of the system.

In the rest of this section, we prove results showing each of the above properties for the system in equations 4. Global stability of a system is usually established by deriving an appropriate Liapunov function. An explicit global Liapunov function is often difficult to specify. However, exploiting the special structure of the system in equations 4, a global Liapunov function can be given as,

\[
L(V) = - \sum_{i=1}^{n} \int_{v_i}^{v_i} [I_i - J_{i*}(\psi)] d\psi + \sum_{i=1}^{n} \sum_{j=i+1}^{n} \int_{v_i - v_j}^{v_i - v_j} J_{ij}(\psi) d\psi.
\]

where \( V = (v_1, \ldots, v_N) \) is the vector of island potentials, and \( L(V) \) is a global Liapunov function, because \( dL(V)/dt \leq 0 \), along every trajectory:

\[
\frac{dL(V)}{dt} = - \sum_{i=1}^{n} v_i [I_i - J_{i*}(v_i)] + \sum_{i=1}^{n} \sum_{j=i+1}^{n} (v_i - v_j) J_{ij}(v_i - v_j)
\]

\[
= - \sum_{i=1}^{n} \left[ I_i - J_{i*}(v_i) + \sum_{j < i} J_{ij}(v_j - v_i) - \sum_{j > i} J_{ij}(v_i - v_j) \right] \frac{dv_i}{dt}
\]

\[
\leq \sum_{i=1}^{n} C_i \left( \frac{dv_i}{dt} \right)^2
\]

Note that \( dL(V)/dt < 0 \) everywhere except at the equilibrium points.

We next address the introduction of multiple equilibrium points. The following theorem shows that our system will exhibit multiple equilibrium points, only if non-monotonic nonlinear functions are allowed in equation 4. We note that this property distinguishes our system from standard neuromorphic systems studied in the literature [16, 23, 26], where it has been shown that strictly monotonic nonlinear activation functions are sufficient to yield multiple stable equilibrium points.

**Theorem 1** If \( J_{i*}(\cdot) \), for all \( i = 1, \ldots, n \), and \( J_{ij}(\cdot) \), for all \( i, j = 1, \ldots, n \), are strictly monotonic functions then the system described in equation 4 has only one equilibrium point.
Proof: We provide a proof by contradiction: we first assume that the system has at least two distinct equilibrium points \( V = (v_1, \ldots, v_n) \) and \( V^* = (v_1^*, \ldots, v_n^*) \), and then derive a contradiction from this assumption.

Without loss of generality, the index set \( S = \{1, \ldots, n\} \) can be decomposed as \( S = S_1 \cup S_1^c \), where \( S_1 = \{1, \ldots, k\} \), \( S_1^c = \{k+1, \ldots, n\} \), \( v_i \geq v_i^* \), for all \( i \in S_1 \), and for all \( j \in S_1^c \), \( v_j < v_j^* \). We first consider the case where \( S_1^c \) is not empty. Since \( V \) and \( V^* \) are both equilibrium points of equation \( 4 \), we have

\[
I_i = J_{is}(v_i) - \sum_{j<i} J_{ij}(v_j - v_i) + \sum_{j>i} J_{ij}(v_i - v_j)
= J_{is}(v_i^*) - \sum_{j<i} J_{ij}(v_j^* - v_i^*) + \sum_{j>i} J_{ij}(v_i^* - v_j^*),
\]
for all \( i = 1, \ldots, n \). Equivalently,

\[
J_{is}(v_i) - J_{is}(v_i^*) = \sum_{j<i} [J_{ij}(v_j - v_i) - J_{ij}(v_j^* - v_i^*)] - \sum_{j>i} [J_{ij}(v_i - v_j) - J_{ij}(v_i^* - v_j^*)]
\]
for all \( i = 1, \ldots, n \). Summing the above equation over the index sets \( S_1 \) and \( S_1^c \), we get

\[
\sum_{i=1}^k [J_{is}(v_i) - J_{is}(v_i^*)] = -\sum_{i=1}^k \sum_{j=k+1}^n [J_{ij}(v_i - v_j) - J_{ij}(v_i^* - v_j^*)],
\]
and

\[
\sum_{j=k+1}^n [J_{js}(v_j) - J_{js}(v_j^*)] = \sum_{i=1}^k \sum_{j=k+1}^n [J_{ij}(v_i - v_j) - J_{ij}(v_i^* - v_j^*)].
\]
Since \( v_i \geq v_f \), for \( 1 \leq i \leq k \), and \( v_j < v_j^* \), for \( (k+1) \leq j \leq n \), it follows that \( (v_i - v_j) > (v_f - v_j^*) \). Moreover, since \( J_{ij}(\cdot) \) is a strictly monotonic function, it follows that \( [J_{ij}(v_i - v_j) - J_{ij}(v_i^* - v_j^*)] > 0 \); hence

\[
\sum_{i=1}^k \sum_{j=k+1}^n [J_{ij}(v_i - v_j) - J_{ij}(v_i^* - v_j^*)] > 0.
\]
Combining equations 7, 8, and 9, we get

\[
\sum_{i=1}^k [J_{is}(v_i) - J_{is}(v_i^*)] < 0 < \sum_{j=k+1}^n [J_{js}(v_j) - J_{js}(v_j^*)].
\]
However, since \( v_i \geq v_f \), for \( 1 \leq i \leq k \), and \( v_j < v_j^* \), for \( (k+1) \leq j \leq n \), and \( J(\cdot) \) is a monotonic function, we get the following inequalities which contradict those in equation 10:

\[
\sum_{i=1}^k [J_{is}(v_i) - J_{is}(v_i^*)] \geq 0, \quad \text{and} \quad \sum_{j=k+1}^n [J_{js}(v_j) - J_{js}(v_j^*)] \leq 0.
\]
If \( S_1^c \) is empty, i.e., \( v_i \geq v_f \), for all \( i = 1, \ldots, n \), then equation 7 reduces to:

\[
\sum_{i=1}^n [J_{is}(v_i) - J_{is}(v_i^*)] = 0.
\]
However, since $V \neq V^*$, there exists at least one $k$ such that $v_k > v_k^*$. Moreover, since $J_{is}()$s are strictly monotonic functions, we get the following contradiction:

$$\sum_{i=1}^{n} [J_{is}(v_i) - J_{is}(v_i^*)] > 0.$$  

One can incorporate non-monotonic functions in the system described in equation 4 in several different ways, and for our purposes we find it convenient to consider $J_{is}(\cdot)$s to be non-monotonic functions of the form shown in figure 4(b), while keeping $J_{ij}(\cdot)$s to be monotonic (in fact linear). Such a choice corresponds directly to physically realizable systems in which there is resonant tunneling between each island and the substrate, together with resistive inter-island coupling. In addition it introduces sufficient complexity for the system to exhibit multiple stable equilibrium points.

We next apply the Liapunov method for deriving sufficient conditions for analyzing the stability of individual equilibrium points of the system described in equations 4.

**Theorem 2** A given equilibrium point $V^* = (v_1^*, \ldots, v_n^*)$ of equation 4 is asymptotically stable if $J_{is}(\cdot)$, for all $i = 1, \ldots, n$, and $J_{ij}(\cdot)$, for all $i, j = 1, \ldots, n$ are monotonic in some neighborhood of $V^*$, i.e., 3 an $\epsilon > 0$ such that for all $||V - V^*|| \leq \epsilon, (v_i - v_i^*)[J_{is}(v_i - v_i^*)] \geq 0$ (1 $\leq i \leq n$), and $[(v_i - v_j) - (v_j - v_j^*)][J_{ij}(v_i - v_j) - J_{ij}(v_i^* - v_j^*)] \geq 0$ (1 $\leq i, j \leq n$).

**Proof:** Let us consider the following quadratic Liapunov function:

$$\mathcal{L}(V - V^*) = \sum_{i=1}^{n} C_i (v_i - v_i^*)^2.$$  

Clearly, $\mathcal{C}(V - V^*) \geq 0$; we have to next show that there always exists a neighborhood of $V^*$ in which $d\mathcal{L}(V - V^*)/dt \leq 0$.

$$\frac{d\mathcal{L}(V - V^*)}{dt} = \sum_{i=1}^{n} (v_i - v_i^*) C_i \frac{dv_i}{dt}$$

$$= \sum_{i=1}^{n} (v_i - v_i^*) [I_i - J_{is}(v_i) + \sum_{j<i} J_{ij}(v_j - v_i) - \sum_{j>i} J_{ij}(v_i - v_j)].$$  

Since $V^*$ is an equilibrium point, it follows that for all $i = 1, \ldots, n$,

$$I_i = J_{is}(v_i^*) - \sum_{j<i} J_{ij}(v_j^* - v_j^*) + \sum_{j>i} J_{ij}(v_i^* - v_j^*).$$

Hence,

$$\frac{d\mathcal{L}(V - V^*)}{dt} = -\sum_{i=1}^{n} (v_i - v_i^*) \left\{ [J_{is}(v_i) - J_{is}(v_i^*)] + \sum_{j<i} [J_{ij}(v_j - v_i) - J_{ij}(v_j^* - v_i^*)] 

- \sum_{j>i} [J_{ij}(v_i - v_j) - J_{ij}(v_i^* - v_j^*)] \right\}$$

$$= -\sum_{i=1}^{n} (v_i - v_i^*) [(J_{is}(v_i) - J_{is}(v_i^*))].$$  

13
Thus, \( \frac{d \mathcal{L}(V - V^*)}{dt} \leq 0 \) for all \( ||V - V^*|| \leq \epsilon \).

We next briefly consider the special case where \( J_{ij}(\cdot) \) is restricted to be linear (hence, strictly monotonic): equation 4 then reduces to:

\[
C_i \frac{dv_i}{dt} = h - J_{is}(v_i) + \sum_{j<i} \left( \frac{v_j - v_i}{R_{ij}} \right) - \sum_{j>i} \left( \frac{v_i - v_j}{R_{ij}} \right),
\]

for all \( i = 1, \ldots, n \). Since \( J_{ij}(\cdot) \)'s are already chosen to be strictly monotonic, Theorem 2 then implies the following: a given equilibrium point \( V^* = (v_1^*, \ldots, v_n^*) \) is stable if \( J_{is}(\cdot) \) is monotonic around \( v_i^* \).

The special case described in equation 11 can also be related to conventional models of CAM studied in the neural network literature. A general CAM model studied in [23] is given as:

\[
\frac{dv_i}{dt} = a_i(v_i) \left[ b_i(v_i) - \sum_{j=1}^{n} c_{ij} d_j(v_j) \right],
\]

where the functions and parameters must satisfy the following conditions to ensure global stability:

1. \( a_i(v_i) \geq 0 \),
2. \( c_{ij} = c_{ji} \), and
3. \( d[d_j(\psi)]/d\psi \geq 0 \) (i.e., the non-linear functions \( d_j(\cdot) \)'s are non-decreasing).

The special case of our system as described in equation 11, reduces to an instance of the general model with the following substitutions:

\[
\begin{align*}
& a_i(v_i) = C_i, \\
& b_i(v_i) = I_i - J_{is}(v_i), \\
& c_{ii} = \frac{1}{R_{ii}}, \\
& c_{ij} = -\frac{1}{R_{ij}} = c_{ji},
\end{align*}
\]

and

\( d_i(v_i) = v_i \).

The above substitutions indicate more clearly the points of departure between our system and the standard CAMs:

1. In our case the self signal function \( b_i \) is nonlinear and non-monotonic, whereas in the standard model \( b_i \) is linear.
2. In our case the other signal function $d_j$ is linear, whereas in the standard model $d_j$ is a strictly monotonic nonlinear function.

Despite these differences our special case performs adequately as a CAM, as we shall show through particular examples in section 3.1.3.

