

# 1 Frontiers in device engineering

*Philip Seliger and A.F.J. Levi*

## 1.1 Introduction

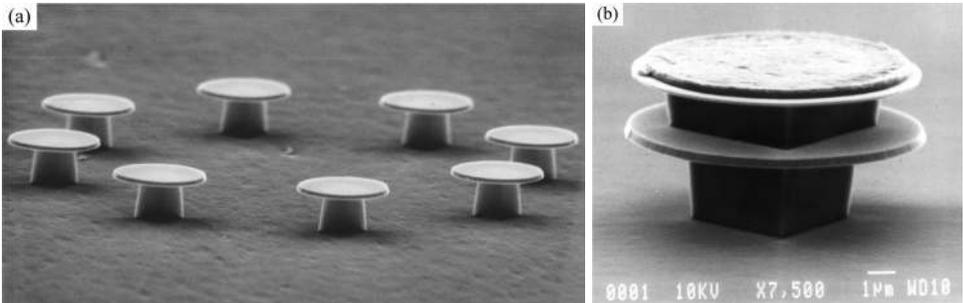
Today, nanoscience promises to provide an overwhelmingly large number of experimentally accessible ways to configure the spatial position of atoms, molecules, and other nanoscale components to form devices. The central challenge of *nano-technology* is to find the best, most practical, configuration that yields a useful device function. In the presence of what will typically be an enormous non-convex search space, it is reasonable to assume that traditional ad hoc design methods will miss many possible solutions. One approach to solving this difficult problem is to employ machine-based searches of configuration space that discover user-defined objective functions. Such an optimal design methodology aims to identify the best broken-symmetry spatial configuration of metal, semiconductor, and dielectric that produces a desired response. Hence, by harnessing a combination of modern computer power, adaptive algorithms, and realistic physical models, it should be possible to seek robust, manufacturable designs that meet previously unobtainable system specifications. Ultimately one can envision a design process that simultaneously is capable of basic scientific discovery and engineering for technological applications.

This is the frontier of device engineering we wish to explore.

### 1.1.1 The past success of ad hoc design

For many years an ad hoc approach to device design has successfully contributed to the development of technology. For example, after identifying the cause of poor device performance one typically tries to create a solution by modifying a process or fabrication step. The result is usually a series of

### Frontiers in device engineering



**Fig. 1.1.** Symmetry, typical of ad hoc device design, is illustrated by SEM micrographs of (a) optically pumped and (b) electrically driven microdisk lasers. The image on the left shows a field of InGaAsP quantum well microdisk lasers supported by InP posts on an InP substrate. Lasing emission is at  $\lambda_0 = 1550$  nm wavelength. Disks are  $2\ \mu\text{m}$  in diameter and  $0.1\ \mu\text{m}$  thick. Nearest-neighbor spacing is  $4.7\ \mu\text{m}$  and the outside diameter of the ring of disks is  $14\ \mu\text{m}$ . Room temperature threshold power is sub-mW for pump radiation at  $850$  nm wavelength. The image on the right is an electrically driven 6-quantum well InGaAsP microdisk laser diode that is  $10\ \mu\text{m}$  in diameter and  $0.3\ \mu\text{m}$  thick. Electrical current is injected from the top metal contact, room temperature threshold current is  $2\ \text{mA}$ , and lasing emission is at  $\lambda_0 = 1550$  nm wavelength.

innovations heavily weighted towards incremental, and hence small, changes in previous practice. The scaling of Complementary Metal Oxide Semiconductor (CMOS) transistors to minimum features sizes of a few nm is a good example of the extraordinary power of such an approach [1, 2].

In addition to incremental improvements there are, of course, new device designs and device concepts that emerge from the research community. Typically, these are also ad hoc in origin and, significantly, tend to have a geometric structure that is highly symmetric. The development of radiation-pressure-driven opto-mechanical resonators is a recent example in which the phenomenon was first explored using highly symmetric toroidal structures [3]. The creation of ultra-small semiconductor lasers, such as the microdisk lasers illustrated in Fig. 1.1, is another [4–6]. The choice of symmetric geometries seems to be a human bias, often driven by ease of analysis. Put simply, symmetric structures are easier to think about. It is symmetric geometries, along with their implicit limited functionality, that are foremost in ad hoc design and are, in general, preferred by the research community.

