

# Wavefunction delocalization in quantum dot arrays: an asymptotic analysis

D. A. Edwards · W. M. Reid · M. F. Doty

Received: 26 January 2012 / Accepted: 4 August 2012 / Published online: 23 October 2012  
© Springer Science+Business Media B.V. 2012

**Abstract** Intermediate-band solar cells using quantum dot arrays (QDAs) are theoretically predicted to significantly increase the efficiency with which solar energy can be harvested. In the limit of identical quantum dots, the wavefunction for electrons in a QDA will be fully delocalized. Fully delocalized wavefunctions have been theoretically shown to reduce thermal losses and consequently increase photovoltaic device efficiency. However, even small nonuniformities can cause electrons to localize in a single quantum dot, negating any advantages from delocalized states. In this work a modified Schrödinger equation is used to model a two-dot array with nonuniform quantum dots and solved using perturbation methods. This result is extended to  $N$ -dot arrays, and several metrics are constructed to characterize the degree of delocalization. Our results, which compare favorably with numerical simulations, show explicitly how the amount of delocalization depends on key design parameters.

**Keywords** Photovoltaics · Quantum dot array · Schrödinger equation · Tunnelling · Wavefunction delocalization

## 1 Introduction

Intermediate-band solar cells (IBSCs) are theoretically predicted to improve the efficiency of photovoltaic devices because they allow a wider portion of the solar spectrum to be harvested [1–3]. The leading material for fabricating intermediate bands is densely packed, vertically aligned arrays of self-assembled InAs quantum dots (QDs) embedded in a GaAs host [4,5]. Present theoretical studies are based on the assumption that the quantum dot arrays (QDAs) are fabricated from identical quantum dots, leading to the formation of delocalized wavefunctions

---

D. A. Edwards (✉)  
Department of Mathematical Sciences, University of Delaware, Newark, DE 19716, USA  
e-mail: edwards@math.udel.edu

W. M. Reid · M. F. Doty  
Department of Materials Science and Engineering, University of Delaware, Newark, DE 19716, USA

W. M. Reid  
e-mail: wmr@udel.edu

M. F. Doty  
e-mail: doty@udel.edu

through coherent tunneling between the individual quantum dots. These delocalized wavefunctions are believed to suppress phonon-mediated relaxation (e.g., loss of energy to heat) and therefore enhance device performance [5–8]. In practice, however, QDs are never identical. Although the period of the array can be well controlled, realistic arrays always have an inhomogeneous distribution of the potential energy levels in the QDs [9]. By analogy with Anderson localization theory [10], the nonuniformity of energy levels causes the wavefunctions to localize, negating any advantages derived from the delocalized wavefunctions.

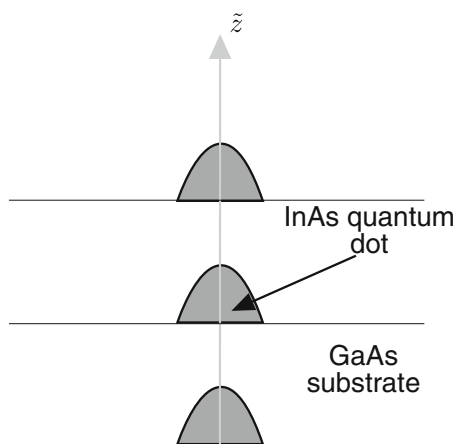
In this work, we create a detailed model for a two-dot array bounded by both infinite- and finite-height potential barriers. The standard Schrödinger equation is modified to take into account that the effective mass of the electron is different in the dot and energy barrier [11]. The model includes several metrics that can be used to measure the amount of delocalization present in an array. We solve the model analytically using perturbation methods in the small parameter  $\delta$ , which represents the inherent variability in the QD potential levels. We provide analytical solutions to indicate how the amount of delocalization depends not only on  $\delta$ , but also on the key parameter  $G$ , the width of the barriers between QDs (which has an experimental lower bound due to strain concerns [9]).

Once we have established results for a two-dot array, we then postulate a simple relationship to extend these results to arrays of general size. Our results compare favorably with the numerical simulations in [11]. The explicit parameter dependence provided by the analytical solution can provide guidance to experimentalists about the sensitivity of delocalization to various fabrication factors. Moreover, the analytical solution provides a means to avoid computationally intensive numerical modeling of delocalization in developing new device designs.

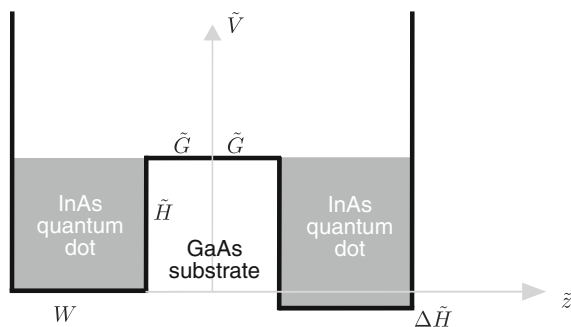
## 2 Governing equations

The profile of the QDA is shown in Fig. 1. An InAs QD (shaded) is grown on GaAs (white). After growing to the desired height, the dot is capped off with more GaAs before the next dot in the array is grown. Here  $\tilde{z}$  is the direction pointing toward the top of the solar cell (and hence the light). We consider a one-dimensional problem along the  $\tilde{z}$ -axis, assuming that variations in the other directions are negligible. This is also justified by the fact that the gap between dots is narrowest (and hence the potential for tunneling greatest) along the  $\tilde{z}$ -axis.

For simplicity, at first we consider a simple two-dot array, as shown in Fig. 2. Though shaded to match Fig. 1, note that in Fig. 2,  $\tilde{z}$  has been rotated and the origin has been shifted. The GaAs provides the internal barriers, while the InAs QDs provide the potential wells. Due to the differing band structure of InAs and GaAs, the electron will have a different effective mass  $\tilde{m}(\tilde{z})$  in each region:



**Fig. 1** Fabrication of quantum dot array



**Fig. 2** Dimensional-variable description of potential in two-dot array with infinite side walls. Note the placement of the origin and that  $\tilde{G}$  is the half-width of the step

**Table 1** Parameter values from the literature

Parameter	Value	Reference
$2\tilde{G}$ (nm)	2–15	[18]
$\tilde{H}$ (meV)	850	[11]
$m_b$ (kg)	$6.10 \times 10^{-32}$	[11] (calculated)
$m_w$ (kg)	$2 \times 10^{-32}$	[11] (calculated)
$W$ (nm)	6.5	[11]
$\Delta\tilde{H}$ (meV)	56.5	[11], designated as $\sigma_0$

$$\tilde{m}(\tilde{z}) = \begin{cases} m_w & \tilde{G} < |\tilde{z}| < W + \tilde{G}, \\ m_b & |\tilde{z}| < \tilde{G}, \end{cases} \quad (2.1)$$

where the subscript “b” stands for “barrier” and the subscript “w” stands for “well.” The effective mass is less than the true electron mass (Table 1). In particular,  $m_w$  is only 2.2% of the true electron mass of  $9.11 \times 10^{-31}$  kg [12, inside cover], while  $m_b$  is only 6.7% of the actual mass.

Given that the mass is now a function of  $\tilde{z}$ , the appropriate one-dimensional Schrödinger equation is given by [13]

$$-\frac{\hbar^2}{2} \frac{d}{d\tilde{z}} \left( \frac{1}{\tilde{m}(\tilde{z})} \frac{du}{d\tilde{z}} \right) + \tilde{V}(\tilde{z})u(\tilde{z}) = \tilde{E}u(\tilde{z}), \quad (2.2)$$

where  $u$  is the wavefunction,  $\tilde{E}$  the energy level, and  $\tilde{V}(\tilde{z})$  the potential.

Because delocalized states are formed by coherent tunneling, the formation of delocalized wavefunctions would be much easier if the half-width  $\tilde{G}$  were very small. Unfortunately, there is an experimental lower bound on  $\tilde{G}$ . The InAs QDs assemble because InAs has a different lattice spacing than GaAs. Relatively thick GaAs layers or strain-relief layers are required to avoid the accumulation of strain, which could lead to the formation of crystal defects that degrade device performance [9]. These considerations are particularly acute when one considers that the arrays must contain approximately 50 QDs to absorb enough photons [9].

QDAs are formed by the sequential deposition of QD layers. The QDs in each layer self-assemble to minimize strain and surface energy related to the mismatch in lattice constants between the deposited InAs and the GaAs surface on which the InAs is deposited. QDs in subsequent layers align because of strain propagation through the GaAs barrier that separates each layer. The thickness  $2\tilde{G}$  of the barrier can be precisely controlled with existing growth methods. Although there can be minor fluctuations in the height  $W$  of the QDs, the primary inhomogeneity in QDAs arises from the random diffusion of InAs during the self-assembly process and the random degree of alloying with the surrounding GaAs after the QD is capped. These effects lead to a fluctuation in the depth  $H$  of the confining potential. Fluctuations in the confining potential, lateral size of the QDs, or the height  $W$  of the QDs, result in fluctuations in the energy of discrete states confined within the QDs. The fluctuation in the depth of the confining potential dominates, so we capture all of these fluctuations with the potential displacement variable  $\Delta\tilde{H}$ , as shown in Fig. 2.

Because the energy levels can be measured experimentally, we take the measured energy levels of an ensemble of real QDs and calculate the corresponding range of  $\Delta\tilde{H}$  required to create an ensemble of square wells having a realistic distribution of energy levels. We can then model realistic arrays as a series of simple square wells of width  $W$  with each dot having a different  $\Delta\tilde{H}$ .

For simplicity of description, in this section we take the external barriers to be infinite. Hence we may model the potential as follows:

$$\tilde{V}(\tilde{z}) = \begin{cases} 0 & -(W + \tilde{G}) < \tilde{z} < -\tilde{G}, \\ \tilde{H} & |\tilde{z}| < \tilde{G}, \\ -\Delta\tilde{H} & \tilde{G} < \tilde{z} < W + \tilde{G}, \\ \infty & \text{else,} \end{cases} \quad (2.3)$$

as shown in Fig. 2.