### 3.1.2 System with Capacitive Networking

We show here that the qualitative stability analyses carried out above, in the absence of capacitive coupling amongst the islands will remain valid, even when a capacitive network is introduced. In particular, the capacitive network enters the dynamical equations through the charge-voltage relation,

$$q_i = C_i v_i + \sum_{j \neq i} C_{ij} (v_i - v_j)$$  \hspace{1cm} (13)

where $C_i$ is the substrate capacitance, and $C_{ij}$ is the inter-island capacitance. The dynamical equations, can thus be rewritten as,

$$\frac{d}{dt} q_i = \sum_j C_{ij} \frac{d}{dt} v_j = I_i - J_{is}(v_i) + \sum_{j < i} J_{ij}(v_j - v_i) - \sum_{j > i} J_{ij}(v_i - v_j),$$  \hspace{1cm} (14)

where $C_{ii} := (C_i + \sum_{j \neq i} C_{ij})$, $C_{ij} = -C_{ij}$, and the matrix $C$ is positive definite. Let $V = (v_1, \ldots, v_n)$ and $I = (I_1, \ldots, I_n)$, then the above system of equations can be written in vector form as follows,

$$C \frac{dV}{dt} = f(V, I),$$  \hspace{1cm} (15)

where, $f(\cdot)$ is a vector function representing the right-hand side of equations 14. Now, using the above notation, equations 4 can also be written as,

$$\frac{dV}{dt} = f(V, I).$$  \hspace{1cm} (16)

Since equation 15 is obtained from equation 16 by multiplying its left-hand side with a positive definite matrix one can conclude the following:

1. The equilibrium points of equations 4 and 14 are identical.

2. The stability of corresponding equilibrium points of equations 4 and equations 14 are unchanged, i.e., a stable (unstable) equilibrium point of equation 4 will remain a stable (unstable) equilibrium point of equation 14.

3. The global Liapunov function for equation 4, as stated in equation 5, is also a global Liapunov function for equation 14. Thus, the system with capacitive coupling also does not exhibit limit cycles.

In the particular examples considered in section 3.1.3, we shall see that capacitive coupling can at most modify the trajectory of the system in phase space, and affect the retrieval times of different memory states, while leaving the locations of the equilibrium points and their stability properties unchanged. The practical consequence of these results is the observation that capacitive coupling alone cannot be used to generate complex dynamical effects. However, in section 3.2 we will find that when single-electron effects become strong the sizes of the capacitance parameters can crucially alter the nature of the equilibrium points.
3.1.3 Illustrative Examples

The Lyapunov theory presented in section 3.1.1 guarantees the existence of multiple stable equilibrium points if non-monotone nonlinearities are permitted into our network. This analysis, however, does not provide a prescription for programming particular types of stable points, which will in general depend on the current pumps $I_i$, the resistive network $[R_{ij}]$, and the properties of the substrate nonlinearities. This issue can, however, be addressed on a case-by-case basis, as we shall show below. Our first example is illustrated in figure 5(a), in which the two islands are coupled with each other through a linear resistor $R_{12}$ and a capacitor $C$. In addition, we assume that each island is coupled to the substrate through a capacitance $C_0$, and a nonlinear resistor whose characteristics are shown in figure 5(b). Since the results presented in this section are numerical in nature, we have minimized the number of parameters by choosing a piecewise linear function, in which each of the three segments have slopes of the same magnitude $R$:

$$J_s(v) = \begin{cases} 
\frac{v}{R} & v < v_0 \\
\frac{2v_0 - v}{R} & v_0 \leq v \leq 2v_0 \\
\frac{v - 2v_0}{R} & v > 2v_0 
\end{cases} \quad (17)$$

Figure 5: (a) A network of two islands which are laterally coupled via a linear resistance $R_{12}$ and a capacitor $C$. The islands are also coupled vertically to the substrate via a capacitance $C_0$, and a nonlinear resistive element. (b) The substrate nonlinearity is modeled as a piecewise linear function. We show theoretically, in section 3.1.1, that more realistic nonlinearities due to resonant tunneling, for example, will yield qualitatively the same dynamics as generated by this nonlinearity. The three segments of the function are denoted by $b_1, b_2$ and $b_3$.

The equations describing the dynamics of the two node network can be written as,

$$\begin{bmatrix} 
C_0 + C & -C \\
-C & C_0 + C
\end{bmatrix} \begin{bmatrix} \dot{v}_1 \\
\dot{v}_2
\end{bmatrix} = \frac{1}{R_{12}} \begin{bmatrix} 
-1 & 1 \\
1 & -1
\end{bmatrix} \begin{bmatrix} v_1 \\
v_2
\end{bmatrix} - \begin{bmatrix} 
J_s(v_1) - I_0 \\
J_s(v_2) - I_0
\end{bmatrix} \quad (18)$$

where we have assumed that both nodes are driven by the same current $I_0$. The equilibrium points of this system can be found by setting $\dot{v}_1 = \dot{v}_2 = 0$, and the stability properties of those equilibrium points can be ascertained by examining the eigenvalues of the system matrix near the equilibrium points [28, 2].

Provided that the driving current $I_0 < v_0/R$, each island can in principle be on any one of the three branches of the nonlinear function $J_s(v)$. Now from theorem 2, we know that the system will
be unstable if either island is operated on branch $b_2$. So, it follows that this system can have at most four different globally stable points, since each island is restricted to being on either branch $b_1$ or $b_3$. Two of these stable points are trivial ones corresponding to both islands operating on the same branch: either $b_1$, or $b_3$. Since we are assuming that both islands are driven by the same current, it follows that both of these stable points will correspond to no current flow in the resistive link $R_{12}$. It is desirable now to determine the conditions on $R_{12}$, R and $I_0$, which will allow one to program the other two stable points, corresponding to one island on $b_1$ and the other on $b_3$. This is readily accomplished by taking island 1 on $b_1$, island 2 on $b_3$, solving equations 18 in steady-state, and by finally imposing the consistency check that $v_1 < v_0$ and $v_2 > 2v_0$. The result of this calculation is:

$$f \equiv \frac{R_{12}}{R} > 2 \quad \text{and} \quad \left(\frac{2}{f+2}\right)\left(\frac{v_0}{R}\right) < I_0 < \left(\frac{f}{f+2}\right)\left(\frac{v_0}{R}\right)$$

(19)

In figure 6 we have taken $f = 5$, and $I_0 = v_0/2R$ consistent with the above constraints, and have plotted the phase diagram corresponding to the dynamics of the network. Figure 6(b) differs from 6(a), only due to the presence of a non-vanishing coupling capacitance $C$ between the islands. As was proved in section 3.1.2, the capacitive coupling does not alter the location, nor the stability properties of the equilibrium points. It only alters the basins of attraction, phase trajectories and the retrieval times.

As an additional example illustrating the possibility of programming particular stable points, we consider a network of three islands. In figure 7 we have shown three-dimensional phase diagrams which demonstrate the possibility of altering the number, and locations of stable equilibrium points when the resistance parameters are chosen appropriately. In these examples all islands of the network have been assumed to be driven by the same current $I_0$. While spatially inhomogeneous current pumps will improve the flexibility with which the stable points can be programmed, uniform
Figure 7: This figure indicates the possibility of programming a network by choosing, resistance parameters appropriately, for the case of a three island system. Only substrate coupling $C_0$ is permitted. The stable equilibrium points are indicated by the shaded circles. (a) eight states, with $R_{12} = R_{23} = R_{13} = 10R$, (b) four states, with $R_{12} = R_{23} = 10R$, and $R_{13} = 5R$. (c) two states, with $R_{12} = R_{23} = R_{13} = 5R$.

Pumping currents may alleviate, somewhat, the practical difficulties associated with contacting a large number of islands.

A potential application of the networks discussed in this section is suggested by the two examples shown in figure 8. The network is assumed to be near-neighbor coupled, and is assumed to be pumped with the same current $I_0 = v_0/2R$, at all islands. The islands are color coded, based on their potential $v_i$ as follows: (i) white = $v_i < v_0$, (ii) cross-hatched = $v_0 < v_i < 2v_0$, and (iii) black = $v_i > 2v_0$. The network in example (a) is characterized by a single inter-island resistance $R_{i:1,i:1} = 5R$, and it is initialized with predominantly black islands on the left half of the plane,
Figure 8: Rudimentary image processing capability in a near-neighbor connected network of 10 x 10 islands, which are all pumped by the same current $I_0 = \frac{v_0}{2R}$. Each island is colored in accordance with the particular branch $b_1, b_2, \text{ or } b_3$, in figure 5(b) - of the substrate nonlinearity, the island potential lies on. (a) The resistive network is uniform $R_{l,j} = 6R$. The input contains domains which are either predominantly black, or white, and the resulting output recovers domains which are either all black, or all white. (b) The network is partitioned into three concentric regions. The islands in the outermost and innermost regions are coupled with low resistances $R_{l,j} = 3R$, and the islands in the intermediate region are coupled with high resistances $R_{l,j} = 13R$. Parts of the network with low resistance produce regions which are either all white, or all black depending on which was predominant in the initial state.

and predominantly white islands on the right of the plane. The final stable state corresponds to the entire left half settling on black, and the right half settling on white. The network in example (b) is characterized by two separate inter-island resistance parameters: low values in the outer annulus, and the inner square, and high values in the intermediate annulus. The outer annulus settles into all white which was the predominant color of the region initially, while the inner square stabilizes to all black, which again was the predominant color of that region in the initial state. We emphasize that these elementary computational characteristics were obtained with only near-
neighbor connectivity, and a single pumping current. This can in fact be implemented in an array of islands, with near-neighbor conduction arising from the tunneling of electrons. In particular, we show in section 3.2 that as long as the islands are not too small, mappings of one state to another, such as illustrated in figure 8, will remain valid even in the presence of the Coulomb blockade.

3.1.4 Technological Issues

We consider briefly some of the technological issues which will be encountered if an experimental realization of this network is attempted. The non-monotonic substrate nonlinearity can be arranged by growing the islands on top of a heterostructure which can provide a resonant tunneling path to ground. More elaborate nonlinearities, with multiple peaks and valleys can also be introduced if a vertical superlattice sustaining several conduction subbands, is grown beneath the array of islands.

The principal practical difficulty will arise from our assumption that all islands are externally pumped, and from the need to probe all of the islands. There is reason to believe, however, that in particular applications it will be sufficient to pump and probe only a subset of all of the islands. However, we note that this system cannot in general be edge-driven. This can be seen by noting that the effects of a particular current pump will be felt only over a set of islands which are within a characteristic length of the pumped island, since the current will leak into the substrate exponentially rapidly as one moves away from the pumped island. In view of this fact, it is impossible to drive large arrays within an edge-driven picture. However, if non-monotone nonlinearities can be introduced between islands, by using molecular wires, for example, we can eliminate substrate currents, and return to the edge-driven array concept. A detailed numerical investigation of this issue is planned for future work, and will be briefly discussed again in section 3.5.1.

3.2 Limits to Continuous Models : Single-Electron Effects

We will consider here the same network as in section 3.1, with the assumption that single electron effects have become pertinent either due to the lowering of temperature, or due to the physical scaling of the metallic islands down to \( d \approx 10 \text{nm} \). Single-electron effects will become relevant, when the change in potential \( \delta V = q/C \), associated with the addition of a single charge \( q \) to an island, becomes comparable to \( kT/q \), the thermal potential. This condition can be met even at room temperature if islands with an effective capacitance smaller than \( C \approx 5 \times 10^{-18} \text{F} \) are fabricated. Experimental evidence exists, indicating the possibility that a range of single-electron systems small enough to be operated at room temperature can at present be fabricated \([69, 51, 56]\).

In our approach single-electron effects are investigated with the following objectives in mind:

1. To determine if single-electron effects will fundamentally alter the collective behavior deduced from classical nonlinear circuit theory.

2. To determine if the single-electron effect by itself can generate novel, and potentially useful collective behavior.

In this section we shall be addressing the first objective, while in sections 3.3 and 3.4 we will study examples which address the second objective.