#### 1.1.2 Looking beyond ad hoc design

Rather than speculate on the reasons for the past success of an ad hoc design methodology, it is more interesting to explore the possibility of an alternative path to design and discovery using new and emerging capabilities such as

---

## 1.2 Example: Optimal design of atomic clusters

---

nanoscience and access to large computing resources. Part of the motivation comes from the fact that while nanoscience has successfully developed a large number of degrees of freedom with which to create new structures, much of what has been proposed (with the notable exception of quantum computing) is focused on replacing existing electronic and photonic devices such as transistors and lasers with their nearest nanotechnology equivalent. The shortcoming in such an approach is a failure to discover new functions, devices, and systems specific to and only achievable using nanoscience. It is hoped that unbiased machine-based searches for functionality will reveal original, nonintuitive, designs characterized by broken-symmetry geometries.

The rapid and successful development of nanoscale fabrication methods has exposed a critical gap in understanding that appears to represent a very important barrier to fully exploiting nanoscience. Absent from largely experimentally-driven nanoscience research is a methodology or procedure to create new functionalities, new devices, and new system architectures. What is needed is a systematic experimental and theoretical approach that results in the efficient discovery of atom, molecular, and macro-molecular based configurations that exhibit the desired, user specified, functionality. It is the ability to provide made-to-order functions, devices, and systems that will enable the true potential of nanoscience and ensure its adoption in practical systems.

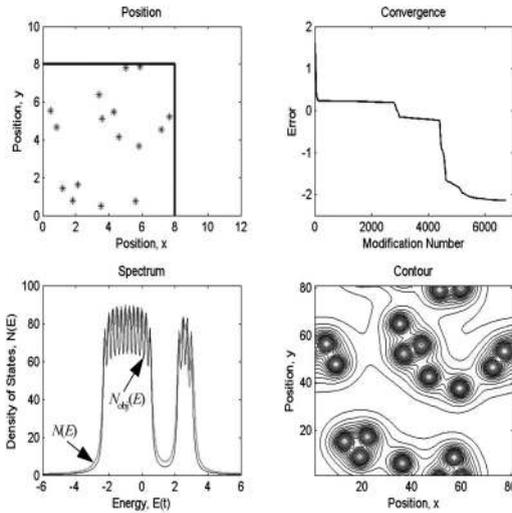
Two key elements of this approach to design are efficient adaptive search algorithms and realistic physical models. Combined they form the basis for the development of optimal design software for small quantum systems. Implemented, such algorithms are capable of discovering initially nonintuitive designs for a given functionality. Typically these designs are highly non-symmetric and usually difficult to interpret. However, as will be illustrated in Section 1.2, sometimes it is possible to analyze the machine-generated solutions and gain new insight into the underlying physical mechanisms driving the system to a given optimal configuration. This potential for learning is another motivation to explore the possibilities of optimal design in nanoscience.

---

## 1.2 Example: Optimal design of atomic clusters

The density of electronic states in a solid is a basic attribute that plays a key role in determining material properties. For example, a singular behavior in the density of quasi-particle states can result in enhanced optical activity at a specific photon energy. In fact, a periodic array of atoms in a crystal gives rise to just such peaks in the density of states. This may be illustrated by

## Frontiers in device engineering



**Fig. 1.2.** Optimized position of 16 atoms in a square of dimension  $8 \times 8$  (upper left) giving asymmetric density of states  $N(E)$  very close to the objective spectrum  $N_{\text{obj}}(E)$  [7]. Energy scale is in units of  $t$ . Contour plot of interaction potential for atoms in optimized positions is shown in lower right panel. In the calculations  $\alpha = 3$ , periodic boundary conditions are used, and  $\Gamma = 0.2828 \times t$ . Convergence as a function of modification number is shown in upper right panel.

considering a two-dimensional square lattice in the nearest-neighbor tight-binding approximation with  $s$ -orbitals and eigenenergies  $E_i$ . Crystal symmetry and interaction mechanism determine the density of states spectrum. In this case, there is a peak in the density of states at the center of the band.

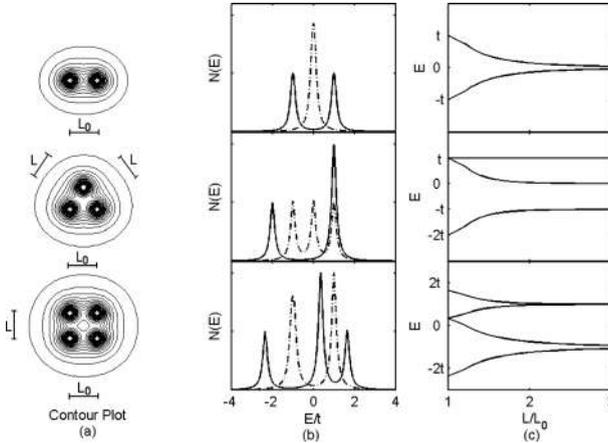
What we would like to do is find configurations of atoms that are not constrained by crystal symmetry. A key idea is that breaking the spatial symmetry of atom positions creates a truly vast number of possibilities, making it feasible to find configurations of atoms with essentially any desired density of states. The ability to control the response of a material in a user-defined way is a powerful concept which has the potential to change the way one views materials, devices, and systems.