The piecewise nature of the potential and effective mass poses the question of what conditions to impose at the interface. From a mathematical perspective, the wavefunction  $u$  should be continuous at any interface:

$$u(\tilde{z}_i^-) = u(\tilde{z}_i^+), \tag{2.4a}$$

where  $\tilde{z}_i$  is the position of any interface. To find the appropriate derivative condition, we integrate (2.2) across the interface to obtain

$$\frac{1}{\tilde{m}(\tilde{z}_i^-)} \frac{du}{dz}(\tilde{z}_i^-) = \frac{1}{\tilde{m}(\tilde{z}_i^+)} \frac{du}{dz}(\tilde{z}_i^+), \tag{2.4b}$$

where we have used (2.4a). To understand the physical interpretation of these conditions, we rewrite (2.4) as

$$\tilde{z}_i^- u(\tilde{z}_i^-) = \tilde{z}_i^+ u(\tilde{z}_i^+), \quad \frac{1}{\tilde{m}(\tilde{z}_i^-)} \left[ -i\hbar \frac{du}{dz}(\tilde{z}_i^-) \right] = \frac{1}{\tilde{m}(\tilde{z}_i^+)} \left[ -i\hbar \frac{du}{dz}(\tilde{z}_i^+) \right]. \tag{2.5}$$

The first condition states that the position operator is continuous across any interface. Keeping in mind that the momentum operator is given by

$$-i\hbar \frac{d}{dz}, \tag{2.6}$$

we see that the second expression in (2.5) implies that the *velocity* operator is continuous across any interface. For more details about these operators, see [14, §4.3], [15, Chap. 2], [16, §8.2].

We now introduce scalings to produce a simplified dimensionless problem. As the solution in the dots will be trigonometric (to be shown below), the appropriate scaling for the length is one that puts a dot of width  $W$  on a scaled length of width  $\pi$ :

$$z_c = \frac{W}{\pi}, \quad z = \frac{\tilde{z}}{z_c}, \quad G = \frac{\tilde{G}}{z_c}. \tag{2.7}$$

For the energy scale, we use the lowest energy level for a single well of width  $W$ , as given by [15, eq. (2.13b)]:

$$E_c = \frac{\hbar^2 \pi^2}{2m_w W^2}, \tag{2.8a}$$

which we will verify later. Hence we introduce the following scalings for the energy variables:

$$\tilde{V} = V E_c, \quad \tilde{E} = E E_c, \quad \tilde{H} = H E_c, \quad \Delta \tilde{H} = \delta E_c. \tag{2.8b}$$

Here  $\delta$ , which will be small, models the relative size of the potential well variance as compared to the characteristic one-well energy level. The limit  $\delta = 0$  corresponds to the case of perfectly uniform fabrication whose solution is symmetric and, hence, fully delocalized.

Substituting (2.7) and (2.8) into (2.1)–(2.3), we have

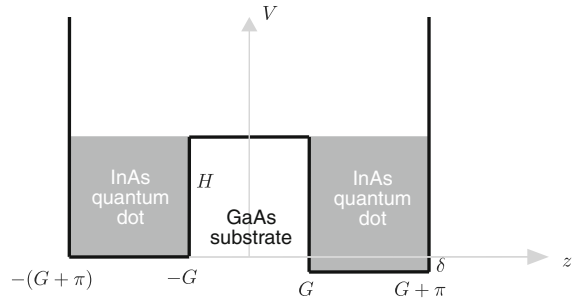
$$\frac{d}{dz} \left( \frac{1}{m(z)} \frac{du}{dz} \right) + [E - V(z)]u = 0, \tag{2.9}$$

$$V(z) = \begin{cases} 0 & -(\pi + G) < z < -G, \\ H & |z| < G, \\ -\delta & G < z < G + \pi, \\ \infty & \text{else,} \end{cases} \tag{2.10a}$$

$$m(z) = \begin{cases} 1 & G < |z| < G + \pi, \\ M \equiv m_b/m_w & |z| < G. \end{cases} \tag{2.10b}$$

The dimensionless version is illustrated in Fig. 3. As shown in Table 2,  $G$  and  $H$  are  $O(1)$ ,  $M > 1$ , and  $\delta$  is small. Hence we will use  $\delta$  as the small parameter in the forthcoming perturbation analysis.

**Fig. 3** Dimensionless-variable description of double-well potential with infinite side walls. Note the placement of the origin and that  $G$  is the half-width of the step



**Table 2** Calculated parameter values

Parameter	Value	Note
$E_c$ (meV)	405	
$G$	$4.83 \times 10^{-1}$ –3.62	Extreme values from Table 1
$G$	2.42	Corresponds to 10-nm-wide barrier
$G$	$7.25 \times 10^{-1}$	Corresponds to 3-nm-wide barrier
$H$	2.10	
$M$	1.75	
$z_c$ (nm)	2.07	
$\delta$	$1.40 \times 10^{-1}$	

We now derive the boundary conditions on the problem. At the interface with the infinite-height external barriers, we have

$$u(\pm(\pi + G)) = 0. \tag{2.11}$$

The conditions at the other interfaces may be found by substituting our scalings into (2.4):

$$u(-G^-) = u(-G^+), \quad u(G^-) = u(G^+), \tag{2.12a}$$

$$\frac{du}{dz}(-G^+) = M \frac{du}{dz}(-G^-), \quad \frac{du}{dz}(G^-) = M \frac{du}{dz}(G^+). \tag{2.12b}$$

Equation (2.12b) indicates that at the dot/barrier interface, the derivative in the barrier must always be  $M$  times the derivative in the dot.

To simplify the algebra, we exploit the piecewise nature of the problem and write

$$\frac{d^2u}{dz^2} + B(z)u = 0, \quad B(z) = \begin{cases} E & -(\pi + G) < z < -G, \\ -M(H - E) & |z| < G, \\ E + \delta & G < z < G + \pi, \\ \infty & \text{else.} \end{cases} \tag{2.13}$$

Solving (2.13) in the dots subject to (2.11), we obtain

$$u_l = S_l \sin \lambda_l (z + \pi + G), \quad \lambda_l = \sqrt{E}, \tag{2.14a}$$

$$u_r = S_r \sin \lambda_r (\pi + G - z), \quad \lambda_r = \sqrt{E + \delta}, \tag{2.14b}$$

where the subscripts “l” and “r” stand for “left” and “right.” We have written  $u_r$  in a slightly nonstandard form that clarifies the near-symmetry of  $u$  under the transformation  $z \mapsto -z$  (which has exact symmetry when  $\delta = 0$ ).

We are most interested in the degree of localization of the wavefunction  $u$  to one dot or the other. This can be characterized by the following ratio:

$$r = \left| \frac{S_r}{S_l} \right|. \tag{2.15}$$

Delocalized states have  $r$  close to 1. In fact, since with  $\delta = 0$  the problem is totally symmetric, we expect that  $r(\delta = 0) = 1$ , as consistent with the preceding discussion.

Satisfying the continuity of  $u$  at  $z = \pm G$  as given by (2.12a), we have

$$\begin{aligned} u_l &= (-S_m + C_m) \sin \lambda_l (z + \pi + G) \sin \lambda_r \pi, \\ u_m &= \left( S_m \frac{\sinh \alpha z}{\sinh \alpha G} + C_m \frac{\cosh \alpha z}{\cosh \alpha G} \right) \sin \lambda_l \pi \sin \lambda_r \pi, \quad \alpha = \sqrt{M(H - E)}, \\ u_r &= (S_m + C_m) \sin \lambda_l \pi \sin \lambda_r (\pi + G - z). \end{aligned} \quad (2.16)$$

Here the subscript ‘‘m’’ stands for ‘‘middle.’’ Given the values in Tables 1 and 2, we infer that  $H > E$  and hence  $\alpha$  is real.

Then enforcing (2.12b), we obtain

$$s_l S_m - c_l C_m = 0, \quad s_r S_m + c_r C_m = 0, \quad (2.17)$$

$$s_l = M \lambda_l \cos \lambda_l \pi + \alpha \coth \alpha G \sin \lambda_l \pi, \quad (2.18a)$$

$$c_l = M \lambda_l \cos \lambda_l \pi + \alpha \tanh \alpha G \sin \lambda_l \pi. \quad (2.18b)$$

Here  $s_r$  and  $c_r$  are of the same form as  $c_l$  and  $s_l$ , but with the subscript on all terms changed from ‘‘l’’ to ‘‘r.’’ Equation (2.17) has a nontrivial solution if and only if  $s_l c_r + s_r c_l = 0$ , which yields the eigenvalue problem in its full form:

$$\begin{aligned} \alpha M (\coth \alpha G + \tanh \alpha G) (\lambda_r \cos \lambda_r \pi \sin \lambda_l \pi + \lambda_l \cos \lambda_l \pi \sin \lambda_r \pi) \\ + 2M^2 \lambda_l \lambda_r \cos \lambda_l \pi \cos \lambda_r \pi + 2\alpha^2 \sin \lambda_l \pi \sin \lambda_r \pi = 0. \end{aligned} \quad (2.19)$$

If (2.19) has a nontrivial solution, then the two equations in (2.17) are redundant. Hence we may solve the first equation for  $S_m$  in terms of  $C_m$  and use the result in (2.15) to obtain

$$r = \left| \frac{(s_l + c_l) \sin \lambda_l \pi}{(s_l - c_l) \sin \lambda_r \pi} \right|. \quad (2.20)$$

### 3 Asymptotic limits

To better understand the localization of the wavefunction in the generic case, we first examine several asymptotic limits.