In the following we will review basic concepts in the modeling of single-electron phenomena, and then describe a numerical simulator based on a monte carlo technique which captures the stochastic dynamics of single electrons in detail. We will then use this numerical simulator to investigate single-electron effects in the nonlinear networks studied in section 3.1. In particular, we show through specific examples that there exists a limit to the scaling of islands below which
the design features deduced in section 3.1 within a continuous-charge picture, will be overtaken by effects stemming from the discreteness of electronic charge. In the continuous-charge picture the actual size of the effective capacitance of each island determines only the rapidity with which each stable point is reached, and does not affect the system behavior in a qualitative way. In contrast, in the single-electron picture, the actual size of the capacitances can alter the global stability properties of the system.

3.2.1 Modeling Single-Electron Effects

We wish to study the network shown in figure 4(a), within a single-electron picture, and draw comparisons with our classical circuit theoretic results of section 3.1. In order to make the single-electron description precise, we need to replace the inter-island resistive links with tunnel barriers. Throughout this section we shall assume the applicability of the 'orthodox' theory, which places an absolute lower bound on the resistance parameters $R > \frac{\hbar}{4e^2}$, which ensures that an electron will be localized on one or the other of the two electrodes forming a tunnel junction. However, we shall retain our assumption that the resistance across these links is constant, which is in fact consistent with common practice in the modeling of single-electron effects since only very small voltages are typically dropped across the tunnel barriers. Our use of a nonlinear function to describe transport between each island and the substrate, however, is less common in the literature. Recently, though, it has been shown that novel electric field quantization effects can prevail in slim semiconductor superlattices when single-electron effects coexist with non-monotonic nonlinear tunnel rates between wells [36]. In the following we will use, nonlinearities of the kind shown in figure 4(b) together with equation 2 to account for the discrete tunneling of electrons into the substrate. As was suggested previously the substrate nonlinearity can arise from the resonant tunneling of electrons through a quantum well formed in the heterolayers of the substrate. In previous studies of quantum dots [4], single-electron effects and resonant tunneling were allowed to coexist in a very different capacity from that required in our case. In particular, in quantum dots the small capacitance region, in fact harbors the discrete energy level through which electrons need to resonantly tunnel to establish conduction. In our case, the metallic island though small enough to be subject to the Coulomb blockade, is assumed to be large enough to be treated as an electrode from which electrons are resonantly forced through a semiconductor quantum well of large capacitance. The charging of the quantum well itself will be assumed not to be subject to the Coulomb blockade.

A strategy for analyzing this network might consist in writing a Kirchoff current balance condition for each island $i$, by analogy with equation 3,

$$\frac{d}{dt}q_i = -(dq_i/dt)_{\text{sub}} - \sum_{j \in \text{n.n.}} (dq_i/dt)_j + I_i$$

(20)

where the differential terms on the right hand side denote the discrete transfer of electrons to the substrate, and to the nearest neighbors. This equation is only symbolic, and cannot be treated as a proper differential equation since the island charges belong to the discrete set,

$$q_i(t) = \int_0^t I_i(\tau) d\tau + nq \text{ where, } n \in \{\pm 1, 52, ...\}$$

(21)

where we have made explicit the fact that the continuous delivery of charge by a current source is punctuated by discrete charge transfer through tunnel junctions. Single electron tunneling is a stochastic process which can be described rigorously in terms of a master equation, written for the
probability \( p(Q, t) \) that the array is in the charge state \( Q \) at the time \( t \) \cite{21, 20},

\[
\frac{\partial}{\partial t} + \sum_{i=1}^{N} I_{i} \frac{\partial}{\partial q_{i}} \] \( p(Q, t) = \sum_{m} \Gamma_{m}(Q - q_{m})p(Q - q_{m}, t) - p(Q, t) \sum_{n} \Gamma_{n}(Q, t) \)

\( (22) \)

where, the two summations on the right-hand side take stock of the single particle tunnel events which will either bring, or take away the system from the charge state \( Q \). This master equation \cite{5} will require a numerical approach due to the lack of a regular method to solve systems of this kind which can have a very large state space even for a very small number of islands, as can be seen from equation 21. The best procedure then is to develop a monte carlo technique which mimics in detail the physics of single-electronics.

3.2.2 Monte Carlo Simulation Technique

In this section we will outline a monte carlo simulation technique \cite{6, 21}, for the simulation of a current biased network of islands, shown in figure 4(a). The state of the system of islands is fully described at time \( t \) by the vector of island charges \( Q(t) \). Over a very small time interval \( \delta t \), each island \( i \) is assumed to be delivered with an increment of charge \( \delta q_{i} = I_{i} \delta t \) by a current source \( I_{i} \). The new charge state then is \( Q(t + \delta t) = Q(t) + \delta t I_{i} \). Next, the entire system is swept, and a vector of numbers \( \{ \Delta E^{m} \} \) corresponding to the energy dissipated as a result of tunnel events (indexed by the superscript \( m \)), between each island and its nearest neighbors, as well as between each island and the substrate are accumulated. From the vector of dissipated energies, a vector of tunnel rates \( \{ \Gamma_{m} \} \) is then generated using equation 2. A further vector of cumulative tunnel rates is then generated by replacing each element of the vector \( \{ \Gamma_{m} \} \) with the sum of all previous rates:

\[
S_{i+\delta t}^{n} = \sum_{n=1}^{n=m-1} \Gamma_{i+\delta t}^{n} \]

\( (23) \)

The probability that any one of these tunnel events should proceed is then determined by calculating,

\[
P(t + \delta t) = e^{-\delta t S_{i+\delta t}^{l}} \]

\( (24) \)

where the total tunnel rate \( S_{i+\delta t}^{l} \) is the last entry in the vector of cumulative rates. A random number \( r_{i} \), distributed uniformly on the unit interval, is now drawn, and a decision to carry out a tunnel event is then made if the condition \( P < r_{i} \) is met. This prescription conforms to the intuitive idea that the larger the total tunnel rate, the greater the frequency with which the events are carried out. If a decision to carry out a tunnel event is made, then a particular event needs to be selected from among all the events which were considered. This selection is made with probability proportional to the individual tunnel rates. This is numerically implemented by picking again a random number \( r_{\phi} \) distributed uniformly on the unit interval, and then selecting the tunnel event with the lowest index \( i \), which meets the condition \( (S_{i+\delta t}^{l}/S_{i+\delta t}^{l}) > r_{\phi} \). The above procedure is then repeated until convergence, or until adequate numerical evidence has been accumulated. It is evident from the above description that the demands on computer time can become quite considerable even for systems with only a small number of islands. This computational challenge is in general met by implementing this algorithm on several different parallel computing platforms.

3.2.3 Illustrative Examples

We will again consider the two island example shown in figure 5, with a linear tunnel rate between islands and a piecewise linear tunnel rate between each island and the substrate. The single-electron
Figure 9: Phase portrait for single-electron stochastic dynamics, with $R_{12} = 5R$, and $I_0 = v_0/2R$ which are the same as in figure 6. (a) with $\eta = q/C_0 v_0 = 10^{-3}$, $kT = 0$, (b) with $\eta = 10^{-3}$, $C_0 = C$ and $kT = 0$, (c) with $\eta = 10^{-2}$ and $kT = 0$, and (d) with $\eta = 10^{-3}$, $kT = 1$.

Phase diagrams shown in figure 9 use the same resistive parameter choice $R_{12} = 512$, and bias current $I_0 = v_0/2R$ as used for the continuous-charge phase diagrams in figure 6. These diagrams were generated using the Monte Carlo algorithm described in section 3.2.2. The parameter which dictates the importance of single-electron effects to the dynamics illustrated by the phase diagrams is the granularity $7 = q/C_0 v_0$, where $C_0$ is the substrate capacitance, and $v_0$ is the location of the current peak in the substrate nonlinearity shown in figure 5(b). For large values of the capacitance, the granularity $7$ will be small, and the single-electron effect will only be reflected marginally in the phase diagrams. Figures 9(a)&(b) show phase diagrams corresponding to $\eta = 10^{-3}$, obtained at the absolute zero of temperature. Upon comparing these two diagrams with those shown in figure
we find that the discrete stochastic dynamics is evident in these two diagrams only in the noisy aspect of the trajectories.

However, a closer examination of the trajectories near the stable points reveals that these trajectories do not terminate at the stable points, but rather circulate in the neighborhood of the stable points. This aspect of the discrete dynamics is made more clear in figure 9(c), in which a higher granularity parameter $\eta = 10^{-2}$, has been used. In effect each of the stable points calculated within the continuous charge model is now replaced with a limit cycle whose radius is inversely related to the effective capacitance of each of the islands. Now, similar behavior is known in the signal processing literature in the context of digital filters, where quantization errors introduced by finite-precision arithmetic lead to limit cycle oscillations in otherwise stable filters [50]. Furthermore, the single-electron tunneling oscillations which have received much experimental and theoretical analysis [18], are also a manifestation of this effect. In particular, when a junction with an effective resistance $R$ is biased with a continuous current $I_0$, the potential dropped across the junction will converge to a stable value at $I_0R$, within a continuous charge picture. However, when single electron effects are present the potential dropped on the junction will oscillate with amplitude $q/C$ about the stable point $I_0R$, with frequency $I_0/q$. The phase diagram in figure 9(c) is an additional manifestation of this effect, which one may use to derive applications. A more immediate consequence of this effect is the fact that for larger values of $\eta$, the limit cycle will grow larger, and will eventually exceed the size of the basins of attraction. This results in the system having a finite probability of exiting each of the four basins of attraction and entering others. In this strongly single-electronic regime the Liapunov theory presented in section 3.1.1 will lose its applicability, and alternative analytical techniques may need to be introduced. Approximate characterization of this behavior can be obtained from a differential equation formulation of single-electronics [10], which despite some fundamental shortcomings [5], offers adequate qualitative explanations.

3.2.4 Technological Issues

The systems considered in this section are broadly similar to those studied in section 3.1. Consequently, the practical issues pointed out in section 3.1.4 are equally pertinent to the networks discussed in this section. If it is desired to bring out the single-electron effects in a strong way it is essential to isolate the islands from external electrodes in order to ensure that each island has a low effective capacitance. This can be arranged for example by attaching a chain of islands laterally to each island shown in the figure, and then current biasing the tips of each chain. This kind of experimental arrangement has been used to facilitate the observability of single-electron-tunneling oscillations [18]. An alternative method might be to bias the network with the tips of scanning tunneling microscopes.

3.3 Ising-type Single-Electron Networks

In section 3.1 we introduced a nanoelectronic neuromorphic model in which the charge was treated as a continuous variable. The consequences of discrete electron dynamics on the associative memory effects in these networks, were then examined in section 3.2. We showed, in particular, that there exists a strongly single-electronic regime in which the collective properties deduced from a continuous charge model will be qualitatively undermined. In this section we will approach single-electronics in a more direct way, and will establish a correspondence between the dynamics of single electrons in networks of islands, and well-known models of discrete neural networks. In particular, we will show that the evolution of an initial charge distribution towards a stable final equilibrium distribution can be given a neuromorphic interpretation. These properties emerge purely as a
result of the discreteness of electronic charge, and the only assumption we make in establishing this connection is that the inter-island capacitances can be chosen arbitrarily. In view of this latter assumption this system cannot at present be considered a viable technological option. However, we believe that our rigorous results are indicative of the potential of single-electronics.

3.3.1 A Single-Electron Computational Module

![Diagram of a single-electron computational module](image)

Figure 10: Schematic of system which exhibits associative memory effects. Electron tunnel events are permitted between each island and the floating plate, and inter-island tunnel events are disallowed. Capacitive coupling is allowed between all islands, the floating plate and a grounded plate. We show that it is possible to choose capacitance parameters consistent with electrostatics to achieve associative memory effects.