As a first step, consider an algorithm that seeks spatial configurations of atoms characterized by a user-specified or objective density of electronic states  $N_{\text{obj}}(E)$ . The essential physics underpinning the approach is illustrated by considering a *long-range* version of the atomic tight-binding model with Hamiltonian

$$\hat{H} = - \sum_{i,j} t_{i,j} (\hat{c}_i^\dagger \hat{c}_j + \hat{c}_i \hat{c}_j^\dagger), \quad (1.1)$$

where  $c_i^\dagger$  and  $c_i$  are creation and annihilation operators respectively at the

1.2 Example: Optimal design of atomic clusters



**Fig. 1.3.** Calculated density of states  $N(E)$  for dimer, trimer, and quadrumer with  $\alpha = 3$  and showing a hierarchy of non-symmetric contributions as a function of atom separation  $L$  normalized to  $L_0$ . Energy scale is in units of  $t$ . An isolated pair of atoms (dashed line) is symmetrically split (solid line) by a dimer. A trimer forming an equilateral triangle ( $L/L_0 = 1$ ) has an asymmetric density of states (solid line), becoming essentially symmetric when  $L/L_0 = 3$  (dashed line). Peak positions for the quadrumer are also controlled by atom separation in the range  $L/L_0 = 1$  (solid line) to  $L/L_0 = 3$  (dashed line) [7].

atom site  $\mathbf{r}_i$ . The overlap integrals  $t_{ij}$  between an atom at position  $\mathbf{r}_i$  and an atom at position  $\mathbf{r}_j$  are parameterized by a power law  $t_{ij} = t/|\mathbf{r}_i - \mathbf{r}_j|^\alpha$ , where  $t$  sets the energy scale. The choice of exponent  $\alpha$  depends on details of the experimental situation. Here, the Hamiltonian matrix in the basis of single particle states is non-sparse because interaction with all atoms is included. For simplicity only  $s$ -orbitals are considered, so it is not necessary to include directionality of atomic electron wave functions. The density of states is

$$N(E) = \sum_i \frac{|\Gamma|/\pi}{(E - E_i)^2 + (\Gamma/2)^2}, \tag{1.2}$$

and  $\Gamma$  is the characteristic energy broadening of each eigenenergy,  $E_i$ .

To demonstrate the power of optimal design, consider the non-symmetric objective density of states spectrum in two dimensions,  $N_{\text{obj}}(E)$ , indicated in the lower left of Fig. 1.2. The optimization algorithm finds a spatial configuration of 16 atoms in an  $8 \times 8$  area with periodic boundary conditions that has a density of states,  $N(E)$ , essentially identical to the desired or objective spectrum [7]. The implication is both apparent and dramatic: a user who requires new material with a specific quasi-particle density of states can use optimal design software to discover configurations of atoms with the desired behavior. The objective functionality is obtained by broken symmetry so, in

---

## Frontiers in device engineering

---

this sense, broken symmetry *is* function.

It is clear from the atom positions indicated in Fig. 1.2 that one could not have guessed the result. However, the output of the computer program can be used to gain new insight into configurations that result in the desired spectrum  $N_{\text{obj}}(E)$ . In this particular case, and as illustrated in Fig. 1.3, one learns that a hierarchy of primitive configurations exists that form the building blocks for any objective density of states. Dimers can be used for symmetric  $N(E)$ , trimers and larger molecular configurations provide asymmetry to  $N(E)$ . While, in a strict sense, these heuristics only apply to the dilute limit in which the normalized average spacing between atoms is much greater than unity, it is apparent one may appeal to this insight to explain the more complicated structures that occur in the dense limit.

Remarkably, some aspects of the model have been confirmed experimentally using scanning tunneling microscopy (STM) to precisely position gold atoms on the surface of a nickel-aluminum crystal. STM measurements [8] show that the splitting in the value of eigenenergies  $E_i$  for Au dimers on NiAl depends inversely on Au atom separation corresponding to  $\alpha = 1$  in the expression  $t_{ij} = t/|\mathbf{r}_i - \mathbf{r}_j|^\alpha$ .