#### 3.1 Localized limits

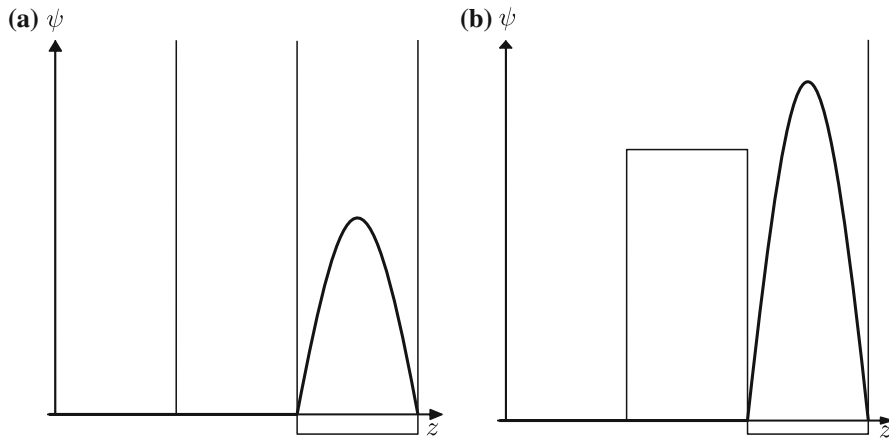
**Large  $H$ .** In the limit  $H \rightarrow \infty$ ,  $\alpha \rightarrow \infty$  while  $M$  remains finite. Thus the leading order of (2.19) is

$$2\alpha^2 \sin \lambda_l \pi \sin \lambda_r \pi = 0. \quad (3.1)$$

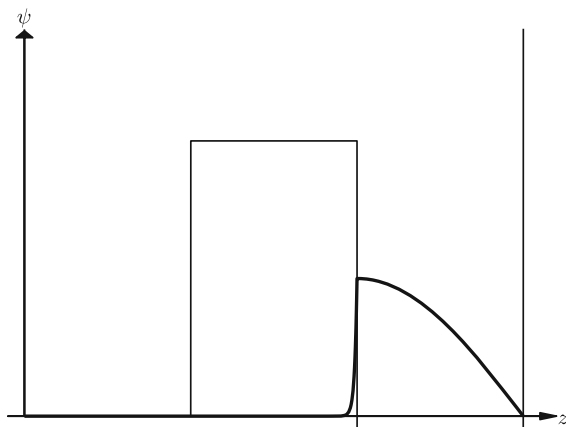
Equation (3.1) has two classes of solutions. The first has  $\sin \lambda_l \pi = 0$ , which yields an infinite set of solutions (namely, the positive integers). However, by (2.14a) we see that higher values of  $\lambda_l$  correspond to larger values of  $E$ . These higher energy solutions correspond to excited states confined within individual QDs. Charges confined in high-energy states rapidly relax to the lowest energy (ground) state, and it is the degree of delocalization of these ground states that is relevant to QD-based photovoltaics.

Consequently, we take  $\lambda_l = 1$ . Then by (2.16) we have that  $u_m = u_r = 0$ , so the wavefunction is localized in the left dot, and hence  $r = 0$ . This value of  $\lambda_l$  corresponds to the theorized infinite-well value of  $E = 1$ , as expected since the wavefunction is confined to one dot.

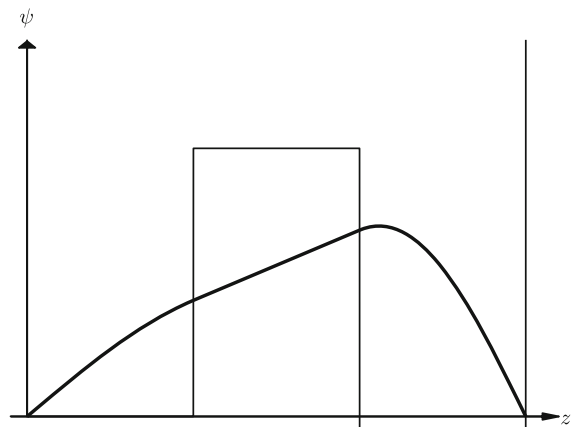
The other solution of (3.1) under consideration is  $\lambda_r = 1$ , in which case  $u_l = u_m = 0$ , so the wavefunction is localized in the right dot, and hence  $r = \infty$ . This case is shown in Fig. 4a. Physically, the large potential barrier keeps the wavefunction confined to one dot.



**Fig. 4** Schematic of wavefunction localized in right dot when **a**  $H \rightarrow \infty$  and **b**  $M \rightarrow 0$



**Fig. 5** Schematic of wavefunction localized in right dot when  $M \rightarrow \infty$



**Fig. 6** Schematic of wavefunction delocalized when  $\alpha \rightarrow 0$

**Small  $M$ .** In the limit  $M \rightarrow 0$ ,  $M$  vanishes faster than  $\alpha$ , so the leading order of (2.19) is still given by (3.1). Hence we once again have the wavefunction isolated in one dot or the other, as shown in Fig. 4b. Physically, when the effective mass tends to zero in the barrier, the momentum operator must also tend to zero to keep the velocity operator  $O(1)$  to match with the finite velocity operator in the dot. This forces  $u_m = 0$  in the barrier, just as in the infinite-potential case. Hence the wavefunction is again confined to one dot.

**Large  $M$ .** In the limit  $M \rightarrow \infty$ ,  $M$  diverges faster than  $\alpha$ , so the leading order of (2.19) is

$$2M^2 \lambda_l \lambda_r \cos \lambda_l \pi \cos \lambda_r \pi = 0. \tag{3.2}$$

One solution of (3.2) is given by  $\lambda_r = 1/2$ . Then from (2.16) we have that

$$\frac{du_r}{dz}(G) = -\frac{1}{2}(S_m + C_m) \sin \lambda_l \pi \cos \frac{\pi}{2} = 0, \tag{3.3}$$

and hence the derivative of the wavefunction is zero at the interface (Fig. 5).

Next we examine the behavior in the barrier. When  $M \rightarrow \infty$ , then  $\alpha \rightarrow \infty$ , and the leading order of (2.13) is  $u_m = 0$ . Thus there must be a boundary layer near the interface (Fig. 5). In the boundary layer,  $du_m/dz = O(M^{1/2})$ ,

which can be seen from (2.13) or (2.16). Hence (2.12b) becomes

$$\frac{du_r}{dz}(G^+) = O(M^{-1/2}) \rightarrow 0, \quad (3.4)$$

consistent with (3.3). Could such a boundary layer exist near  $z = -G^+$ ? No, because it would force  $\lambda_1 = 1/2$ , which violates (2.14b) since  $\delta \neq 0$ . In particular, we see that with  $\lambda_1 \neq 1/2$ ,  $s_1 = c_1$  in the limit  $M \rightarrow \infty$  [from (2.18)], which means that  $S_1 = C_1$  [from (2.17)], and hence  $u_1 \equiv 0$  [from (2.16)]. This is consistent with the fact that with no boundary layer in the barrier,  $u_1$  must satisfy the following conditions:

$$u_1(-G^-) = \frac{du_1}{dz}(-G^-) = 0. \quad (3.5)$$

Thus the wavefunction is localized in the right dot. If instead  $\lambda_1 = 1/2$ , then the wavefunction is localized in the left dot and the diagram is the mirror image of Fig. 5.

Physically, since the effective mass is large in the barrier, the velocity operator will vanish there. Since the effective mass is finite in the dot, the momentum operator (2.6) must vanish on the dot side of the interface to enforce continuity of the velocity operator. Hence  $du_m/dz$  must vanish there.

Note that the preceding three asymptotic limits all resulted in the wavefunction's being localized to a single dot. Now we consider two limits that result in the wavefunction's being delocalized.

### 3.2 Delocalized limits

**Small  $\alpha$ .** If  $H \rightarrow E^+$ , then  $\alpha \rightarrow 0$  while  $M$  remains finite. Hence (2.13) becomes

$$\frac{d^2 u_m}{dz^2} = 0,$$

which causes the boundary conditions in (2.12b) to become

$$\frac{du_1}{dz}(-G^-) = \frac{du_r}{dz}(G^+).$$

Moreover, since the slope of  $u_m$  (and hence the momentum flux) is constant in the barrier, from a physical perspective for all practical purposes the barrier is not there. (This is consistent with our choice of limit  $H \rightarrow E^+$ , which means that the barrier provides no resistance to the wavefunction.) Hence the wavefunction is delocalized (Fig. 6).

$\delta = 0$ . When  $\delta = 0$ , the problem is symmetric by (2.13). Hence  $u_1 = \pm u_r$ . This limit does not affect  $u_m$ , so the wavefunction will decay in the barrier as governed by the other parameters (Fig. 7).

## 4 General discussion of asymptotic analysis

To achieve wavefunction delocalization, experimentalists must keep  $r$  near 1 even in the case where  $\delta \neq 0$ . To see why that is so difficult, note from (2.18) that

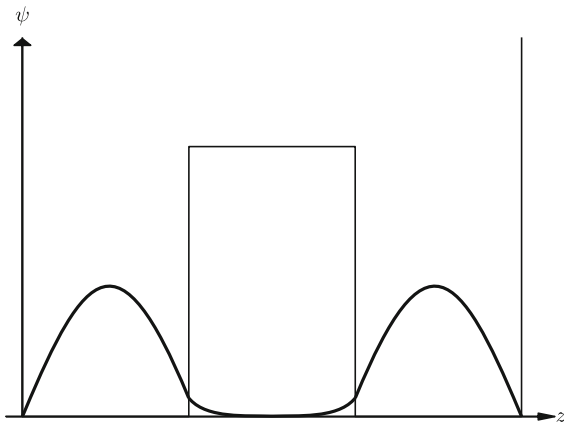
$$s_1 - c_1 = \epsilon_1, \quad \epsilon_1 = \frac{2\alpha \sin \lambda_1 \pi}{\sinh 2\alpha G}. \quad (4.1)$$

Substituting (4.1) into (2.20), we obtain

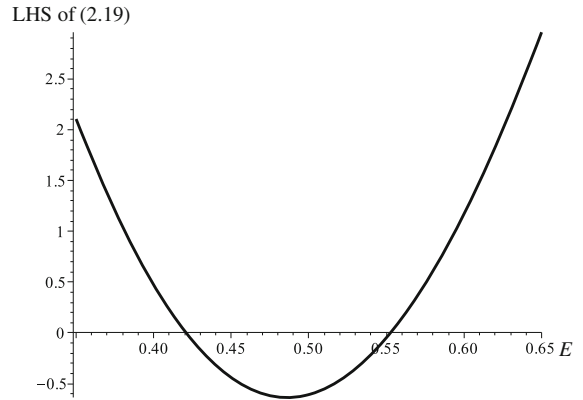
$$r = \left| \frac{(2c_1 + \epsilon_1) \sin \lambda_1 \pi}{\epsilon_1 \sin \lambda_r \pi} \right|. \quad (4.2)$$

But  $\epsilon_1$  is extremely small for even moderate  $\alpha G$ . Hence by (4.2) we have that  $r \rightarrow \infty$ , which corresponds to a highly localized wavefunction.





**Fig. 7** Schematic of symmetric wavefunction delocalized when dots are symmetric



**Fig. 8** Plot of left-hand side of (2.19) versus  $E$  using the parameters in Table 2. Intersections with the horizontal axis are allowable energy levels

To demonstrate this, we plot the left-hand side of (2.19) in Fig. 8 for the parameters in Table 2. Note that there are two values of  $E$  bunched close together, as predicted physically for weakly coupled dots with a small potential difference between them. In particular, the allowable energy levels are

$$E_{<} = 4.21 \times 10^{-1}, \quad E_{>} = 5.53 \times 10^{-1}. \tag{4.3}$$

For the lower value of  $E$ , we have  $\epsilon_1 = 1.44 \times 10^{-4}$ , which is small, as theorized. (The localization is similarly strong for the higher energy value, as shown subsequently.)