We consider the network of metallic islands shown in figure 10, which can in general be modelled in terms of a set of capacitance parameters, and a set conductance parameters. We define two different types of capacitance parameters: (1) $C_{ij} = \text{coupling between islands } i \text{ and } j$, including the floating plate, and (2) $C_i = \text{coupling between island } i \text{ and the grounded plate}$. In this section we shall disallow lateral conduction between adjacent islands, and consider only systems in which electron transfer is permitted between any given island and the floating plate, as shown in figure 10. An algorithm first developed for lateral transport studies [7, 6] can be readily adapted to the description of electron dynamics in our system. A tunnel event between island $i$ and the floating plate will change the charge $q_i$ on that island by the discrete amount $q$, the charge of a single electron. These tunnel events are chosen in such a way as to lower a free energy which can be written compactly as,

$$E(Q) = \frac{1}{2} Q^T C^{-1} Q - \sum_{p=1}^{P} q_p^s v_p^s$$

(25)

where the first term is the electrostatic energy stored in all of the capacitances in the network, and the second term is the work done by the external voltage sources, each of which has injected a net charge of $q^s_p$ at the potential $v^s_p$. Denoting the vector of island charges and voltages respectively by $Q = [q_1, \cdots, q_N]$, and $V = [v_1, \cdots, v_N]$, the capacitance matrix $C$ in equation 25 can be shown to satisfy the charge-voltage relation $Q = CV$, which can also be written as,

$$q_i = C_i v_i + \sum_{i \neq j} C_{ij}(v_i - v_j)$$

(26)
Table 2: A comparison between the properties of the single-electron module in figure 10 and those of the discrete Hopfield network.

<table>
<thead>
<tr>
<th>Single Electron Network</th>
<th>Hopfield Network</th>
</tr>
</thead>
<tbody>
<tr>
<td>minimize energy function</td>
<td>minimize energy function</td>
</tr>
<tr>
<td>( E = \frac{1}{2} Q^T W Q )</td>
<td>( H = -\frac{1}{2} S^T W S )</td>
</tr>
<tr>
<td>charges of opposite sign align themselves</td>
<td>like spins align themselves</td>
</tr>
<tr>
<td>the inverse of the synaptic matrix is restricted to ( M )-matrices</td>
<td>synaptic matrix can be arbitrary</td>
</tr>
<tr>
<td>the synaptic matrix has to be found directly from the energy function</td>
<td>matrix can be found from desired stable states using Hebb's rule</td>
</tr>
<tr>
<td>electron number ( n_i = q_i/q ) can assume any integer value</td>
<td>spin ( S_i ) can assume only two values: (+1, -1).</td>
</tr>
<tr>
<td>update rule is global: find tunnel rates at all nodes, and select one</td>
<td>update rule is local: select a node and carry out transition if threshold condition is met</td>
</tr>
</tbody>
</table>

From which it follows that the off-diagonal entries \((i \neq j)\) of \( C \) are \( C_{ij} = -C_{ji} \), and the diagonal entries are \( C_{ii} = (C_{1} + \sum_{j \neq i} C_{ij}) \).

Now, using equations 25 and 26, we can derive a simple relation for the change in energy when an electron is exchanged between island \( i \) and the floating gate \( f \). Denoting the potential before and after the tunnel by superscripts \( b \) and \( a \), respectively, we can write the energy loss when an electron tunnels from island \( i \) to the floating plate as,

\[
E^b - E^a = \Delta E_i = \frac{q}{2} \left\{ (v^b_i + v^a_i) - (v^b_f + v^a_f) \right\}
\]  

(27)

An electron tunneling from the floating plate into the island will yield the same equation with the opposite sign. The rate at which the electron tunnel event occurs can now be written in terms of \( AE \); as,

\[
\Gamma_i = \frac{J_{if}(\Delta E_i/q)}{q \left( 1 - \exp(-\Delta E_i/kT) \right)}
\]  

(28)

where \( J_{if}(V) \) is the current voltage characteristics of the junction between island \( i \) and the plate in the absence of Coulomb blockade. In usual studies of Coulomb blockade the tunneling current is assumed to grow linearly with \( V \), which is valid for very small biases. For larger biases, or in the presence of discrete energy levels in the islands it is essential to introduce the non-linear \( J_{if}(V) \) characteristics [37, 41]. In this section we shall only be concerned with the local minima of the energy function in equation 25, and the exact forms of the tunnel rates will not influence the locations nor the stability properties of these minima. However, the exact forms of the tunnel rates may influence the basin sizes and the frequency of occurrence of spurious states, which will be investigated in a separate study. Now, a system in a given charge state \( Q \) can lower its energy by undertaking several different tunnel events. At any given time a particular tunnel event is chosen with probability proportional to the tunnel rates using a monte-carlo prescription [6].

The evident similarity between the above dynamics and that of the well-known neural network model of Hopfield [26, 24] is tabulated in 2. In Hopfield networks the metastable spin states \( S \) which minimize an energy function \( H(S) \), are treated as memories which can be retrieved with high probability, when the network is presented with partially corrupt input patterns. This function
is achieved, due to the fact that input patterns which do not exactly match a stored metastable pattern, will be characterized by a non-minimal energy and will therefore evolve towards the closest minimum. The actual performance of these networks, however, depends crucially on the properties of a weight or synaptic matrix $W$, which in our case is the inverse capacitance matrix $W = C^{-1}$. In Hopfield networks the entries in the synaptic matrix can be chosen arbitrarily, and can in principle be determined to suit a set of metastable states one intends to store in the network. The capacitance matrix of the single-electron module in figure 10, on the other hand, is limited to a narrow class of matrices due to the basic rules of electrostatics:

- The diagonal entries,
  \[(C)_{ij} = C_i + \sum_{j=1}^{n} C_{ij},\]
  are positive, and off-diagonal elements, $-C_{ij}$, are all non-positive. Such matrices are referred to as $Z$ matrices in the literature [46].
- Sum of the elements of each row of $C$ equals $C_i$, i.e.,
  \[C1 = [C_1 \cdots C_n]^T \geq 0,\]
  where $1 = [1 \cdots 1]^T$. This property also implies that $C$ has positive principal minors [46]. Such a matrix with non-positive off-diagonal elements and positive principal minors is referred to as an $M$-matrix.

We also note that any $M$-matrix that satisfies the second property corresponds to a valid capacitance matrix for the single electron computing module. In the following we will investigate the possibility of programming this network, subject to the above constraints, to achieve associative memory effects.

3.3.2 Networks With Local Capacitive Coupling

We study here the structure of the synaptic matrix when only local capacitive coupling is considered. The results indicate that local capacitive coupling might lead to low capacity and large number of spurious memories in the network.

If only near neighbor capacitive coupling is considered, then the capacitance matrix for an one-dimensiona1 regular array is a tridiagonal matrix of the following form:

\[
C = \begin{bmatrix}
C_0 + C & -C & 0 & 0 & \cdots & 0 \\
-C & C_0 + 2C & -C & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & 0 & -C & C + C_0
\end{bmatrix},
\]

where $C_0$ is the capacitance to the substrate, and $C$ is the internode capacitance. Let us examine the first column of $W$, which satisfies

\[
\begin{bmatrix}
C_0 + C & -C & 0 & 0 & \cdots & 0 \\
-C & C_0 + 2C & -C & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & 0 & -C & C + C_0
\end{bmatrix} \begin{bmatrix}
W_{11} \\
W_{12} \\
\vdots \\
W_{1n}
\end{bmatrix} = \begin{bmatrix}
1 \\
0 \\
\vdots \\
0
\end{bmatrix}.
\]
The above matrix equation corresponds to a second-order linear difference equation of the following form:

\[ W_{1i} = \delta(W_{1(i-1)} + W_{1(i+1)}), \quad \text{for } i = 2, \ldots, (n - 1), \]

where \( \delta = \frac{C}{C_0 + 2C}, \) and \( W_{10} = 1. \) The solution can be given as

\[ W_{11} = \frac{1}{C_0 + C(1 - \alpha)}, \]

\[ W_{1i} = W_{11} \alpha^{(i-1)}, \quad \text{for } i = 2, \ldots, (n - 1), \]

and

\[ W_{1n} = \frac{W_{11} C}{C_0 + C} \alpha^{(n-2)}, \]

where

\[ \alpha = \frac{1}{2\delta} (1 - \sqrt{1 - 4\delta^2}) < 1. \]

One can carry out a similar analysis for the other columns of \( W, \) and show that the elements in every row/column of \( W \) decrease exponentially with their distance from the diagonal.

In general, if every island is capacitively coupled to \( k \) of its neighbors, then the preceding analysis can again be carried out. It will involve \( k^{th} \) order linear difference equations, whose solutions will determine the entries of \( W. \) One can then use this correspondence to show that the elements in every row/column of \( W \) decrease exponentially with their distance from the diagonal. Hence, the synaptic matrix \( W \) can be modeled essentially as a banded matrix with a width of \( O(\log n). \) That is, all elements in a row/column that are more than \( c \log n \) (where, \( c \) is determined by \( k \)) distance away from the diagonal element equal 0.

One can now utilize the correspondence of the single electron system with Hopfield networks, and deduce its performance as an associative memory. The banded structure of \( W \) corresponds to having a network with \( O(\log n) \) connectivity. That is, in the equivalent network, each node is connected to at most \( O(\log n) \) other nodes. The dependence of storage 'capacity' of Hopfield networks on the degree of connectivity has been studied extensively in the literature [65]. In particular, one of the results states that if each node in the network is connected to at most \( d \) other nodes, then the storage capacity of the network is \( O(d/\log n), \) where \( n \) it; the total number of nodes. For the single electron system, this result implies a constant (i.e., independent of the number of islands) storage capacity.

We note that the 'capacity' of an associative memory system indicates its ability to store any given set of desired patterns. Thus, if a network has a fixed capacity, then it implies that only a constant number of patterns can be chosen arbitrarily and programmed as the stable states of the system.. Moreover, low capacity of a system usually indicates that the system will have a large number of 'spurious' memories [13]. That is, if a desired set of patterns is stored in the system, then it will also introduce a very large number of other stable states in the network.

### 3.3.3 Programming Stable States

We study here ways of storing patterns in the system when the capacitive coupling network is permitted to be nonlocal.

Suppose we have a system of \( n \) dots and one floating plate. Let \( Q = [q_0, q_1, \ldots, q_n] \) be the charge vector of the system, where \( q_0 \) is the charge of the floating plate and \( q_1, \ldots, q_n \) are the charges of
the \( n \) dots. Then the charge vector of the system after a single tunnel event involving the \( i \)th island has occurred can be represented as

\[
Q + (\Delta Q)_i = [q_0 \pm 1 \ q_1 \ \cdots \ q_i \pm 1 \ \cdots \ q_n].
\]

Of course, these tunnel events are energetically favorable only if

\[
\text{Note that the above equation in fact represents two cases corresponding to whether the } i \text{th island either gains or loses an electron.}
\]

If the charge vector \( Q \) is such that \( E(Q + (\Delta Q)_i) - E(Q) > 0 \) for all \( i \), then it is a metastable state of the system. Since all possible transitions from the charge state \( Q \) only increase the energy function, we will also refer to a metastable state as a local minimum of the energy function \( E \).

We can define the problem of storing patterns as follows.

**Problem 1** Given a set of vectors \( M = \{M_1, \ldots, M_T\} \), where \( M_i \in \{0,1\}^{(n+1)} \) and \( \sum_{j=0}^{n} m_{ij} = k \), determine a synaptic matrix \( W \) such that

1. \( M_i \)'s are local minima of the energy function \( E(Q) = Q^T W Q \).
2. \( W^{-1} \) is a valid capacitance matrix, i.e., \( W^{-1} \) is an \( \mathbb{M} \)-matrix and \( W^{-1} 1 \geq 0 \).

Note that in a Hopfield network, it is sufficient to meet only the first set of constraints that ensure the local minima property of the metastable states.

Using the results of section 3.4 we can also show that for the single electron computing module it is sufficient to first determine a matrix \( A \) that satisfies only the local minima constraints. It can then be modified to yield a valid synaptic matrix as stated below.

**Theorem 3** Let \( A \) be an \( (n + 1) \times (n + 1) \) matrix such that a given set of vectors \( M = \{M_1, \ldots, M_T\} \) (where \( M_i \in \{0,1\}^{(n+1)} \) and \( \sum_{j=0}^{n} m_{ij} = k \)) are local minima of the energy function \( Q^T A Q \). Then one can determine constants \( \lambda \) and \( c (\geq 0) \) such that the matrix \( W = A + \lambda I_n + c 11^T \) have the following properties:

1. \( M_i \)'s are local minima of the energy function \( E(Q) = Q^T W Q \).
2. \( W^{-1} \) is a valid capacitance matrix.