Chapter 2 discusses optimal design of atomic clusters in more detail.

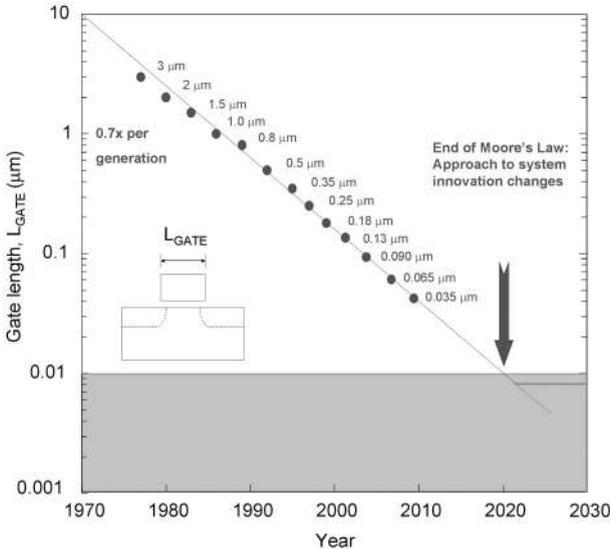
---

### 1.3 Design in the age of quantum technology

One of the greatest achievements of semiconductor technology has been the continuous reduction in transistor minimum feature size over the past 35 years. Often described as Moore's Law [1] or scaling, Fig. 1.4 illustrates the historical exponential reduction in CMOS gate length with time. Of course, at some point physical and other limitations will force such geometric scaling to end. Today there seems to be a consensus that a manufacturable technology with minimum feature sizes below 10 nm is achievable [2]. This confidence is based partly on improvements in lithography tools and partly on experience overcoming previously declared limits to scaling [9]. Nevertheless, sometime after 2020 Moore's Law will come to an end and new paths to system innovation will have to be found. Our concern is not how to achieve minimum feature size below 10 nm but rather the approach to design when such capability is available because it is on these nm length scales that the best opportunities to exploit quantum effects will occur.

When simple physical scaling of device geometry no longer provides a path to increased system functionality, improved device performance and function might be achieved by manipulating new quantum degrees of freedom. Examples might include controlling the single electron states of atomic

### 1.3 Design in the age of quantum technology



**Fig. 1.4.** Illustrating reduction in CMOS gate length with time. Gate length has been decreasing, or scaling, consistently for the past 35 years. However, physical and other limitations to continued scaling will impact by about the year 2020 and geometric scaling will come to an end. As this end-point nears, fundamental changes in the approach to system innovation are required.

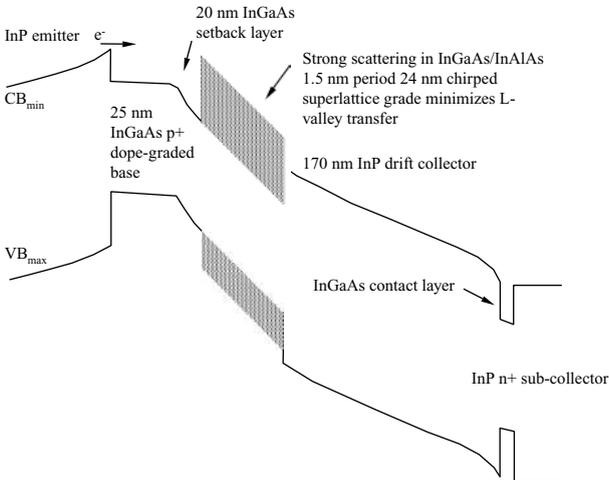
and nm-sized particles via geometry [10, 11], using interacting electrons in the presence of the coulomb interaction to exploit collective excitations such as plasmons [12], hybridization to control bonding and chemical specificity, using electron and orbital spin to control magnetic response [13], strong light-matter interaction in nm-sized geometries [14, 15], and non-equilibrium processes on fs time scales [16].

Because of the large number of variables, it seems reasonable to consider avoiding ad hoc design and trying to apply a systematic approach to problem solving and analysis. Such an approach is more likely to reap dividends as we move away from devices that behave semi-classically and into a less familiar quantum regime where our intuition might fail. However, to date, much of what has been explored in nanoscience is the result of curiosity-driven research with little direct connection to practical technology development. This approach to discovery may not only be inefficient but may also be susceptible to replacement by more effective methods.