For generic  $H$  and  $M$ , localization of the wavefunction occurs when  $\epsilon_1 = o(c_1)$  [when by (4.2)  $r \rightarrow \infty$ ]. Unfortunately, this case is the current experimental state of the art. If we substitute the parameters in Table 2 and the lower value of the energy level in (4.3) into (2.18b), we have  $c_1 = 1.1290$ , which is much larger than the value of  $\epsilon_1$  given previously. Hence it is no surprise that when we substitute the values of  $c_1$  and  $\epsilon_1$  into (4.2), we obtain

$$r_{<} = 1.97 \times 10^4, \tag{4.4a}$$

which corresponds to the wavefunction's being localized in the right dot. Using the larger value of  $E$  leads to the wavefunction's being localized in the left dot because we have

$$r_{>} = 7.42 \times 10^{-5}. \tag{4.4b}$$

The preceding idea shows that if  $\epsilon_1 = O(c_1)$ , then we will have some level of delocalization. To track the transition between the delocalized case  $r = 1$  and the localized case where  $r = 0$  or  $r = \infty$ , we must consider this intermediate case. As discussed previously, we choose  $\delta$  as the perturbation parameter for our analysis. As can be seen from (2.14b), in the limit  $\delta \rightarrow 0$ ,  $\lambda_l \rightarrow \lambda_r$ . Hence to leading order we drop subscripts if the variables have them and use a subscript zero if they do not:

$$\lambda_{\cdot} = \lambda + \lambda_{\cdot,1}\delta + o(\delta), \quad \alpha = \alpha_0 + \alpha_1\delta + o(\delta), \tag{4.5}$$

where the dot refers to “l” or “r.” Note that, though there is an explicit  $\delta$ -dependence in  $\lambda_r$ ,  $\lambda_l$  also depends on  $\delta$  implicitly through  $E$ .

We must next relate  $\epsilon_1$  to  $\delta$ . The difference between  $s_1$  and  $c_1$  is largely dictated by the difference between  $\coth \alpha G$  and  $\tanh \alpha G$ . But even for moderate  $\alpha G$  we have

$$\coth \alpha G \sim 1 + 2e^{-2\alpha G}, \quad \tanh \alpha G \sim 1 - 2e^{-2\alpha G}, \tag{4.6a}$$

so we let

$$\epsilon = 2e^{-2\alpha_0 G}. \tag{4.6b}$$

Taking  $\epsilon \rightarrow 0$  implies that  $G \rightarrow \infty$ , yielding the limiting case of isolated wells from the standard theory. Then we relate  $\epsilon$  to  $\delta$  by letting

$$\epsilon = \epsilon_0 \delta, \quad \epsilon_0 = O(1). \tag{4.7}$$

Using these definitions, we have the following expression for  $s_1$ , to leading two orders:

$$s_1 = M\lambda \cos \lambda\pi + \alpha_0 \sin \lambda\pi + \delta \left\{ M \left[ \lambda_{1,1} \cos \lambda\pi - \lambda \lambda_{1,1} \pi \sin \lambda\pi \right] + (\alpha_0 \epsilon_0 + \alpha_1) \sin \lambda\pi + \alpha_0 \lambda_{1,1} \pi \cos \lambda\pi \right\}. \tag{4.8}$$

Since the case of interest has  $c_1 = O(\epsilon_1)$ , we must have that  $s_1 = O(\epsilon) = O(\delta)$ . This forces the leading-order term of (4.8) to be zero, giving a relationship between  $\lambda$  and  $\alpha_0$ :

$$\lambda \cot \lambda\pi = -\frac{\alpha_0}{M}. \tag{4.9}$$

We also note from (2.16) that

$$\alpha_0 = \sqrt{M(H - \lambda^2)}. \tag{4.10}$$

Hence upon substituting (4.10) into (4.9), we have a single equation that can be solved for  $E$ .

### 5 Next-order correction, infinite-height barriers

Since we have zeroed out the leading order of  $s_1$  using (4.9), we obtain

$$s_1 = \delta(a_1 + \alpha_0 \epsilon_0 \sin \lambda\pi), \tag{5.1a}$$

$$a_1 = (\alpha_1 - M\lambda \lambda_{1,1} \pi) \sin \lambda\pi + (M + \alpha_0 \pi) \lambda_{1,1} \cos \lambda\pi. \tag{5.1b}$$

Given the fact that  $s_r$  is the same as  $s_1$  with a change of subscript, we have

$$s_r = \delta(a_r + \alpha_0 \epsilon_0 \sin \lambda\pi), \tag{5.2}$$

where we have used the fact that the leading order of both  $s$  terms is the same, and we have defined  $a_r$  similar to  $a_1$ .

Expanding the  $c$  terms in a similar matter, we have from (2.18) that the only difference is the change from  $\coth$  to  $\tanh$ , which from (4.6) just means that  $-\epsilon_0$  replaces  $\epsilon_0$ . Hence when (4.9) is satisfied, the leading order still vanishes and we have

$$c_1 = \delta(a_1 - \alpha_0 \epsilon_0 \sin \lambda\pi), \quad c_r = \delta(a_r - \alpha_0 \epsilon_0 \sin \lambda\pi).$$

With these definitions, (2.20) becomes

$$r = |r_*|, \quad r_* = \frac{a_1}{\alpha_0 \epsilon_0 \sin \lambda\pi}, \tag{5.3}$$

where we have used the fact that to leading order,  $\lambda_1 = \lambda_r$ . In the limit that  $\epsilon_0 \rightarrow \infty$  ( $\delta = o(\epsilon)$ ), we get  $r = 1$  consistent with the  $\delta = 0$  case discussed in Sect. 3. If, instead,  $\epsilon_0 \rightarrow 0$  ( $\epsilon = o(\delta)$ ), then  $r \rightarrow \infty$ , and the wavefunction is localized.

Now satisfying the determinant condition (2.19), we have

$$2a_1 a_r - 2\alpha_0^2 \epsilon_0^2 \sin^2 \lambda\pi = 0$$

and hence

$$\frac{a_r}{a_1} r_*^2 - 1 = 0. \tag{5.4}$$

This expression depends on both  $\lambda_{r,1}$  and  $\lambda_{1,1}$ , so we relate them by noting that

$$\lambda_{r,1} = \lambda_{1,1} + \frac{1}{2\lambda}. \tag{5.5}$$

Substituting (5.5) into (5.1b) adapted for  $a_r$ , we obtain the following relationship between  $a_r$  and  $a_1$ :

$$a_r = a_1 \left( 1 - \frac{b}{r_*} \right), \tag{5.6a}$$

$$b = \frac{1}{2\epsilon_0} \left( \frac{M + \alpha_0 \pi}{M\lambda^2} + \frac{M\pi}{\alpha_0} \right), \tag{5.6b}$$

where we have used (4.9). Substituting (5.6a) into (5.4) yields a simple quadratic equation for  $r_*$ :

$$r_*^2 - br_* - 1 = 0, \quad (5.7a)$$

and hence

$$r_* = \frac{b \pm \sqrt{b^2 + 4}}{2}. \quad (5.7b)$$

Here there are two solutions, corresponding to the two different eigenvalues (energy levels) and wavefunction delocalizations. Since we could easily switch the roles of the left and right dots, the two values of  $r$  should be reciprocals of one another. This is confirmed by (5.7a), which indicates that the product of the two solutions for  $r_*$  is  $-1$ .

Substituting the parameters in Table 2 into (4.9) yields

$$E = 5.53 \times 10^{-1}, \quad (5.8)$$

which is simply  $E_>$  from (4.3). This can be seen by noting that the second equations in (2.14) may be rewritten as

$$E_> = \lambda_l^2, \quad E_< = \lambda_r^2 - \delta. \quad (5.9)$$

Hence when  $\delta = 0$ , the two energy levels converge on the higher level. The lower potential energy in the perturbed well (modeled with  $\delta > 0$ ) then allows a lower energy level to be obtained. We may also calculate  $\epsilon = 1.14 \times 10^{-4}$ , which is small, consistent with the results in Sect. 4. Calculating the  $r$  values, we substitute the relevant parameters into (5.7b) to obtain

$$r_< = 1.30 \times 10^4, \quad r_> = 7.5 \times 10^{-5}. \quad (5.10)$$

These values (especially  $r_<$ ) are not as close to the values in (4.4) as one might expect. The discrepancy can be explained by noting that our approximate theory uses the assumption that  $\epsilon_0 = O(1)$ , while for the actual parameters we have  $\epsilon_0 = 8.14 \times 10^{-4}$ .

## 6 Bounds in infinite-height case

Given that the current state of experiments leads to undesirably localized wavefunctions, we use the perturbation theory to provide guidance as to how physical parameters can be changed to ensure delocalization. Therefore, we write the parameters in the full problem in terms of  $r_*$ , the delocalization parameter. Solving (5.7a) for  $b$ , we have

$$b = \frac{r_*^2 - 1}{r_*}. \quad (6.1)$$

We substitute this result into (5.6b) and solve for  $\delta$ . Then we use (2.8b) to rewrite our expression in terms of the true experimental tolerance  $\Delta\tilde{H}$ , yielding

$$\Delta\tilde{H} = 4E_c e^{-2\alpha_0 G} \left( \frac{r_*^2 - 1}{r_*} \right) \left[ \frac{\alpha_0 M \lambda^2}{\pi(\alpha_0^2 + \lambda^2 M^2) + \alpha_0 M} \right], \quad (6.2)$$

where we have used (4.6b) and (4.7).

We examine each factor in turn. By (2.8a) we have that  $E_c$ , the characteristic energy level, depends only on the material properties of the dot. The parenthetical term depends only on the bound we set. The remaining terms depend on the whole system through the eigenvalue  $\lambda$ . Obviously the most critical dependence is on the exponential term, which is the only term that depends on  $G$ .

Figure 9 shows a graph of (6.2) using (5.8) and various values of  $G$ . [Recall that (5.8) is independent of  $G$  since  $G$  is used only to determine the size of  $\epsilon_0$ .] The left endpoint corresponds to a peak in the second well equal to 60% of that in the first. With a typical value of 10 nm for the barrier (middle line),  $\Delta\tilde{H}$  would have to be less than 0.01 meV to achieve this bound. With the thinnest barrier (3 nm) currently available, a tolerance of only around 10 to 20 meV is required. Hence, as discussed previously, the solution is highly sensitive to the value of  $G$ .