We next describe a linear programming formulation for determining a single electron module that stores \( \tau \) memories \( M_1, M_2, \ldots, M_T \), where \( M_i = [0, m_{i1}, \ldots, m_{iN}], 1 \leq i \leq \tau, m_{ij} \in \{0,1\} \), and \( \sum_{j=0}^{n} m_{ij} = k \). Without loss of generality, \( q_0 \) has been chosen to be \( 0 \) in all the desired patterns; the same approach works for arbitrary values of \( q_0 \).
First we need to make all the \( M_i \)'s as metastable states, i.e., all the possible transitions from the state \( M_i \) are unfavorable:

\[
E([0, m_{i1}, \ldots, m_{in}]) - E([1, m_{i1}, \ldots, m_{ij-1}, m_{ij} - 1, m_{ij+1}, \ldots, m_{in}]) \leq -\Delta
\]

\[
E([0, m_{i1}, \ldots, m_{in}]) - E([-1, m_{i1}, \ldots, m_{ij-1}, m_{ij} + 1, m_{ij+1}, \ldots, m_{in}]) \geq -\Delta
\]

where \( A \) is a positive number, \( 1 \leq i \leq \tau \) and \( 1 \leq j \leq n \). If \( E = QA \), then Eqn. (29) shows that each of the constraint is linear in the elements of the matrix \( A \).

We can also add constraints that make sure that every 0-1 charge vector that is within a Hamming distance \( d \) from each stored vector \( M_i \) is not a metastable state of the system. Since the total charge in the system is \( k \), the shortest distance between any two 0-1 charge vectors is 2. The following set of linear constraints ensure that for any \( Q \), if the Hamming distance between \( Q \) and \( M_i \) is 2, then a transition from \( Q \) to \( M_i \) is favorable

\[
E([1, m_{i1}, \ldots, m_{ij}, \ldots, m_{in}]) - E([0, m_{i1}, \ldots, m_{ij}, \ldots, m_{in}]) \geq \Delta
\]

where \( m_{ij} = 1 \).

\[
E([0, m_{i1}, \ldots, m_{ij}, \ldots, m_{in}]) - E([1, m_{i1}, \ldots, m_{ij}, \ldots, m_{in}]) \geq A
\]

where \( m_{ij} = 1 \) and \( m_{il} = 0 \).

Therefore, one has a total of \( 2\tau n + r(k - k + 1) \) linear constraints, when every charge state within Hamming distance of 2 from each \( M_i \) converges to \( M_i \). If one wants to increase the basin of attraction for each \( M_i \) then additional linear constraints can be added. These linear constraints can be solved to yield a feasible matrix, which can then be modified to yield a valid synaptic matrix (as stated in Theorem 3).

Example \((n = 9, k = 4, \tau = 3)\): Let the three memories to be programmed be

\[
M_1 = [0 \ 1 \ 1 \ 1 \ 1 \ 0 \ 0 \ 0 \ 0]^T
\]

\[
M_2 = [0 \ 1 \ 0 \ 1 \ 0 \ 1 \ 0 \ 0 \ 0]^T
\]

\[
M_3 = [0 \ 0 \ 0 \ 0 \ 1 \ 1 \ 1 \ 0 \ 0]^T
\]

To get the \( W \) matrix we first find a feasible solution to a linear programming problem with 126 constraints. By suitably modifying the matrix obtained after solving the linear programming, we get the following valid synaptic matrix

\[
W = \begin{bmatrix}
7.1180 & 1.048 & .010 & 1.054 & 1.038 & 1.080 & 1.064 & 1.000 & 1.000 & 1.051 \\
1.048 & 7.174 & 1.000 & 1.000 & 1.070 & 1.054 & 1.120 & 1.000 & 1.000 & 1.071 \\
1.010 & 1.000 & 7.150 & 1.070 & 1.070 & 1.096 & 1.000 & 1.000 & 1.000 & 1.067 \\
1.054 & 1.000 & 1.070 & 7.114 & 1.060 & 1.000 & 1.054 & 1.000 & 1.000 & 1.041 \\
1.038 & 1.070 & 1.070 & 1.060 & 7.114 & 1.054 & 1.000 & 1.000 & 1.000 & 1.041 \\
1.080 & 1.054 & 1.096 & 1.000 & 1.054 & 7.182 & 1.000 & 1.000 & 1.000 & 1.083 \\
1.064 & 1.120 & 1.000 & 1.054 & 1.000 & 1.000 & 7.174 & 1.000 & 1.000 & 1.071 \\
1.000 & 1.000 & 1.000 & 1.000 & 1.000 & 1.000 & 7.564 & 1.000 & 1.000 & 1.000 \\
1.000 & 1.000 & 1.000 & 1.000 & 1.000 & 1.000 & 1.000 & 7.564 & 1.000 & 1.000 \\
1.051 & 1.071 & 1.067 & 1.041 & 1.041 & 1.083 & 1.071 & 1.000 & 1.000 & 7.038
\end{bmatrix}
\]

Our simulation results show that starting from any 0-1 charge vector \( Q \), such that \( \sum_{i=0}^{N} q_i = 4 \), the single electron module evolves to one of the three programmed memories with high probability.

30
3.4 Single-Electron Networks for Combinatorial Optimization

In this section we will show that the discrete stochastic dynamics of electrons can be viewed as a procedure for simulated annealing. Several problems from combinatorial optimization are then mapped to different single-electron networks. A novel mapping methodology tailored specifically for single-electron networks of the kind shown in figure 10 is developed. The results of this section are predicated on the possibility of arbitrarily choosing the inter-island capacitances. In addition, a particular choice of capacitance parameters will enable the solution of only a particular optimization problem. In view of these limitations these results are to be taken as indicative of the potential of single-electronics, and not as a conclusive proposal to achieve technologically viable computation at present.

3.4.1 A General Methodology

From a computational perspective, the single electron computing module shown in Fig. 10 uses a stochastic simulated annealing algorithm to solve the following constrained quadratic optimization problem:

\[
\text{Minimize } Q^T W Q \text{ such that } \sum_{i=1}^{n} q_i = k, \text{ and } q_i \in \mathcal{Z},
\]

where \( k \) is the initial number of electrons in the array, \( Q = (q_1, \ldots, q_n) \) is the charge vector, and \( \mathcal{Z} \) is the set of integers (allowing the islands to have multiple number of electrons). A more general statement of the underlying optimization problem, when multiple modules of the form shown in Fig. 10 are used, is:

\[
\text{Minimize } Q^T W Q \text{ such that } \sum_{j \in S_i} q_j = k_i, \text{ and } q_i \in \mathcal{Z},
\]

where \( k_i \) is the initial number of electrons in the \( i \)th module, and \( S_i \) comprises the indices of the islands in the \( i \)th module.

It is well known that if a given combinatorial problem can be expressed as a 0–1 (i.e., variables are restricted to be either 0 or 1) quadratic minimization problem with linear constraints, then one can always construct a discrete Hopfield network such that the energy function of the resultant network corresponds to the given quadratic objective function (suitably modified to account for the linear constraints) \[26, 58\]. Since the nodes in a Hopfield network can only assume binary values (i.e., \( \{0,1\} \) or \( \{1,-1\} \)), and the underlying dynamics minimizes its energy function, the derived Hopfield network finds an approximate solution to the given combinatorial problem. An optimum solution is obtained only if the network evolves to a state that achieves the global minimum of its energy function. In the literature, several NP-complete problems have been formulated as 0–1 quadratic minimization problems with linear constraints, thereby demonstrating the potential of Hopfield (and related) networks for performing combinatorial optimization.

However, the dynamics of the single electron array, as stated in Eqn. (31), has a number of built-in restrictions that do not permit a direct adaptation of the methodology used for mapping combinatorial problems onto Hopfield type networks. Some of these restrictions can be enumerated as follows:

1. The discrete valued variables in Eqn. (31) can potentially assume any integer (both positive and negative) value, and are not restricted to be only binary valued. Quadratic optimization
formulations of almost all combinatorial problems, on the other hand, use only binary valued
variables.

2. The cost function is purely quadratic in Eqn. (31), and only a restricted type of linear con-
straints is permitted. The reduction of combinatorial problems usually involves a general
quadratic term, i.e., the objective function has both quadratic as well as linear terms. More-
over, as it will be pointed out later, the linear constraints required for combinatorial problems
are more general than the type permitted in Eqn. (31).

3. In Eqn. (31), the coefficient matrix, W, must satisfy the following:

(a) $W^{-1}$ is an M-matrix, i.e., it has positive diagonal elements and negative off-diagonal
    elements.

(b) $W^{-1}1 \succeq \mathbf{0}$, where $1 = (1 \cdots 1)^T$ (i.e., the all-1 vector), and $\mathbf{0} = (0 \cdots 0)^T$.

The above two conditions ensure that $W^{-1}$ corresponds to a valid capacitance matrix for the
single electron array.

This appears to be a binding restriction, since the coefficient matrix resulting from a combi-
natorial problem is usually tailored to that particular problem, and in general its inverse is
not an M-matrix. On the other hand, the weight matrix of a Hopfield network has very few
restrictions, and can directly accommodate any coefficient matrix required by the applica-
tion. For the single electron array, however, the inverse of the coefficient matrix must always
correspond to a valid capacitance matrix.

In the following we develop a general methodology which accommodates the above restrictions,
and at the same time enables one to map a class of quadratic minimization problems onto single
electron networks. We then provide examples of a number of combinatorial problems which can be
formulated as instances of the class of quadratic minimization problems that we consider.

In particular, we study constrained quadratic minimization problems of the following form:

\[
\text{Minimize } X^T A X \text{ such that } \sum_{j \in S_i} x_j = k_i, \quad x_i \in \{0,1\},
\]

where $A$ is an $n \times n$ real symmetric matrix, $X \in \{0,1\}^n$, $k_i$'s are positive integers, and $S_i$'s partition
the index set $\{1, \ldots, n\}$, i.e., $\cup S_i = \{1, \ldots, n\}$, and $S_i \cap S_j = \emptyset$.

The methodology discussed next reduces the above minimization problem to the following:

\[
\text{Minimize } Y^T (A + \lambda I_{n \times n} + c11^T) Y \text{ such that } \sum_{j \in S_i} y_j = k_i, \quad y_i \in \mathbb{Z},
\]

where $I_{n \times n}$ is the $n \times n$ identity matrix, $1 = (1 \cdots 1)^T$, and $\lambda \geq 0, c \geq 0$ are constants
determined by the reduction process.

The transformation methodology ensures that the two optimization problems described in
Eqns. (32) and (33) are related as follows:

1. If $X = X_0 \in \{0,1\}^n$ is a solution for Eqn. (32), then $Y = X_0$ is also a solution for Eqn. (33).
   Moreover, if $E_0$ is the global minimum of the objective function in Eqn. (32), then the global
   minimum of the objective function in Eqn. (33) equals $E_0 + \lambda (\sum k_i) + c (\sum k_i)^2$.

Note that this condition ensures that even though the variables $y_i$'s in Eqn. (33) can assume
arbitrary integer values, the global minimum is achieved only for $0 - 1$ assignments.
2. The coefficient matrix, \( W = (A + \lambda I_{n \times n} + c11^T) \), of Eqn. (33) corresponds to a valid synaptic matrix for single electron arrays:

- \( W^{-1} \) is an M-matrix, and \( W^{-1}1 \geq 0 \).

The following series of technical results provides rigorous proofs for the above claims.