#### 1.3.1 High performance heterostructure bipolar transistors

An example of a high-performance electronic device designed in an ad hoc fashion but with nm control of material composition in one dimension is the heterostructure bipolar transistor (HBT).

### Frontiers in device engineering



**Fig. 1.5.** A high-performance heterostructure bipolar transistor (HBT) band profile. The InP emitter injects electrons into the heavily doped p-type base. The collector contains a complex chirped InGaAs/InAlAs superlattice. Designs similar to this have achieved extrapolated characteristic frequency response in excess of 500 GHz at room temperature [17].

The semiconductor band profile of just such an HBT is illustrated in Fig. 1.5. As may be seen, the band profile is quite complex. It is, in fact, the result of many years of iterative improvement in design culminating in the demonstration of an extrapolated characteristic frequency response in excess of 500 GHz at room temperature [17].

Despite this remarkable success, the emitter, base, and complex collector structure has never been systematically optimized. In part this is due to the fact that no reliable, efficient, physically realistic model of electron transport has been developed. Today, no one knows if the designs that have been implemented are optimal, or even close to optimal.

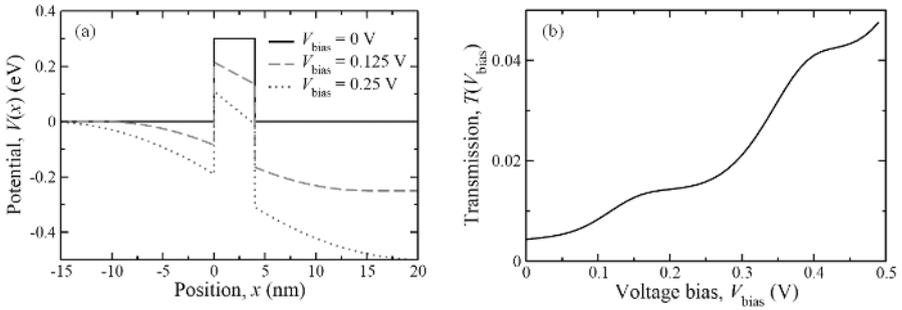
The next section illustrates how one might approach an optimal design strategy for one aspect of nanoscale electronic device design.

### 1.3.2 Control of electron transmission through a tunnel barrier

As a prototype system, consider the semiconductor  $\text{Al}_x\text{Ga}_{1-x}\text{As}$ , which has a lattice constant 0.5653 nm and atomic layer separation 0.2827 nm. Atomically precise layer-by-layer crystal growth is possible using molecular beam epitaxy (MBE) and the  $\text{Al}_x\text{Ga}_{1-x}\text{As}$  alloy can be used to form heterojunctions with controlled conduction and valence band off-sets.

As a specific example [18], consider electron transport through a rectangular  $\text{Al}_x\text{Ga}_{1-x}\text{As}$  barrier of energy  $V_0 = 0.3$  eV and width  $L = 4$  nm, as shown