Though we have derived a relationship for a two-dot system, realistic QDAs may have 50 dots or more. For simplicity, we focus on the case where  $r < 1$ . In the two-dot system, the amplitude in one dot is roughly  $r$  times the amplitude in the other. Now consider an  $N$ -dot array where  $\psi$  is at its maximum in dot  $j$ . (Figure 10 shows a schematic where  $N = 5$  and  $j = 3$ .) If we examine the dots in pairwise fashion, it is reasonable to assume that the amplitude in one dot would be  $r$  times smaller than the amplitude in its nearest neighbor. Extending this result, we have that

$$\frac{\text{amplitude in dot } j \pm i}{\text{amplitude in dot } j} = r^i. \tag{6.3}$$

We consider a wavefunction to be delocalized if it is “spread out” over several dots, which is equivalent to its amplitude being “large enough” in those dots. Formally, we define the *delocalization probability cutoff*  $p$  such that if the probability ratio in a certain dot exceeds  $p$ , we say that the wavefunction is delocalized to that dot. Hence the wavefunction with a maximum at dot  $j$  is delocalized to dot  $j \pm i$  if

$$r^{2i} \geq p, \tag{6.4a}$$

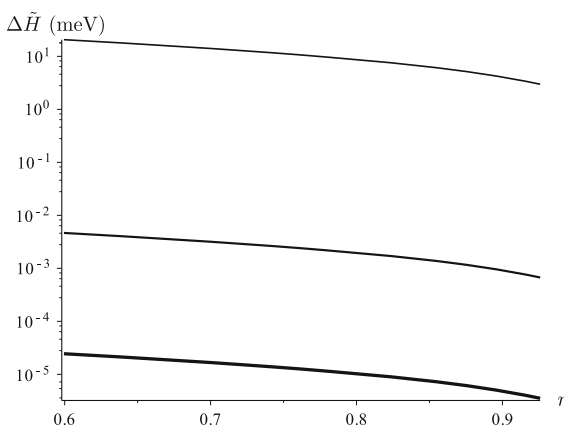
where the 2 in the exponent comes from the fact that we are analyzing the probability amplitude, which is the square of the wavefunction amplitude. Note from (5.7b) that if  $r < 1$ , then  $r_* < 0$ , and hence the critical value of  $r_*$  in (6.4a) is

$$r_* = -p^{1/2i}. \tag{6.4b}$$

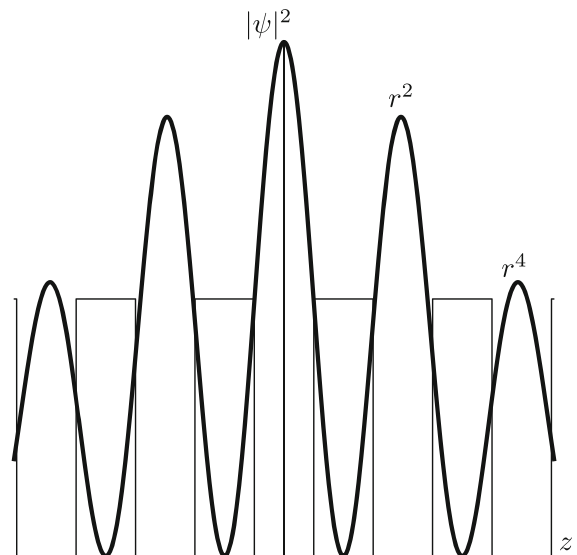
Given a cutoff  $p$ , we then define the *delocalization length*  $n$ . Here  $n$  is defined as the number of dots over which the wavefunction is delocalized. To calculate  $n$ , note that if the wavefunction’s maximum is far enough from the boundaries, the amplitude decay is symmetric and  $n = 2i + 1$ , where  $i$  is the largest integer satisfying (6.4a). Hence we define

$$r_o = -p^{1/(n-1)} \tag{6.5a}$$

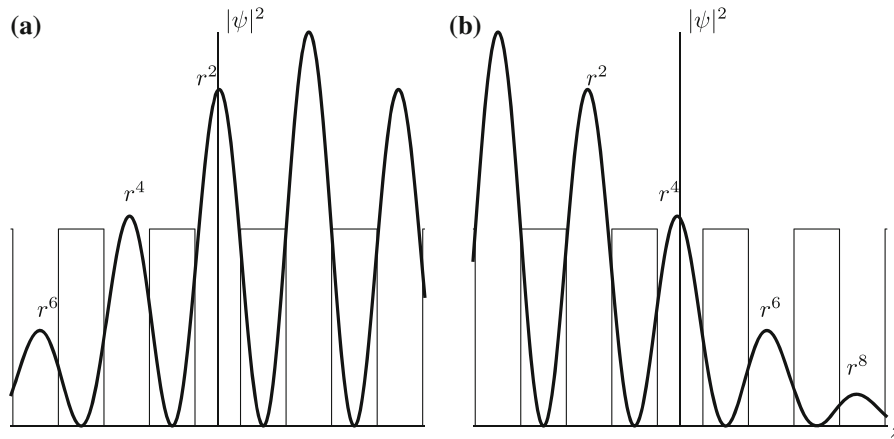
as the value of  $r_*$  needed to guarantee that *at least one* wavefunction has delocalization length  $n$ . Here the subscript “o” stands for “one.”



**Fig. 9** Plot of experimental tolerance (6.2) versus  $r = |r_*|$ . In increasing order of thickness: minimum value of  $G$  from Table 2,  $G$  corresponding to a 10-nm barrier, maximum value of  $G$  from Table 2

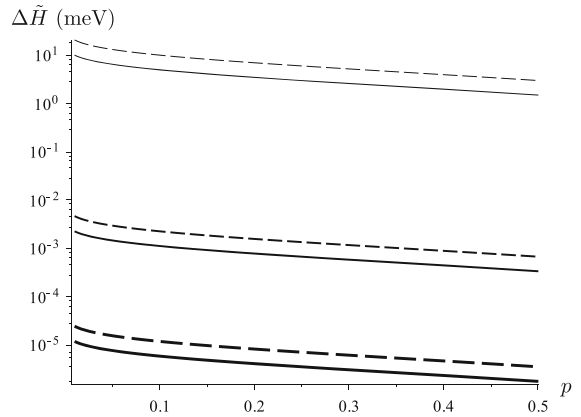


**Fig. 10** Schematic of five-dot system with peak in dot 3. The ratio is squared since we are considering the probability amplitude



**Fig. 11** Schematics of five-dot system with peaks in dot 4 (a) and 1 (b). The ratio is squared since we are considering the probability amplitude

**Fig. 12** Plot of one-wave-delocalized bound (6.6a) (dashed lines) and all-waves-delocalized bound (6.6b) (solid lines) versus  $p$  for  $n = 10$ . In increasing order of thickness: minimum value of  $G$  from Table 2,  $G$  corresponding to a 10-nm barrier, maximum value of  $G$  from Table 2



However, the wavefunction may also have its maximum near the boundary, as shown in Fig. 11. The worst-case scenario is shown in Fig. 11b. When the peak adjoins a boundary, the entire delocalization must happen at one side, so  $n = i + 1$ , where  $i$  is the largest integer satisfying (6.4a). Hence we define

$$r_a = -p^{1/(2n-2)} \tag{6.5b}$$

as the value of  $r_*$  needed to guarantee that *all* wavefunctions have delocalization length  $n$ . Here the subscript “a” stands for “all.” Note that  $|r_a| < |r_o|$  because it is the bound for the more restrictive case.

Substituting (6.5) into (6.2) yields corresponding bounds on  $\Delta \tilde{H}$ :

$$\Delta \tilde{H}_o = 4E_c e^{-2\alpha_0 G} \left[ \frac{1 - p^{2/(n-1)}}{p^{1/(n-1)}} \right] \left[ \frac{\alpha_0 M \lambda^2}{\pi(\alpha_0^2 + \lambda^2 M^2) + \alpha_0 M} \right], \tag{6.6a}$$

$$\Delta \tilde{H}_a = \Delta \tilde{H}_o \left[ \frac{p^{1/(2n-2)}}{1 + p^{1/(n-1)}} \right]. \tag{6.6b}$$

As expected,  $\Delta \tilde{H}_a < \Delta \tilde{H}_o$  because it is the bound for the more restrictive case.

Figure 12 shows a graph of (6.6) using (5.8) and various values of  $G$ . Using  $\Delta \tilde{H}_o$  relaxes the bound by a factor of approximately 2.5 relative to  $\Delta \tilde{H}_a$ . The graph for each  $G$  is relatively flat since  $p$  values in this range correspond to a much smaller range of  $r_*$ . In particular, the graph has

$$0.01 \leq p \leq 0.5 \implies -0.93 \leq r_o \leq -0.60, \quad -0.96 \leq r_a \leq -0.77.$$

With  $n = 10$ , the range of  $p$  in Fig. 12 corresponds exactly to the range in  $r$  in Fig. 9. The difference in concavity between the two figures is due to the nonlinear scaling now used for the horizontal axis.

Figure 12 illustrates the limitations of current fabrication technology. The left endpoint corresponds to a 1% probability that an electron wavefunction is delocalized over a 10-dot array. With a typical value of 10nm for the barrier (middle line),  $\Delta\tilde{H}$  would have to be less than 0.01 meV. (Contrast that with the current tolerance of 56.5 meV stated in Table 1.) Fortunately, with the thinnest barrier (3 nm) currently available, a tolerance of only around 10 to 20 meV is required. Due to this high sensitivity to  $G$ , if barriers even slightly thinner than 3 nm can be constructed while handling the concomitant strain issues, delocalization is possible with current experimental tolerances in  $\Delta\tilde{H}$ .

We next compare our analytical results with numerical simulations. In [11], a series of  $N$  dots was constructed with potentials normally distributed with mean zero and standard deviation  $\Delta\tilde{H}$  as given in Table 1. Then the first  $N$  wavefunctions (each one having a peak in a different dot) were computed, along with their delocalization lengths. The delocalization length was computed by measuring the distance at which the threshold value  $p$  was reached and dividing by the length of the dot. Hence the computed values of  $n$  were decimals and could be less than 1 (Fig. 14). The average of each of those  $N$  lengths was then computed to produce  $\langle n \rangle$ , the *average delocalization length of the array*. This numerical experiment was repeated over many realizations, and the average of all those realizations was reported in the manuscript.