**Lemma 3.** Given an \( n \times n \) real matrix \( A \) that is positive and symmetric (i.e., \( a_{ij} = a_{ji} > 0 \)), let

\[
m_{\text{max}} = \max_{1 \leq i \leq n} \left( \sum_{j=1}^{n} a_{ij} \right).
\]

Then for any \( \lambda \geq m_{\text{max}} \),

\[
(A + \lambda I_{n \times n})^{-1} 1 \geq 0.
\]

**Proof:** We observe that \((A + \lambda I_{n \times n})^{-1} \geq 0\) if and only if there exists a non-negative vector, \( Y \in \mathbb{R}^n \), such that \((A + \lambda I_{n \times n})Y = 1\). In other words, we want to show that the following Linear Programming (LP) problem has a feasible solution if \( \lambda \geq m_{\text{max}} \):

Maximize \( \mathbf{0}^T \mathbf{Y} \) such that

\[
(A + \lambda I_{n \times n}) \mathbf{Y} = 1 \quad \text{and} \quad \mathbf{Y} \geq 0.
\]

The feasibility of the above LP can be established by checking the boundedness of the cost function of its dual LP, as given below:

Minimize \( \mathbf{1}^T \mathbf{U} \) such that

\[
\mathbf{U}^T(A + \lambda I_{n \times n}) \geq 1.
\]

The duality theorem [52] states that the LP in Eqn. (34) has a feasible solution if and only if its dual (i.e., the LP in Eqn. (35)) has a bounded objective function. Since Eqn. (35) minimizes \( \mathbf{1}^T \mathbf{U} \), it is sufficient to show that \( \mathbf{1}^T \mathbf{U} \geq 0 \). Using the assumptions that \( A \) is a positive matrix and that \( \lambda \geq m_{\text{max}} \), one can indeed show that if \( \mathbf{U}^T(A + \lambda I_{n \times n}) \geq 1 \) then \( \mathbf{1}^T \mathbf{U} \geq 0 \).

Hence, Eqn. (35) has a bounded objective function, and Eqn. (34) has a feasible solution. \( \square \)

**Lemma 2.** Given an \( n \times n \) real matrix \( A \) (elements denoted as \( a_{ij} \)), let

\[
a_{\text{max}} = \max \{ a_{ij} : 1 \leq i, j \leq n \}.
\]

and

\[
a_{\text{min}} = \min \{ a_{ij} : 1 \leq i, j \leq n \}.
\]

Also, let

\[
c = \begin{cases} 
1 + |a_{\text{min}}| & \text{if } a_{\text{min}} < 1, \\
0 & \text{if } a_{\text{min}} \geq 1.
\end{cases}
\]

Then for \( \lambda \geq n(a_{\text{max}} + c)^2 \), the matrix \((A + \lambda I_{n \times n} + c11^T)^{-1}\) is an M-matrix, i.e., its off-diagonal elements are negative and the diagonal elements are positive.

**Proof:** Let \( B = A + c11^T \), then the choice of \( c \) implies that \((B)_{ij} \geq 1\), and that \( \lambda \geq nb_{\text{max}}^2 \), where \( b_{\text{max}} \) is the maximum value of the entries in \( B \). Using a power series expansion we obtain

\[
(A + \lambda I_{n \times n} + c11^T)^{-1} = \frac{1}{\lambda} \left( I_{n \times n} + \frac{1}{\lambda} B \right)^{-1}
\]

\[
= \frac{1}{\lambda} \left( I_{n \times n} - \frac{1}{\lambda} B + \frac{1}{\lambda^2} B^2 - \frac{1}{\lambda^3} B^3 + \frac{1}{\lambda^4} B^4 - \cdots \right).
\]
Since, B is a positive matrix, we obtain that for every \( k \geq 2 \), \( B^k \) is a positive matrix and that 
\[
(B^k)_{ij} \leq n^{k-1}. 
\]
Let 
\[
C = -\frac{1}{\lambda} B + \frac{1}{\lambda^2} B^2 - \frac{1}{\lambda^3} B^3 + \frac{1}{\lambda^4} B^4 - \cdots
\]
It follows then that
\[
-(C)_{ij} = \frac{(B)_{ij}}{\lambda} - \frac{(B^2)_{ij}}{\lambda^2} + \frac{(B^3)_{ij}}{\lambda^3} - \frac{(B^4)_{ij}}{\lambda^4} + \cdots
\]
\[
\geq \frac{(B)_{ij}}{\lambda} - \frac{b_{\text{max}}^2 n}{\lambda^2} - \frac{b_{\text{max}}^3 n^3}{\lambda^4} - \cdots
\]
If \( \lambda \geq n b_{\text{max}}^2 \), then the above series converges, and one gets
\[
1 > (-C_{ij}) > 0
\]
Hence,
\[
(A + \lambda I_{n \times n} + c1 1^T)^{-1} = \frac{1}{\lambda} (I_{n \times n} + C)
\]
has positive diagonal elements and negative off-diagonal elements. Note that the preceding two lemmas together imply that by appropriately choosing \( \lambda \) and \( c \) one can always obtain a matrix \( W = (A + \lambda I_{n \times n} + c1 1^T) \), such that \( W^{-1} \) is an M-matrix, and \( W^{-1} 1 \geq 0 \).

**Lemma 3** Consider the following quadratic minimization problem defined over binary variables:

\[
\text{Minimize } X^T A X \text{ such that } \sum_{j \in S_i} x_j = k_i, \text{ and } x_i \in \{0,1\}, \quad (36)
\]
where \( A \) is an \( n \times n \) real, positive definite, and symmetric matrix, \( k_i \)'s are positive constants, and \( S_i \)'s partition the index set \( (1, \ldots, n) \). Let \( E_0 \) be the minimum value of the objective function.

Next, consider a modified minimization problem in which variables can assume any integer value:

\[
\text{Minimize } Y^T (A + \lambda I_{n \times n}) Y \text{ such that } \sum_{j \in S_i} y_j = k_i, \text{ and } y_i \in \mathbb{Z}. \quad (37)
\]

If \( \lambda > E_0 \), then any optimal solution \( Y_0 \), of the above minimization problem has only binary entries, i.e., \( Y_0 \in \{0,1\}^n \). Moreover, if \( X_0 \in \{0,1\}^n \) is a solution of the 0-1 optimization problem stated in Eqn. (36), then \( Y = X_0 \) is also an optimal solution of the general problem stated in Eqn. (37).

**Proof:** First we observe that for any binary vector \( X \) that satisfies the linear constraints in Eqn. (37) (hence, also in Eqn. (36)),

\[
X^T (\lambda I_{n \times n}) X = \lambda \sum_{i=1}^{n} x_i^2 = \lambda \sum k_i.
\]
However, if a nonbinary integer vector \( Y \) (i.e., \( Y \not\in \{0,1\}^n \)) satisfies the same constraints, then one can verify that

\[
Y^T(\lambda I_{n\times n})Y \geq \lambda(1 + \sum k_i).
\]

Thus, for any nonbinary integer vector \( Y \), we have the following (we also use the fact that \( A \) is positive definite, and \( Y^TAY \geq 0 \)):

\[
Y^T(A + \lambda I_{n\times n})Y = Y^TAY + Y^T(\lambda I_{n\times n})Y \\
\geq \lambda + \lambda \sum k_i.
\]

However, if \( Y = X_0 \), where \( X_0 \) is an optimal solution of Eqn. (36) (hence, \( X_0^TAX_0 = E_0 \)), then

\[
X_0^TA + \lambda I_{n\times n}X_0 = E_0 + \lambda \sum k_i.
\]

Thus, if \( \lambda > E_0 \) then

\[
X_0^T(A + \lambda I_{n\times n})X_0 < Y^T(A + \lambda I_{n\times n})Y
\]

for every nonbinary integer vector \( Y \). Hence, \( Y = X_0 \) is an optimal solution of Eqn. (37) if \( \lambda > E_0 \).

Note that the global minimum of the cost function in Eqn. (37) is \( E_0 + \lambda \sum k_i \).

Combining the results of the above three lemmas, we can state the main result in the form of the following theorem.

**Theorem 4**  Given a \( 0-1 \) optimization problem as stated in Eqn. (32), one can always choose constants \( \lambda \) and \( c \), and transform it into an optimization problem as stated in Eqn. (33) such that

1. If \( X = X_0 \in \{0,1\}^n \) is a solution for Eqn. (32), then \( Y = X_0 \) is also a solution for Eqn. (33). Moreover, if \( E_0 \) is the global minimum of the objective function in Eqn. (32), then the global minimum of the objective function in Eqn. (33) equals \( E_0 + \lambda(\sum k_i) + c(\sum k_i)^2 \).

2. If \( W = (A + \lambda I_{n\times n} + c11^T) \), then \( W^{-1} \) is an M-matrix, and \( W^{-1}1 \geq 0 \).

**3.4.2 Examples**

We next show how different combinatorial problems can be cast as quadratic minimization problems of the form given in Eqn. (32). The results of Theorem 4 can then be used to derive corresponding single electron computing modules that execute stochastic simulated annealing algorithms for solving the related minimization problems.

3.4.2.1 Cliques  Given a graph \( G = (V,E) \), where \( V \) is the set of nodes/vertices and \( E \) is the set of edges, a completely connected subgraph of \( G \) is called a clique. A formal definition can be presented as follows:

**Definition 1** (Clique) A set \( S \) of nodes in a graph forms a clique if there is an edge between every pair of nodes nodes in \( S \). The size of the clique is the number of nodes in \( S \).

The problem of determining a clique of maximum size for a given graph \( G \) is a well known combinatorial problem, and belongs to the class of NP-complete problems [19].

It is customary to express NP-complete problems as yes-no problems [19]:

**Problem 2**  Given a graph \( G = (V,E) \) with \( n \) nodes, and a positive integer \( k (\leq n) \), is there a clique of size \( k \) in \( G \)?
The maximum size of the clique can be determined by asking the preceding yes-no question at most \((\log n)\) times; hence, it is sufficient to solve Problem 2 efficiently. We next show how Problem 2 can be reduced to a quadratic minimization problem of the form Eqn. (32).

First we establish a correspondence between subgraphs of a given graph \(G = (V, E)\), and \(n\)-dimensional binary vectors, \(X = [x_1, \ldots, x_n]\), where \(x_i \in \{0, 1\}\). Given any subgraph comprising a set of nodes \(S \subseteq V = \{v_1, \ldots, v_n\}\), one can define a corresponding binary vector as follows:

\[
x_i = \begin{cases} 
1 & \text{if } v_i \in S, \\
0 & \text{if } v_i \notin S.
\end{cases}
\]

Similarly, given any \(X = [x_1, \ldots, x_n] \in \{0, 1\}^n\), one can define a corresponding subgraph comprising the following set of nodes:

\[
S = \{v_j : x_j = 1 \text{ and } 1 \leq j \leq n\}.
\]

The above correspondence will be used for all the examples in this section, and we shall refer to a subgraph either explicitly by the set of nodes \(S\) it comprises, or equivalently, by the corresponding binary vector \(X \in \{0, 1\}^n\).

Next, let us define an \(n \times n\) incidence matrix \(A\) as follows:

\[
a_{ij} = \begin{cases} 
-1 & \text{if there is an edge in } G \text{ connecting nodes } i \text{ and } j, \\
0 & \text{otherwise}.
\end{cases}
\]

Note that \(A\) is symmetric, and its diagonal elements are all 0.

**Theorem 5** A given graph \(G = (V, E)\) with \(n\) nodes has a clique of size \(k\) \((\leq n)\) if and only if the result of the following minimization problem is \(-k(k - 1)\):

\[
\text{Minimize } X^TAX \text{ such that } \sum_{j=1}^{n} x_j = k, \text{ and } x_i \in \{0, 1\}.
\]

(38)

Moreover, if \(X_0 \in \{0, 1\}^n\) is a solution that achieves the desired minimum \((-k(k - 1))\), then the subgraph corresponding to it is a clique of size \(k\).

**Proof:** Given any \(X \in \{0, 1\}^n\), let \(P_X\) be the set of indices for which \(x_i = 1\), i.e.,

\[
P = \{j : x_j = 1 \text{ and } 1 \leq j \leq n\}.
\]

Then it follows that for any \(X \in \{0, 1\}^n\),

\[
X^TAX = \sum_{i,j \in P_X} a_{ij} = -2 \times \text{[no. of edges in the subgraph defined by } X]\.
\]

In our case, the linear constraint ensures that any feasible solution \(X\) has exactly \(k\) 1's (i.e., \(|P_X| = k\)), and hence the corresponding subgraph has \(k\) nodes. Since a clique of size \(k\) is a completely connected subgraph, it has the maximum number of edges \((= k(k - 1)/2)\) among all the subgraphs with \(k\) nodes. Thus, \(X^TAX\) attains the minimum value of \(-k(k - 1)\) if and only if the corresponding subgraph is a clique of size \(k\).