### 1.3 Design in the age of quantum technology

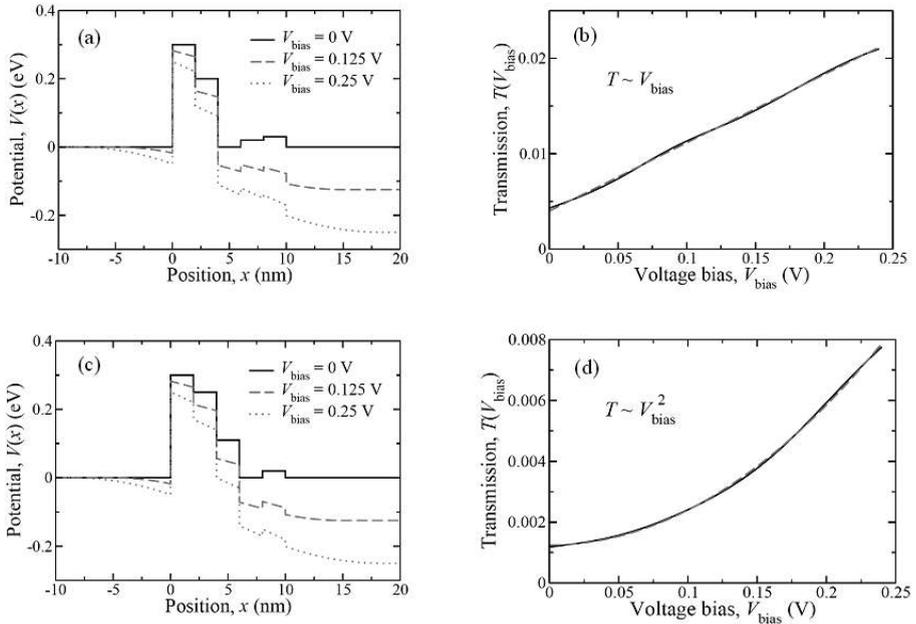


**Fig. 1.6.** A rectangular potential barrier of energy  $V_0 = 0.3$  eV and width  $L = 4$  nm gives rise to rapid increase in electron transmission with increasing voltage bias,  $V_{\text{bias}}$  and resonances. Effective electron mass is  $m = 0.07 \times m_0$ , where  $m_0$  is the bare electron mass. (a) Conduction band profile of the rectangular potential barrier for the indicated values of  $V_{\text{bias}}$ . (b) Transmission probability as a function of  $V_{\text{bias}}$  for an electron of energy  $E = 26$  meV incident from the left.

in Fig. 1.6. The barrier is sandwiched between  $n$ -type GaAs electrodes with carrier concentration  $n = 10^{18} \text{ cm}^{-3}$ . Applying a bias voltage,  $V_{\text{bias}}$ , results in a depletion region on the right side and an accumulation region on the left side of the barrier. The form of the conduction band profile  $V(x)$  in these regions is calculated by solving the Poisson equation. Net electron motion is in the  $x$  direction, normal to the barrier plane and there is no confinement in the  $y$  and  $z$  directions, thereby avoiding possible detrimental consequences of quantized conductance [19–21]. A numerical solution to the Schrödinger equation is obtained piecewise by discretizing the potential profile into 4,000 steps, matching boundary conditions at each interface, and implementing the propagation matrix method [22–24]. An electron of energy  $E = 26$  meV incident from the left is partially reflected and partially transmitted, as determined by the wave function boundary conditions  $\psi_j = \psi_{j+1}$  and  $\partial\psi_j/\partial x = \partial\psi_{j+1}/\partial x$  at each interface. Here  $\psi_j$  is a solution of Schrödinger's equation in region  $j$  with wave vector  $k_j = \sqrt{2m(E - V_j)}$ , where  $V_j$  is the local potential in the conduction band and  $m$  is the effective electron mass.

Exponential increase in electron transmission with bias voltage is a generic feature of the simplest barrier profiles. Potential wells, on the other hand, are known to produce bound-state resonances, leading to sharp transmission peaks. Hence, design of structures with linear and other power-law transmission-voltage characteristics likely involves broken-symmetry potential barrier profiles. As an initial challenge in our exploration of this possibility we use an adaptive quantum design approach to find a potential profile with a transmission function  $T(V_{\text{bias}})$  that increases linearly with bias voltage in the window  $0 \text{ V} < V_{\text{bias}} < 0.25 \text{ V}$ .

### Frontiers in device engineering



**Fig. 1.7.** (a) and (c) are solutions from exhaustive numerical searches for conduction band profiles  $V(x)$  that yield linear and square dependences of electron transmission as a function of bias voltage,  $V_{\text{bias}}$ .  $V(x)$  is constrained to a region that is 10 nm wide and the maximum local potential is 0.3 eV. The resulting  $T(V_{\text{bias}})$  for an electron of energy  $E = 26$  meV incident from the left are shown as solid lines in (b) and (d). Broken line is the objective response.

The conduction band potential energy profile is defined on a grid with  $\Delta x = 2$  nm (about 8 monolayers in GaAs) spatial increments and  $\Delta V = 0.01$  eV energy increments. The numerical search for optimal broken-symmetry barrier profile is constrained to take into account physical as well as computational limitations. Physically, varying the composition of an  $\text{Al}_x\text{Ga}_{1-x}\text{As}$  alloy controls the conduction band potential profile. Each alloy plane normal to the growth direction has an average local potential,  $V(x)$ . Fabrication inaccuracies of 1–2 monolayers may occur in the epitaxial growth processes, and hence the targeted transmission functionality needs to remain stable against such variations. Moreover, the Al concentration can only be controlled to within a few percent. Computationally, the dimensionality of the search space needs to be constrained in order to match the available computer hardware capabilities. In this example, to keep the search space finite, we focus on nanoscale barrier structures of total width  $L = 10$  nm with a maximum on-site potential of 0.3 eV measured from the GaAs conduction band minimum.