How then to compute  $\langle n \rangle$  in our circumstance? Consider the five-dot case diagrammed in Figs. 10 and 11, and assume that  $r^4 = p$ . Then the wavefunction in Fig. 10 would have  $n = 5$ , the wavefunction in Fig. 11a (and its mirror image) would have  $n = 4$ , and the wavefunction in Fig. 11b (and its mirror image) would have  $n = 3$ . Hence in this case  $\langle n \rangle = 19/5$ .

Generalizing this result, each wavefunction has  $n = 2i + 1$  unless that width impinges on a boundary. As the peak approaches the boundary, there will be two wavefunctions (one on the left, one on the right) where the last dot over which it would be delocalized extends outside the boundary. These two dots would have  $n = 2i = (2i + 1) - 1$ . Similarly, there will be two with  $n = (2i + 1) - 2, n = (2i + 1) - 3$ , etc., until finally, when dot  $j$  adjoins the boundary, those two dots will have  $n = i + 1 = (2i + 1) - i$ . Hence we have

$$\langle n \rangle = 2i + 1 - \frac{1}{N} [2(1) + 2(2) + \dots + 2(i)] = 2i + 1 - \frac{i(i + 1)}{N}, \quad 1 \leq i \leq N. \tag{6.7}$$

Note that letting  $i = 2, N = 5$  in (6.7) yields  $\langle n \rangle = 19/5$ , consistent with the preceding discussion. If  $i = 0$ , then  $\langle n \rangle = 1$  (each wavefunction confined to its own dot). Also, as  $N \rightarrow \infty$ , the edge effects become negligible, and the average quickly approaches the two-sided value.

The derivation of (6.7) implicitly assumed some wavefunctions could be delocalized to their full extent, in other words, that  $2i + 1 \leq N$ . However, even in the case where  $2i + 1 > N$ , the formula still holds. That is because the same overlap calculation applies, but now *every* wavefunction has an overlap, and some may have overlap on both sides. But the counting works out the same, and (6.7) still holds.

Again considering the five-dot case as an example, let us assume that  $r^6 = p$ , so  $i = 3$ . In that case, the wavefunction in Fig. 11b (and its mirror image) would have  $n = 4$ . The remaining wavefunctions would be delocalized over the whole array, so they would have  $n = 5$ . Hence  $\langle n \rangle = 23/5$ , which is exactly the result one obtains by substituting  $i = 3, N = 5$  into (6.7). Moreover, we note that if  $i = N$ , then  $\langle n \rangle = N$  (each wavefunction delocalized over the whole array). This is also the case if  $i > N$ , so we have

$$\langle n \rangle = N, \quad i > N. \tag{6.8}$$

Therefore, to compare our analytical solutions to the numerical results in [11], we first write the corresponding form of (6.2) for general  $i$  (which for the purposes of the plot we treat as a continuous variable):

$$\Delta\tilde{H} = 4E_c e^{-2\alpha_0 G} \left( \frac{1 - p^{1/i}}{p^{1/2i}} \right) \left[ \frac{\alpha_0 M \lambda^2}{\pi(\alpha_0^2 + \lambda^2 M^2) + \alpha_0 M} \right]. \tag{6.9}$$

Then we simply plot (6.9) on the  $x$ -axis and (6.7) on the  $y$ -axis, as shown in Fig. 13. There is reasonable agreement between theory and experiment, even in the unphysical case of infinite side walls.

In general, for moderate values of  $\langle n \rangle$  the analytical solutions provide more conservative predictions than the numerical simulations. This can be explained by noting that while the analytical solution assumes that every well is offset by  $\Delta\tilde{H}$ , the numerical simulations have a range of potential levels whose standard deviation is  $\Delta\tilde{H}$ . Hence most of the dots would have smaller offsets, which would then increase delocalization.

For small values of  $\langle n \rangle$  the numerical solutions are more conservative. This is more pronounced in Fig. 14. In this graph,  $\Delta\tilde{H}$  and  $p$  are fixed, and  $\langle n \rangle$  is plotted as a function of the barrier width  $2\tilde{G}$ . Note that with the change in scale, we are now focusing on small values of  $\langle n \rangle$ , which is where the analytical results were overestimates in Fig. 13.

In such a regime, the delocalization length is going to be highly sensitive to the particular values of  $\Delta\tilde{H}$  in the few dots over which the delocalization occurs. Hence a large value of  $\Delta\tilde{H}$ , randomly selected, can shut down delocalization entirely. Recall also that due to the way  $\langle n \rangle$  is calculated in [11], these will contribute values smaller than 1 to the average, hence reducing it beyond what the analytical results will allow.

### 7 Finite barriers

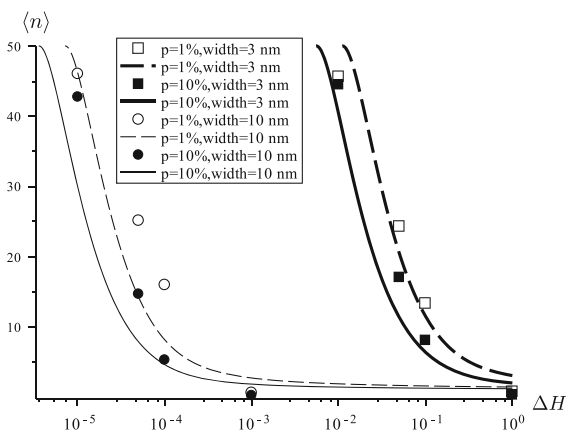
A true experimental system will have finite barriers as  $|z| \rightarrow \infty$ . To model such a case, we adapt the previous problem as shown in Fig. 15. To point out relationships with the infinite-height case, we begin by taking the external barriers to be of arbitrary (but equal) height  $H_{\pm}$ . The true solar cell has  $H_{\pm} = H$ .

Since (2.11) is now replaced by continuity conditions that match the wells to the external barriers, the expressions in (2.14) must become more general. To preserve the interpretation of  $S$  as an amplitude so that we may continue to use the interpretation of  $r$  in (2.15), we generalize the infinite-barrier solutions through an arbitrary phase shift  $\phi$ :

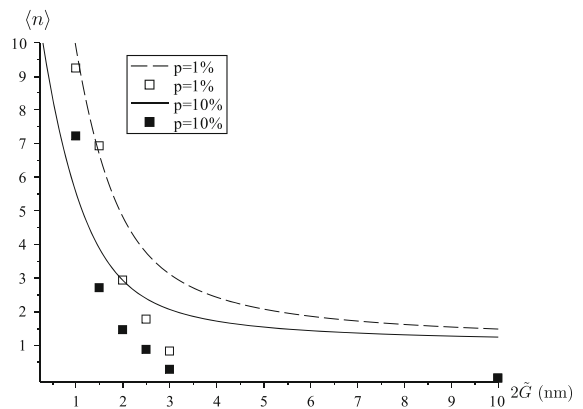
$$u_l = S_l \sin \lambda_l (z + \pi + G + \phi_l), \tag{7.1a}$$

$$u_r = S_r \sin \lambda_r (\pi + G + \phi_r - z). \tag{7.1b}$$

As in Sect. 2, the solution to (2.13) in the outer barriers will be an exponential. As in (2.12a), the function values at  $z = \pm(G + \pi)$  must be continuous. Hence we have



**Fig. 13** Lines Plot of (6.7) versus (6.9) for  $N = 50$ . Points Results from computations in [11]. Here the  $x$ -axis is  $\Delta H$  normalized by the higher value of  $\Delta\tilde{H}$  given in Table 1



**Fig. 14** Lines Plot of (6.7) versus barrier width for the infinite-barrier case. Points Results from computations in [11]. Note the numerical values with  $\langle n \rangle < 1$

$$u_- = S_l \sin \lambda_1 \phi_l \exp(\alpha_{\pm}(z + \pi + G)), \quad \alpha_{\pm} = \sqrt{M(H_{\pm} - E)}, \tag{7.2a}$$

$$u_+ = S_r \sin \lambda_r \phi_r \exp(\alpha_{\pm}(\pi + G - z)), \tag{7.2b}$$

where  $u_{\pm}$  holds as  $z \rightarrow \pm\infty$ . As in Sect. 2, we choose these forms to illustrate clearly the symmetry of the problem.

Similar to (2.12b), at the interfaces  $z = \pm(G + \pi)$  the derivative of the wavefunction in the GaAs must be  $M$  times the derivative of the wavefunction in the dot. Satisfying these conditions, we obtain

$$\tan \lambda_l \phi_l = \frac{\lambda_l M}{\alpha_{\pm}}, \quad \tan \lambda_r \phi_r = \frac{\lambda_r M}{\alpha_{\pm}}. \tag{7.3}$$

Equations (7.3) then define the phase shifts  $\phi$  in terms of the other parameters in the problem. In the case of an infinite wall,  $H_{\pm} \rightarrow \infty$ , so  $\alpha_{\pm} \rightarrow \infty$  and the right-hand side of (7.3) becomes zero, which forces the  $\phi$  to 0 as in Sect. 2.