Note that if there is no clique of size \(k\), then the minimization problem determines a subgraph with the maximum number of edges in it. From the optimal solution vector, one can then easily verify that it does not correspond to a clique. \(\square\)
Since, Eqn. (38) is in the form of Eqn. (32), one can follow the general methodology discussed in the preceding section, and transform it into an equation of the form given in Eqn. (30), which corresponds to the dynamics of a single electron computing module.

The entries of the matrix $A$ in Eqn. (38) are restricted to be either $0$ or $-1$, and for this special case one can show the following:

$$W = A + \lambda I_{n \times n} + c 11^T$$ satisfies the conditions of Theorem 4, if $c = 2$ anil $A = 4n$.

![Graph of 10 nodes](image1.png)

(a) A graph of 10 nodes

![Incidence matrix](image2.png)

(b) The incidence matrix

![W matrix](image3.png)

(c) The $W$ matrix

![C matrix](image4.png)

(d) The $C$ matrix

Figure 11: Steps involved in the derivation of capacitance parameters for a single-electron computational module that determines if there is a clique of a given size in a graph. (a) & (b) A graph comprising ten nodes and its incidence matrix. (c) A synaptic matrix derived from the incidence matrix following the procedure stated in Theorem 4. (d) The capacitance matrix for the single-electron computational module.

Fig. 11 shows the incidence matrix, the $W$ matrix, and the capacitance matrix (i.e., $W^{-1}$) of a corresponding single electron module for a graph with ten nodes. Fig. 12 shows the initial distribution of electrons ($k = 4$) in the single electron array, and also the occupancy when the system has evolved to a minimum-energy configuration. As shown in the figure, the minimum-energy configuration corresponds to a clique of size 4 in the graph. Fig. 13 shows several snapshots of the configuration of electrons as the array evolves from the initial configuration to a minimum-energy configuration. Note that while the minimum-energy configuration has only single occupancy of electrons (as guaranteed by our analysis), the intermediate steps can have multiple electrons on individual islands. This mechanism helps the system to get out of local minima (at finite
Figure 12: (a) & (b) The initial configuration of electrons in the computational module, and the set of nodes it corresponds to in the original graph. *'s indicate the islands occupied by the four electrons. As to be expected, the initial set of nodes do not define a clique. (c) & (d) The minimum-energy configuration of electrons, and the corresponding clique (comprising four nodes) in the original graph.

temperature), and could potentially let the system perform better than conventional simulated annealing.

3.4.2.2 Independent Sets  We next study a combinatorial problem closely related to the problem of determining a clique of maximum size.

Definition 2 (Independent Set) A set $S$ of nodes in a graph is an independent set if no two nodes in $S$ are connected to each other.

The problem of determining an independent set of maximum size can be expressed as a yes-no problem as follows:

Problem 3  Given a graph $G = (V, E)$ with $n$ nodes, and a positive integer $k$ ($k \leq n$), is there an independent set of size $k$ in $G$?

Note again that by asking the above question at most $(\log n)$ times, one can determine the size of the maximum independent set, and hence, it is sufficient to solve Problem 3.
Figure 13: Snapshots of a typical sequence of charge states that the computational module evolves through at a finite temperature. Since, each island in the computational module corresponds to a node in the graph, the islands with non-zero charge are highlighted directly on the graph, and their charges as well as the total energy of the configuration are listed. Note that while the minimum-energy configuration (as illustrated in (e)) corresponds to a charge state in which multiple electrons on the same island are not permitted, intermediate charge states encountered during the dynamics are not so constrained. This, in principle, might allow the system to escape local minima more easily.

We next define an incidence matrix $A$ that is the negation of the matrix defined for solving Problem 1, i.e.,

$$a_{ij} = \begin{cases} 1 & \text{if there is an edge in } G \text{ connecting nodes } i \text{ and } j, \\ 0 & \text{otherwise}. \end{cases}$$
We next state the following theorem which maps Problem 3 to a minimization problem of the form given in Eqn. (32). The proof of the theorem is very similar to that of Theorem 38, and is skipped here.

**Theorem 6** A given graph \( G = (V, E) \) with \( n \) nodes has an independent set of size \( k (\leq n) \) if and only if the result of the following minimization problem is \( 0 \):

\[
\text{Minimize } X^T A X \quad \text{such that} \quad \sum_{j=1}^{n} x_j = k, \quad \text{and} \quad x_i \in \{0, 1\},
\]

(39)

Moreover, if \( X_0 \) is an optimal solution that achieves the desired minimum (\( = 0 \)), then the set of nodes corresponding to \( X_0 \) is an independent set of size \( k \).

**3.4.2.3 Minimum Cut** We consider here graphs in which every edge has a cost function assigned to it.

**Definition 3 (Cut and Capacity)** Let \( G \) be an undirected graph with vertex set \( V = \{v_1, \ldots, v_n\} \). Let \( c_{ij} \) be the cost associated with an edge connecting \( v_i \) and \( v_j \). A cut is a partition \( (S, \overline{S}) \) of the nodes in \( V \) into sets \( S \) and \( \overline{S} \) such that \( S \cup \overline{S} = X \) and \( S \cap \overline{S} = \emptyset \). The *capacity* of the cut \( C(S, \overline{S}) \) is sum of costs on the edges which connect the nodes in \( S \) to the nodes in \( \overline{S} \), i.e.,

\[
\sum_{v_i \in S \land v_j \in \overline{S}} c_{ij}.
\]

Note that our definition allows either \( S \) or \( \overline{S} \) to be empty. In that case the partition will be called a null cut and the associated capacity will be zero.

**Problem 4 (Minimum Cut)** Given a graph \( G = (V, E) \), determine a cut that has the minimum capacity among all other cuts in the graph.

Given any cut \( \{S, \overline{S}\} \), we define \( 2n \) binary variables as follows:

\[
x_{i1} = \begin{cases} 
1 & \text{if node } i \text{ is in } S, \\
0 & \text{if node } i \text{ is in } \overline{S}, 
\end{cases}
\]

and

\[
x_{i2} = 1 - x_{i1}.
\]

Correspondingly, if we define a vector \( X = [x_{11} \cdots x_{n1} \ x_{12} \cdots x_{n2}] \), then one can easily verify that any assignment of \( X \in \{0, 1\}^{2n} \) (i.e., each coordinate of \( X \) is assigned either 0 or 1) with the additional constraint that \( x_{i1} \neq x_{i2} \), corresponds to a cut of \( G \).

In order to set up the minimization problem, we define a modified symmetric cost-function matrix \( D \) as follows:

\[
d_{ij} = \begin{cases} 
A & \text{if } i = j, \\
c_{ij} & \text{otherwise},
\end{cases}
\]

where \( A \) is a large number, e.g., \( A \geq 1 + \sum_{i,j} |c_{ij}| \). Next let us define a \( 2n \times 2n \) symmetric matrix as follows:

\[
A = \begin{bmatrix} 
0_{n \times n} & D \\
D & 0_{n \times n}
\end{bmatrix}
\]

(40)

where \( 0_{n \times n} \) is an \( n \times n \) matrix with all zero elements.
Theorem 7  The minimum capacity of any n-node graph \( G = (V, E) \) is given by the result of the following minimization problem:

\[
\text{Minimize } X^T A X \text{ such that } \sum_{j=1}^{n} x_{j1} + \sum_{j=1}^{n} x_{j2} = n, \text{ and } x_{i1}, x_{i2} \in \{0, 1\}, \tag{41}
\]

where \( A \) is the matrix defined in Eqn. (40) and \( X = [x_{11} \ldots x_{n1} x_{12} \ldots x_{n2}] \in \{0, 1\}^{2n} \). Moreover, if \( X_0 \) is an optimal solution, then a minimum cut of \( G \) is given by the cut represented by \( X_0 \).

Proof: First we observe that if any \( X = [x_{11} \ldots x_{n1} x_{12} \ldots x_{n2}] \in \{0, 1\}^{2n} \) satisfies the constraint \( x_{i1} \neq x_{i2} \) for every \( i = 1, \ldots, n \), then it also satisfies the constraint \( \sum_{j=1}^{n} x_{j1} + \sum_{j=1}^{n} x_{j2} = n \). Thus every cut of \( G \) satisfies the linear constraint in Eqn. (41). Moreover, if \( x_{i1} \neq x_{i2} \) for every \( i \), then \( X^T A X \) equals twice the capacity of the cut defined by \( X \). That is, let \( (S_X, \overline{S}_X) \) be the cut corresponding to \( X \), then

\[
X^T A X = 2 \sum_{i \in S_X, j \in \overline{S}_X} c_{ij}.
\]

Thus minimizing \( X^T A X \) corresponds to determining a cut with the minimum capacity.

Now, the proof is completed by showing that if for any \( i \), \( x_{i1} = x_{i2} = 1 \) in \( X \in \{0, 1\}^{2n} \), and \( \sum_{j=1}^{n} x_{j1} + \sum_{j=1}^{n} x_{j2} = n \), then it cannot be an optimal solution. From the construction of \( A \) it follows that if \( x_{i1} = x_{i2} = 1 \) in \( X \) for any \( i \), then

\[
X^T A X = 2\lambda + r_X,
\]

where \( r_X \) represents the rest of the product terms. Since, \( x_{i1} = x_{i2} = 1 \) and there are exactly \( n \) \( 1 \)'s in \( X \) there must exist a \( l \) such that \( x_{l1} = x_{l2} = 0 \). Now, construct a modified vector \( X' \) which is the same as \( X \) at every coordinate except for the following two: \( x'_{i1} = 0 \) and \( x'_{i2} = 1 \). The cost function corresponding to \( X' \) is then given by

\[
X'^T A X' = r_X + 2(\sum_{k=1, k \neq l}^{n} c_{kl} x_{k1} - \sum_{k=1, k \neq i}^{n} c_{ki} x_{k1}) < r_X + 2\lambda = X^T A X.
\]

Hence, \( X \) is not an optimal solution.

Often it is useful to calculate the minimum capacity under the constraint that the cuts have a specified number of nodes.

**Problem 5  (Constrained Minimum Cut)** Given a graph \( G = (V, E) \) and a positive integer \( k \ (\leq n) \), determine a cut that has the minimum capacity among all cuts \( (S, \overline{S}) \), where \( |S| = k \). \( \square \)

The case where \( k = n/2 \) has been studied extensively in the literature [52, 19].

One can also formulate the Constrained Minimum Cut problem as a quadratic minimization problem of the form given in Eqn. (32), as stated below.
Theorem 8  Given any $k \leq n$, the constrained minimum capacity of $G = (V, E)$ is given by the result of the following minimization problem:

$$\text{Minimize } X^T A X \text{ such that } \sum_{j=1}^{n} x_{j1} = k, \sum_{j=1}^{n} x_{j2} = (n - k), \text{ and } x_{i1}, x_{i2} \in \{0, 1\},$$  \hspace{1cm} (42)

where $A$ is the matrix defined in Eqn. (40) and $X = [x_{11} \ldots x_{n1} x_{12} \ldots x_{n2}] \in \{0, 1\}^{2n}$. Moreover, if $X_0$ is an optimal solution, then a constrained minimum cut of $G$ is given by the cut represented by $X_0$.

Note that in the above formulation of the minimization problem, there are two different linear constraints. Hence, the above formulation corresponds to two single electron computing modules of the kind shown in Fig. 10.

3.4.2.4 Traveling Salesman Problem  In the following we will show how the travelling salesman problem can be reduced to a single-electron network problem.

Definition 4  Traveling Salesman Problem (TSP): Given $n$ cities and the distances $d_{ij}$ (where, $d_{ij} = d_{ji} \geq 0$) between city $c_i$ and city $c_j$, determine a closed tour (i.e., each city is visited once and the tour ends at its starting point) of minimum length.