With (7.1) replacing (2.14), the rest of the analysis holds once we insert the phase. In particular,  $r$  becomes

$$r = \left| \frac{(s_l + c_l) \sin \lambda_l(\pi + \phi_l)}{(s_l - c_l) \sin \lambda_r(\pi + \phi_r)} \right|, \tag{7.4}$$

which replaces (2.20). Moreover,  $s_l$  and  $c_l$  become

$$s_l = M \lambda_l \cos \lambda_l(\pi + \phi_l) + \alpha \coth \alpha G \sin \lambda_l(\pi + \phi_l), \tag{7.5a}$$

$$c_l = M \lambda_l \cos \lambda_l(\pi + \phi_l) + \alpha \tanh \alpha G \sin \lambda_l(\pi + \phi_l). \tag{7.5b}$$

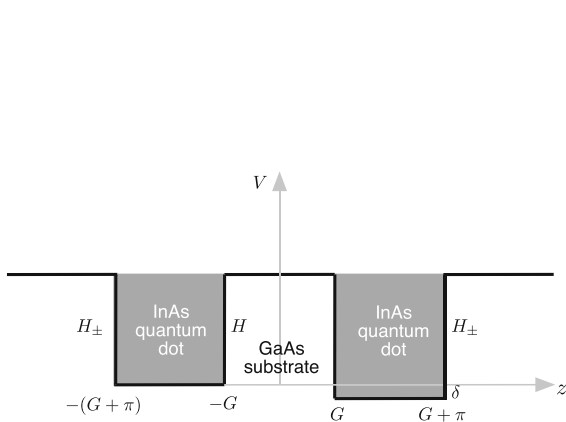
Hence the eigenvalue condition (2.19) becomes

$$\alpha M (\coth \alpha G + \tanh \alpha G) [\lambda_r \cos \lambda_r(\pi + \phi_r) \sin \lambda_l(\pi + \phi_l) + \lambda_l \cos \lambda_l(\pi + \phi_l) \sin \lambda_r(\pi + \phi_r)] + 2M^2 \lambda_l \lambda_r \cos \lambda_l(\pi + \phi_l) \cos \lambda_r(\pi + \phi_r) + 2\alpha^2 \sin \lambda_l(\pi + \phi_l) \sin \lambda_r(\pi + \phi_r) = 0. \tag{7.6}$$

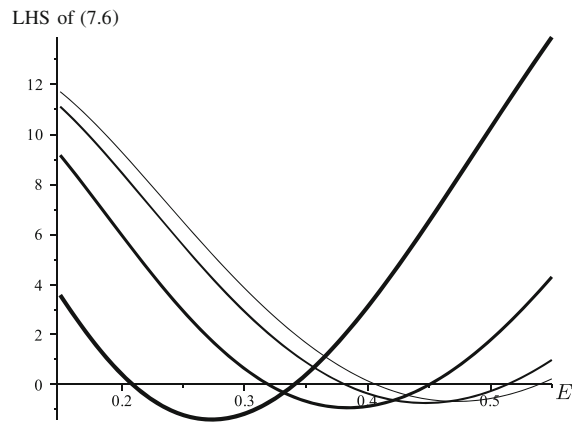
As indicated previously, in a real solar cell,  $\alpha_{\pm} = \alpha$ . In Fig. 16, we demonstrate the effect of finite-height side walls on our solution. In particular, we plot the left-hand side of (7.6) versus  $E$ . The thickest line shows the true solution in the solar cell, which yields the energy levels

$$E_{<} = 2.09 \times 10^{-1}, \quad E_{>} = 3.40 \times 10^{-1}, \tag{7.7}$$

which are shifted significantly from the infinite-height case in (4.3). We then show the effect of increasing the side-wall height upon our solutions. Note that when  $\alpha_{\pm} = 27\alpha$  (the thinnest line), we are nearly in the case where  $\alpha_{\pm} = \infty$  in Sect. 4, as shown in Fig. 8.



**Fig. 15** Dimensionless-variable description of potential in two-dot array with finite side walls



**Fig. 16** Plot of left-hand side of (7.6) versus  $E$  using the parameters in Table 2. In decreasing order of thickness:  $\alpha_{\pm} = \alpha, 3\alpha, 9\alpha, 27\alpha$



Though the energy levels have shifted, the wavefunctions remain highly localized. Using the lower value of the energy level in (7.7) in (7.5), we have

$$s_1 = 1.6868 \times 10^{-1}, \quad c_1 = 1.6867 \times 10^{-1}, \quad (7.8)$$

which force  $r$  to be large from (7.4):

$$r_{<} = 6.40 \times 10^4, \quad (7.9a)$$

which corresponds to the wavefunctions being localized in the right dot. Similarly, using the higher energy value yields

$$r_{>} = 3.49 \times 10^{-5}, \quad (7.9b)$$

which corresponds to the wavefunctions being localized in the left dot. In fact, the localization is even more pronounced in the finite-barrier case.

Due to the symmetry in the exterior walls, in the limit  $\delta \rightarrow 0$ ,  $\phi_l = \phi_r$ . Therefore, we write

$$\phi = \phi + \phi_{,1}\delta + o(\delta). \quad (7.10)$$

Hence the asymptotic analysis in Sect. 4 holds once we introduce the phase, so (4.9) becomes

$$\lambda \cot \lambda(\pi + \phi) = -\frac{\alpha_0}{M}. \quad (7.11)$$

Moreover, the leading order of (7.3) may be rewritten as

$$\lambda \cot(\lambda\phi) = \frac{\alpha_0}{M}. \quad (7.12)$$

Continuing to simplify using (7.11), we obtain

$$\lambda\phi = \frac{\pi(1 - \lambda)}{2}. \quad (7.13)$$

Note that substituting (7.13) into (7.11) yields an equation that, along with our definitions of  $\lambda$  and  $\alpha_0$ , can be solved for  $E$ . Equation (7.13) could allow us to rewrite the solution in terms of  $\lambda\phi$  rather than  $\lambda(\pi + \phi)$ . To better illustrate the phase as a displacement from the infinite-barrier case, we leave our results in their current form.

Continuing the analysis as in Sect. 5, we have that (5.3) becomes

$$r_* = \frac{a_1}{\alpha_0 \epsilon_0 \sin \lambda(\pi + \phi)}, \quad (7.14)$$

where the definition of  $a_1$  is more complicated than (5.1b) due to the perturbation of the phase itself:

$$a_1 = \{\alpha_1 - M\lambda[\lambda_{1,1}(\pi + \phi) + \lambda\phi_{1,1}]\} \sin \lambda(\pi + \phi) + \{M\lambda_{1,1} + \alpha_0[\lambda_{1,1}(\pi + \phi) + \lambda\phi_{1,1}]\} \cos \lambda(\pi + \phi). \quad (7.15)$$

Continuing our analysis requires relating  $\phi_{r,1}$  to  $\phi_{l,1}$ . The relevant expression may be obtained by expanding (7.3) for small  $\delta$ :

$$(\lambda_{r,1} - \lambda_{l,1})\phi + \lambda(\phi_{r,1} - \phi_{l,1}) = \frac{c}{2\lambda}, \quad c = \frac{M\alpha_0}{\alpha_0^2 + M^2\lambda^2}, \quad (7.16)$$

where we have used (5.5), (7.11), and (7.12). Substituting (7.16) into (7.15) adapted for  $a_r$ , Eq. (5.7b) still holds, but (5.6b) is replaced by

$$b = \frac{1}{2\epsilon_0} \left[ \frac{M + \alpha_0(\pi + c)}{M\lambda^2} + \frac{M(\pi + c)}{\alpha_0} \right], \quad (7.17)$$

where we have used (7.11). Note that the explicit dependence upon  $\phi$  vanishes, so the solution depends implicitly on  $\phi$  through  $\alpha_0$  and  $\lambda$ . The reduction to the previous case may be seen by noting that the  $c$  term is generated by the finite barrier and will tend to zero as  $\alpha_{\pm} \rightarrow \infty$ . In that limit,  $b$  in (7.17) reduces to  $b$  in (5.6b).

However, in the experimental case under consideration, we may simply substitute the definition of  $c$  in (7.16) into (7.17), yielding

$$b = \frac{1}{2\epsilon_0} \left( \frac{2M + \alpha_0\pi}{M\lambda^2} + \frac{M\pi}{\alpha_0} \right). \quad (7.18)$$

Since the equation for  $r_*$  is the same as in the previous case, there are only two differences between the finite- and infinite-barrier models:

1.  $M$  is replaced with  $2M$  in the first term in the definition of  $b$  in (7.18).
2. The values of  $\alpha_0$  and  $\lambda$  change because of the introduction of  $\phi$  in the analysis.

Substituting the parameters in Table 2 into (7.11), we obtain

$$E = 3.40 \times 10^{-1}, \tag{7.19}$$

which is simply  $E_>$  in (7.7). Hence a negative perturbation on the right well (with  $\delta > 0$ ) returns a lower energy value, validating that our approach continues to return physically reasonable results. We may also calculate  $\epsilon = 2.76 \times 10^{-5}$ , which is small, consistent with our results given previously. Calculating the  $r$  values, we substitute the relevant parameters into (5.7b) to obtain

$$r_< = 4.31 \times 10^4, \quad r_> = 3 \times 10^{-5}. \tag{7.20}$$

These values (especially  $r_<$ ) are not as close to the values in (7.9) as one might expect. The discrepancy can be explained by noting that the approximate theory uses the assumption that  $\epsilon_0 = O(1)$ , while for the actual parameters we have  $\epsilon_0 = 1.97 \times 10^{-4}$ . Note also that as compared with (5.10), there is a higher degree of localization in the finite-barrier case.

### 8 Bounds in finite-height case

We conclude by examining the bound in the finite-barrier case. Substituting (6.1) into (7.18), and rewriting in terms of  $p$ , we obtain

$$\Delta \tilde{H} = 4E_c e^{-2\alpha_0 G} \left( \frac{1 - p^{1/i}}{p^{1/2i}} \right) \left[ \frac{\alpha_0 M \lambda^2}{\pi(\alpha_0^2 + \lambda^2 M^2) + 2\alpha_0 M} \right], \tag{8.1}$$

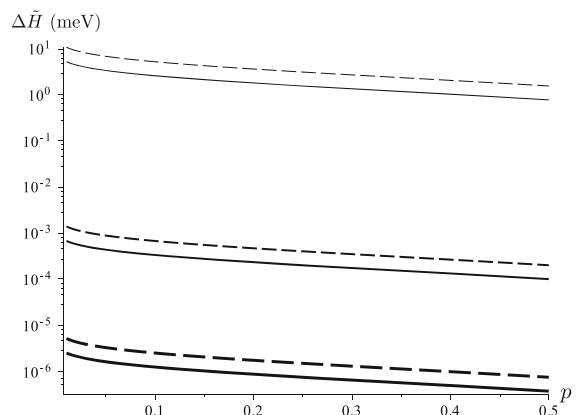
where we have used (6.4b). So the only explicit difference between (8.1) and (6.9) is the extra 2 in the denominator of the bracketed term. Similarly, the definition of  $\Delta \tilde{H}_0$  in (6.6a) changes to

$$\Delta \tilde{H}_0 = 4E_c e^{-2\alpha_0 G} \left[ \frac{1 - p^{2/(n-1)}}{p^{1/(n-1)}} \right] \left[ \frac{\alpha_0 M \lambda^2}{\pi(\alpha_0^2 + \lambda^2 M^2) + 2\alpha_0 M} \right], \tag{8.2}$$

but (6.6b) remains the same.

Figure 17 shows a graph of (8.2) and (6.6b) using (7.19) and various values of  $G$ . [Recall that (7.19) is independent of  $G$  since  $G$  is used only to determine the size of  $\epsilon_0$ .] In this case, the presence of the finite-height barrier means that the bounds are roughly two times smaller. This then forces the tolerance to only approximately 5 to 10 meV with the thinnest barriers available. This more severe bound to enforce delocalization is consistent with the enhanced localization for the finite-barrier case, as discussed in Sect. 7.