A tour comprises $n$ stops, where at each stop one of the cities is visited, and every city is visited only once during the complete tour. Thus, one can define $n^2$ binary variables $z_{ij}$ ($1 \leq i, j \leq n$) as follows:

$$z_{ij} = \begin{cases} 1 & \text{if the } j\text{th city is visited at the } i\text{th stop} \\ 0 & \text{otherwise.} \end{cases}$$

In order to ensure a valid tour, $z_{ij}$'s have to satisfy the following $2n$ constraints:

1. Each stop has only one city assigned to it:

$$\sum_{j=1}^{n} z_{ij} = 1 \text{ for all } i = 1, \ldots, n.$$  \hspace{1cm} (43)

2. Each city is assigned to only one stop:

$$\sum_{i=1}^{n} z_{ij} = 1 \text{ for all } j = 1, \ldots, n.$$  \hspace{1cm} (44)

One can now verify that the total length of any tour can be computed as

$$L = \frac{1}{2} \sum_{i \neq j}^{n} \sum_{k=1}^{n} d_{ij} z_{ik}(z_{j,k-1} + z_{j,k+1}),$$  \hspace{1cm} (45)

where $d_{ij}$ is the distance between cities $i$ and $j$, and we let $z_{j,n+1} = z_{j1}$ and $z_{j0} = z_{jn}$. Minimizing $L$ subject to the preceding constraints will yield a tour of minimum length.

Clearly, Eqns. (45), (44), and (43) together define a quadratic minimization problem with linear constraints, and this set up is used to map Traveling Salesman Problem on Hopfield networks [26, 58]. However, both the cost function as well as the sets of constraints are not of the form permitted in Eqn. (32).
Figure 14: Schematic illustration of the mapping procedure for solving travelling salesman problem. (a) The graph corresponding to four cities, along with the distances between cities. (b) The corresponding single-electron system comprising four computational modules, each of which is initialized with a single electron. (c) A minimum-energy configuration of the electrons, which defines a minimum length tour of the cities.

In order to set up a minimization problem that conforms to Eqn. (32), we first define a modified distance matrix $E$ as follows:

$$E_{ij} = \begin{cases} A & \text{if } i = j, \\ d_{ij} & \text{otherwise,} \end{cases}$$

where $A$ is a large number, e.g., larger than the length of a tour of the cities in order (i.e., $A \geq 1 + \sum_{t=1}^{n} d_{t(t+1)}$, where $d_{n(n+1)} = d_{1n}$). We also define another matrix

$$F = A I_{n \times n}$$
We next define a symmetric \( n^2 \times n^2 \) matrix \( A \) as follows:

\[
A = \begin{bmatrix}
0_{n \times n} & E & F & \cdots & \cdots & F & F & E \\
E & 0_{n \times n} & E & F & \cdots & \cdots & F & F \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
F & F & F & \cdots & \cdots & E & 0_{n \times n} & E \\
E & F & F & \cdots & \cdots & F & E & 0_{n \times n}
\end{bmatrix}
\]

Let \( Z = [z_{11} \cdots z_{1n} z_{21} \cdots z_{2n} \cdots z_{n1} \cdots z_{nn}]^T \in \{0,1\}^{n^2} \), then we can state the following theorem that states: the Traveling Salesman Problem as a quadratic minimization problem of the form given in Eqn. (32).

**Theorem 9** The minimum tour length is given by the result of the following minimization problem:

\[
\text{Minimize } Z^T A Z \text{ such that } \sum_{j=1}^{n} z_{ij} = 1 \text{ for all } i = 1, \ldots, n \text{ and } z_{ij} \in \{0,1\}
\]

(46)

Moreover, if \( Z_0 \in \{0,1\}^{n^2} \) is an optimal solution then it corresponds to a valid tour (i.e., it satisfies constraints in Eqns. (43) and (44)) with the minimum length.

**Proof:** First we note that for any given \( Z \in \{0,1\}^{n^2} \) that satisfies the constraints in Eqns. (43) and (44), the cost function

\[
Z^T A Z = 2[\text{Length of the tour defined by } Z].
\]

Thus minimization of \( Z^T A Z \) leads to a tour of minimum length.

To complete the proof we have to show that any solution \( Z \) of the optimization problem stated in Eqn. (46) will always satisfy the constraints in Eqns. (43) and (44). Since, the constraints in Eqns. (46) and (43) are identical, any solution \( Z \) always satisfies Eqn. (43).

Next, let us consider a \( Z \) that does not satisfy the constraints in Eqn. (44). That is, there exists a \( k \) such that \( \sum_{i=1}^{n} z_{ik} \geq 2 \) (i.e., the \( k \)th city has been assigned to more than one stop). From the construction of matrix \( A \) and the choice of \( A \) it follows that

\[
Z^T A Z \geq 2 \Lambda
= 2[\text{Length of the tour of the cities in the order } 1, 2, \ldots, n]
= Z_1^T A Z_1^T,
\]

where \( Z_1 \) is the vector that defines the tour of the cities in order (i.e., 1, 2, \ldots, \( n \)). Thus, any \( Z \) that does not satisfy the constraints in Eqn. (44) is not an optimal solution.

Note that the \( n \) linear constraints in Eqn. (46) imply that the corresponding single electron system comprises \( n \) computing modules of the kind shown in Fig. 10, and that each such module is initialized with a single electron. Fig. 14 shows the system organization, initial distribution of electrons, and a final minimum-energy configuration (corresponding to an optimal tour) for a four-city Traveling Salesman Problem.
3.5 Alternative Networks

In this section we will briefly describe two alternative approaches, which are currently under investigation, and are also based on technologies involving arrays of metallic islands. In section 3.5.1 we consider networks of islands in which the transport nonlinearities inherent in the molecular links connecting the islands become integral to the functionality of the system. In section 3.5.2 we consider networks in which field-effect transistors of small gate length and width are controlled by single-electron nodes, which are embedded in the insulating region.

3.5.1 Electron Dynamics in Networks with Quantum Links

The exploratory approach discussed in this section is motivated by the observation that wave and particle effect devices afford a rich set of nonlinearities which when coupled together in a network can show complex dynamical effects. In this scheme the devices, as well as the conductive links which connect the devices will possess non-linear transport characteristics, which will become integral to the functionality of the system. We consider, again an array of nanostructured metallic islands in which conduction is governed by Coulomb blockade. The usual model for transport in these arrays uses a linearized formula for the tunnel rates to account for conduction between two adjacent islands. This linearized model leads to an essentially ohmic behavior of the network, once the applied bias exceeds the Coulomb threshold [7]. There is evidence, however, that Coulomb blockade effects can coexist with transport mechanisms other than tunneling. So we begin by hypothesizing that these alternative conduction channels can afford strong non-linearities, including the possibility of unidirectional electron propagation along the channels [34]. If such networks can be fabricated, then the non-linear interactions amongst thousands of devices which are collectively driven far from equilibrium by applied biases of high amplitude, may reveal a fundamental kind of computational effect [49].

A plausible conductive link of this kind, could be established by replacing the tunnel barriers between the islands with short quantum wires, whose phase coherent transport characteristics can be understood using the Landauer formalism [14, 60]. These links can in principle be used to effect charge transfer between non-adjacent islands, which will help overcome topological limitations of typical architectures in which tunnel events are permitted only between nearest neighbors. The usual way of making quantum wires using lithographic procedures, will of course be impractical in the scheme being described here. We have in mind novel quantum wires established, for instance, by molecular links, or by atomically-precise lithography [63] both of which are being researched extensively at present. These links need to be only a few nanometers in length, so that the requisite phase coherence can be maintained across them at room temperature. The two metallic islands which are linked by such a quantum wire can be modelled as reservoirs in which the phase of the electrons is broken. If the metallic islands, are sufficiently small that the charging of them is governed by appreciable Coulomb blockade effects, then an additional non-linearity will become available at the nodes. The fact that the phase coherence is regularly disrupted within the islands, and the fact that transport across large arrays consisting of several thousand islands is typically studied, indicates that strong conductance fluctuations, which are endemic to phase coherent devices [39], will be averaged out. It is not clear at present whether such a network can give rise to useful computational effects. We believe, however, that the complex spatial and temporal variations of the electrostatic potential accompanying electron dynamics in such a network can be interpreted as a computational process.

The fundamental question that remains to be resolved concerns the applicability of the semi-classical 'orthodox theory' of Coulomb blockade, in the presence of leaky of conductive links. A
density matrix approach which takes proper account of the wave mechanics and the Coulomb blockade will become too cumbersome, and will prohibit the numerical study of large enough networks. We therefore favor a first order study which assumes the availability of non-linear conductive links compatible with the 'orthodox' theory. This effort is in a preliminary stage, at present.

3.5.2 Hybrid Functional Devices

An aspect of research in device physics is the fact that there is a substantial gap in dimensions and current drivability between the two classes of devices which are at present being researched extensively. As was discussed earlier, conventional FETs typically involve gate widths of about 10μm, while the gate length is scaled down to as low a value as one can accommodate without losing transistor action. The scaling of gate lengths in this way is likely to encounter a bottleneck well before the 10nm mark. Once this is reached, similar switching action can at present be found only in much smaller devices which operate on the basis of quite new physical effects, such as the Coulomb blockade. It is clear from the foregoing, that if a proposal which seeks to achieve functionality solely on the basis of nanoelectronic primitives is successful, a revolutionary change in electronics will have occurred, amounting to an increase in integration levels by a factor of thousand, or more [34, 5]. An examination of the history of advanced technology, however, indicates that this type of paradigm shift is rare, and suggests instead a more incremental approach [38, 32]. It is useful, therefore, to examine the prospects for more modest proposals which build upon the strengths of microelectronics, while taking advantage of novel nanoelectronic concepts. A signal development which might motivate this approach is the recent demonstration of a single electron non-volatile memory device, which can be operated at room temperature [69]. The appeal of the hybrid device concept, consists in its ability to be integrated within a traditional silicon process technology. This approach indicates a scenario in which nanoelectronic device concepts are gradually incorporated into conventional integrated electronics, which could eventually lead to purely nanoelectronic integrated circuits.

4 Concluding Remarks

The approach to the design of integrated nanoelectronic systems described in this report has been developed around the thesis that it is possible to continue the use of semiclassical device physics well into the nanoelectronic regime. We have been additionally guided by our conviction that significant progress in integrated nanoelectronics will occur by overcoming currently perceived obstacles to the creation of suitable interconnection networks. Accordingly, we believe that exploratory research which relaxes the interconnection bottleneck in various plausible ways, can play a vital role in the visualization of novel possibilities. If useful information handling attributes are eventually uncovered within this picture, this class of networks will in effect become 'artificial solids' [30] which compute. These computing solids conform also to the picture originally envisioned by Morton, who suggested that "substantial inroads might be expected by abandoning entirely classical circuit concepts and going directly to the interactions between energy and matter" to create 'functional devices' [48].

The work discussed in this report has been unified under a particular technology based on the creation of arrays of nanometer-sized metallic islands. We then considered different types of network mechanisms for the transfer of electrons between islands. Depending on the types of transport nonlinearities permitted by the network links, we found that it is possible to generate different kinds of collective computational activity in the networks. In particular, we showed that within a classical circuit theoretic model, non-monotone nonlinearities in the local transport can
yield global associative memory effects. We then found that this picture will breakdown within a single-electronic picture if the islands become too small. We then investigated networks in which the sole nonlinearity arose from the discreteness of electronic charge. These networks were shown to yield associative memory effects, as well as yield approximate solutions to certain NP-complete optimization problems.

Acknowledgements
The authors wish to thank Professor Supriyo Datta for valuable discussions, and Mr. Xiaodong Wang for help with numerical simulations. This work was supported in part by an NSF Postdoctoral Research Associateship (Grant No. ASC-9211073), by the NSF Grant No. ECS-9308814, and by the General Motors Faculty Fellowship from the Schools of Engineering at Purdue University.
References