**Fig. 17** Plot of (8.2) (dashed lines) and (6.6b) (solid lines) versus  $p$  for  $n = 10$ . In increasing order of thickness: minimum value of  $G$  from Table 2,  $G$  corresponding to a 10-nm barrier, maximum value of  $G$  from Table 2



The finite analog to Fig. 13 is Fig. 18. Again note that the bounds are tighter in the finite-barrier case, as the curves have all shifted to the left. In particular, there is a significant gap between the analytical results and the computational ones for the 10-nm case. Though the analytical results are more conservative, and hence could be relied upon to guarantee delocalization, the discrepancy is still noteworthy.

As in Sect. 6, the analytical results assume that each well is offset by  $\Delta\tilde{H}$ . In contrast, the numerical computations use a normal distribution with standard deviation  $\Delta\tilde{H}$ ; hence many of the offsets would be smaller. Since smaller offsets mean less localization, and the localization itself is highly sensitive to the barrier width  $G$ , it is understandable that the analytical results in the 10-nm case would have a larger discrepancy from the numerical results.

The analogous graph to Fig. 14 is Fig. 19. Note that the curves have again shifted to the left.

## 9 Conclusions and further research

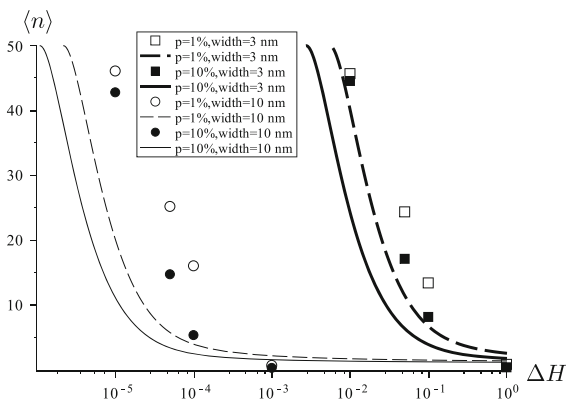
### 9.1 Conclusions

Intermediate-band solar cells have piqued the interest of scientists and engineers because of their potential to enable a larger portion of the solar spectrum to be converted into electricity. QDAs are a promising avenue for fabricating intermediate bands, but the nonuniformity of presently achievable arrays leads to localized wavefunctions that are predicted to enhance thermal relaxation rates and therefore degrade device performance.

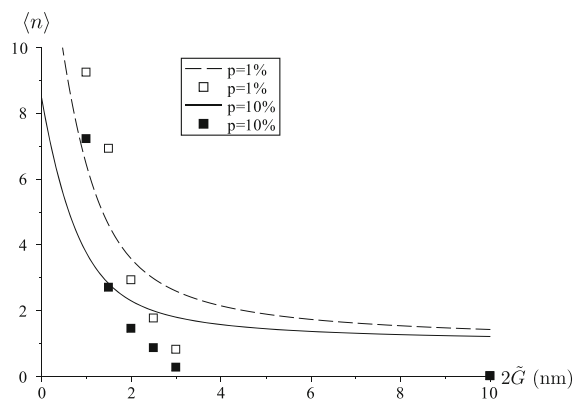
To quantify the level of uniformity needed to produce the desired level of delocalization, we adapted the standard Schrödinger equation to a two-dot QDA system where the effective mass of the electron varies. We solved the resulting equation using perturbation methods, where the zeroth-order state was the perfectly uniform (symmetric) case. The perturbation parameter  $\delta$  characterized the variation in potential from dot to dot due to fabrication issues.

We solved the resulting system in the case of both finite- and infinite-height barriers as  $|z| \rightarrow \infty$ . The two cases are intimately related and in the experimental case are represented by a single change in the expression for  $b$ . The more realistic case of finite-height barriers leads to lower energy levels and higher localization.

By defining the ratio  $r$  between wavefunction amplitudes in each dot, we were able to characterize easily the amount of delocalization. We discovered that in order for there to be a measurable amount of delocalization,  $c_1 = O(\epsilon_1)$  – something not achievable in the experimental system under consideration. Given the result for a two-dot system, we extended our results to  $N$ -dot systems with a simple assumption that agrees reasonably well with more realistic numerical simulations.



**Fig. 18** Lines Plot of (6.7) versus (8.1) for  $N = 50$ . Points Results from computations in [11]. Here the  $x$ -axis is  $\Delta H$  normalized by the value in Table 1



**Fig. 19** Lines Plot of (6.7) versus barrier width for the finite-barrier case. Points Results from computations in [11]

The perturbation approach allowed three separate calculations to characterize delocalization, each of which demonstrated explicit parameter dependence, making the results easier to apply than isolated numerical simulations. First, we constructed a bound in  $\Delta\tilde{H}$  necessary to achieve a certain value of  $r$ , which can be related to  $p$ , the delocalization probability cutoff. As shown in Fig. 17, with the thinnest barriers currently producible, the bound on  $\Delta\tilde{H}$  is approximately an order of magnitude smaller than current technology will allow.

Second, we constructed the average delocalization length of the array  $\langle n \rangle$  as a function of the potential nonuniformity  $\Delta\tilde{H}$ . This calculation provides an estimate for the delocalization length that would occur when  $\Delta\tilde{H}$  is reduced. The analytical estimates are conservative compared to the numerical simulations, as the numerical simulations use a range of values for  $\Delta\tilde{H}$ , while we use a single value. However, the conservative nature of the analytical estimates is useful from an application point of view.

Because the barrier width  $2G$  can be adjusted experimentally (subject to strain considerations), we also showed how  $\langle n \rangle$  varied with  $G$ . Here the analytical calculations were slightly less conservative than the numerical simulations for small  $\langle n \rangle$ , but still more conservative for the experimentally desirable case of large  $\langle n \rangle$ .

The analytical solutions in this work exhibit explicit dependence on the experimental parameters. Hence they can provide guidance to experimentalists about the sensitivity of delocalization to various fabrication factors. In particular, the solution is highly sensitive to  $G$ , which is related to the barrier width. Hence making  $G$  as small as possible (subject to strain constraints) is paramount to increasing delocalization.

## 9.2 Further research

The mathematical model in this paper can be further refined to apply to more realistic systems. Due to alloying between InAs and GaAs, the true shape of the potential well in the QDs is not square. Though the relative effects should be small, such a system can certainly be handled analytically for realistic choices of the well shape.

A key shortcoming of this simplified model is that  $\delta$  (and hence  $p$ ) is assumed the same for each well. This limitation explains much of the discrepancy in our graphs. Using stochastic analysis, one should be able to model a system where  $\delta$  is chosen randomly from a probability distribution, and produce the expected value of  $\langle n \rangle$  for such a system. This could be done by replacing (6.4a) with

$$p = \prod_{j=1}^i r_j^2,$$

where each  $r_j$  is different depending on the value of  $\delta$  for that well. This new form of  $p$  could then be used in the  $\Delta\tilde{H}$  equations [such as (6.9)] to generate more realistic results.

Although we have modeled a specific InAs/GaAs material system, there are a variety of other solar cell device concepts where delocalized wavefunctions in one-, two-, and three-dimensional arrays of QDs would be advantageous for device performance [17]. The approach developed here provides the starting point for mathematical analysis of delocalization in more complex QDAs.

**Acknowledgments** We thank the reviewers for their many insightful comments. The work of Doty and Reid was supported by NSF award DMR-0844747.

## References

1. Lopez N, Reichertz LA, Yu KM, Campman K, Walukiewicz W (2011) Engineering the electronic band structure for multiband solar cells. *Phys Rev Lett* 106:028701
2. Luque A, Marti A (1997) Increasing the efficiency of ideal solar cells by photon induced transitions at intermediate levels. *Phys Rev Lett* 78:5014–5017
3. Luque A, Marti A (2010) The intermediate band solar cell: progress toward the realization of an attractive concept. *Adv Mater* 22:160–174

4. Bailey CG, Forbes DV, Raffaele RP, Hubbard SM (2011) Near 1 v open circuit voltage InAs/GaAs quantum dot solar cells. *Appl Phys Lett* 98:163105
5. Luque A, Linares PG, Antolin E, Canovas E, Farmer CD, Stanley CR, Marti A (2010) Multiple levels in intermediate band solar cells. *Appl Phys Lett* 96:013501
6. Antolin E, Marti A, Olea J, Pastor D, Gonzalez-Diaz G, Martil I, Luque A (2009) Lifetime recovery in ultrahighly titanium-doped silicon for the implementation of an intermediate band material. *Appl Phys Lett* 94:042115
7. Luque A, Marti A, Antolin E, Tablero C (2006) Intermediate bands versus levels in nonradiative recombination. *Physica B* 382:320–327
8. Tomic S (2010) Intermediate-band solar cells: Influence of band formation on dynamical processes in InAs/GaAs quantum dot arrays. *Phys Rev B* 82:195321
9. Popescu V, Bester G, Hanna MC, Norman AG, Zunger A (2008) Theoretical and experimental examination of the intermediate-band concept for strain-balanced (In,Ga)As/Ga(As,P) quantum dot solar cells. *Phys Rev B* 78:205321
10. Lagendijk A, van Tiggelen B, Wiersma DS (2009) Fifty years of Anderson localization. *Phys Today* 62:24–29
11. Reid WM, Driscoll T, Doty MF (2012) Forming delocalized intermediate states with realistic quantum dots. *J Appl Phys* 111:056102
12. Mahan BH, Myers RJ (1987) *University chemistry*. 4. Benjamin Cummings, Menlo Park
13. Levy-Leblond JM (1995) Position-dependent effective mass and Galilean invariance. *Phys Rev A* 52:1845–1849
14. Davies PCW (1984) *Quantum mechanics*. Routledge & Kegan Paul, London
15. Landshoff P, Metherell A (1979) *Simple quantum physics*. Cambridge University Press, New York
16. Misra PK (2010) *Physics of condensed matter*. Academic Press, Boston
17. Hillhouse HW, Beard MC (2009) Solar cells from colloidal nanocrystals: fundamentals, materials, devices, and economics. *Curr Opin Colloid Int Sci* 14:245–259
18. Bailey CG, Hubbard SM, Forbes DV, Aguinaldo R, Cress CD, Polly SJ, Raffaele RP (2009) Effect of barrier thickness on strain balanced InAs/GaAs QD solar cells. In: *Proceedings of the photovoltaic specialists conference (PVSC), 2009 34th IEEE*, pp 000137–000140