Order and disorder in the heteroepitaxy of semiconductor nanostructures

Fulvio Ratto a,*, Federico Rosei b,*

a Istituto di Fisica Applicata, Consiglio Nazionale delle Ricerche, Via Madonna del Piano 10, Sesto Fiorentino, I-50019, Italy
b Institut National de la Recherche Scientifique, Universite du Quebec, 1650 Boul. Lionel Boulet, J3X 1S2, Varennes (QC) Canada

ARTICLE INFO

Keywords:
Semiconductor heteroepitaxy
Epitaxial nanostructures
Self-organization
Substrate patterning
Lithography

ABSTRACT

The heteroepitaxy of semiconductor pairs with a small lattice mismatch is a process of tremendous interest in materials science and technology. The principal mechanism of strain relief in these interfaces is the formation of three dimensional islands either directly on a bare substrate (Volmer–Weber growth mode) or following the formation of an initially flat wetting layer (Stranski–Krastanov growth mode). The elemental and strain inhomogeneities associated with these three dimensional islands may result into a confinement potential for electrons and/or holes, as in a standard quantum well. At variance with a standard quantum well, the confinement in these nanostructures (often referred to as ‘quantum dots’ (QDs)) occurs in all three spatial dimensions and over length scales comparable with the relevant De Broglie wavelength. This strong confinement may give rise to a discrete spectrum of charge carrier energy levels, as in an artificial atom. On the other hand the spectra of these nanostructures may be tuned with their physical and chemical properties, providing an enabling opportunity to design novel optical and electronic components. Epitaxial nanostructures are proposed as the building blocks of a variety of innovative applications, which may represent step-change solutions to many challenges in the fields of photonics and electronics, such as e.g. new possibilities to integrate versatile lasers and transistors in Information and Communication Technologies and to replace MOSFET devices with miniature components capable of sustaining the race to miniaturization of integrated circuits. Examples of possible applications include lasers, optical detectors, white-light sources, single-photon and entangled-pair sources, single electron transistors, quantum cellular automata, quantum bits, etc.

To harness these properties and bring these functionalities to fruition, the ability to manufacture individual QDs may be not enough. Rather, the next critical issue is to gain control over the location of nanostructures with respect to each other and their surroundings, both on a surface and in a three dimensional architecture. The exploration of this issue is essential to engineer nanostructure density, mutual interactions (hybridization of electron energy levels, charge interactions, spin interactions, etc.) and the interface with the external circuitry (electrodes, gates, wires, etc.). Moreover in some applications the principal feature is the layout of a huge number of QDs with respect to each other’s nearest neighbours (as e.g. in a laser), whereas in other applications it is the precise location of a possibly smaller number of QDs within a complex architecture (as e.g. in an SET or QCA platform).

An ample variety of natural (bottom–up, parallel) and artificial (typically integrated top–down and bottom–up, sequential and/or parallel) methods have been reported to yield some extent of control over nanostructure positioning. This review aims at highlighting some of the most relevant concepts developed over recent years. While a significant number of reviews on different aspects of the synthesis and characterization of individual nanostructures are found in the literature, the complexity of the issues mentioned above has never been addressed within a dedicated framework so far.

© 2010 Elsevier B.V. All rights reserved.
1. Introduction

1.1. General framework of semiconductor nanostructures

Ever since the chalcolithic period, the technological and socioeconomic progress of humanity has been accompanied by the invention and exploitation of novel functional materials, including metals, ceramics, and more recently plastics, semiconductors and biomaterials. Over the last two decades the advent of nanotechnology has been a milestone in this race towards innovative materials. A wealth of new concepts have been developed, based on the idea of taking advantage of the size-dependent physical and chemical properties of matter at the nanoscale, which often behaves quite differently with respect to its macroscopic counterpart. In this context, the concept of artificial atoms is one of the most powerful innovations, namely the possibility to engineer elementary building blocks comprising thousands of atoms, yet displaying discrete energy levels.

In 1904 Nagaoka devised the first attempt of a planetary model for natural atoms (the so-called saturnian model), [1] which was later revised by Rutherford and Bohr. According to the now well-established paradigm, the spectral properties and chemical behaviour of ordinary atoms stem from the interaction between light electrons with a negative charge and a massive nucleus with a positive charge, which essentially results into the quantum confinement of the electrons inside the deep potential well associated with the nucleus. This potential well acts as a spherical potential trap for the electrons. In a conceptually similar fashion, comparable quantization features may be achieved by the design of artificial traps, e.g. by the confinement of charge carriers (electrons or holes) into one, two, and even three dimensional ‘boxes’. When the size of these boxes becomes small enough, i.e. at the nanoscale (with lateral dimensions of the order of the charge carriers’ de Broglie wavelength in the medium), the relevant separation between discrete electron levels may become significantly larger than typical thermal energies, which in turn may result into appreciable quantum features. Such three dimensional boxes are usually referred to as Quantum Dots (QDs), and can be considered as artificial atoms. In principle, by attaining a thorough control over the energy barriers, size and shape of the potential wells enclosing individual QDs, the electron levels and orbital symmetries of the corresponding particle system may become broadly accessible [2–4]. Therefore the optoelectronic properties of individual QDs may be tailored to specific needs and applications. The first experimental demonstrations of the possibility to control photophysical and photochemical dynamics of QDs came from the work by Ekimov and Onushchenko in 1981 [5] and Brus in 1983 [6], who investigated size effects in copper chloride and cadmium sulphide colloidal particles respectively (~2.5 and 3.5 nm respectively). Moreover, the perspective to manipulate the relative positions within appropriate systems of QDs may lead to the design of arbitrarily complex and functional architectures, including artificial molecules and even artificial crystals [7,8]. This is expected to lead to new unprecedented opportunities in physics, chemistry, materials science and engineering [9,10].

1.2. Potential applications of quantum confined systems

The potential for applications of artificial atoms, molecules and crystals is enormous. Fields and applications which may become critically affected in the near future include e.g. light emitting diode (LED) and laser technologies, [11,12] single-photon sources, [13] new computation technologies, [14–16] improved catalysis, [17] photovoltaic devices, [18] and environmental and biomedical diagnostics and imaging technologies. [19,20] The latter are only a few notable examples where substantial progress is possible, well beyond possible state-of-the-art solutions.

Suitable nanostructures may emit light as a result of excitation (including e.g. population inversion) and subsequent spontaneous relaxation (and possibly, stimulated relaxation) of their discrete energy levels, i.e. the recombination of exciton pairs where both electrons and holes may be trapped in the same volume. Several reasons motivate the intense research in this field. QD-based LEDs and lasers are expected to exhibit much improved performance with respect to bulk-active emitting materials: a lower threshold power and higher differential gain, a weaker temperature dependence, longer carrier lifetime in the excited states (due to minimal carrier–phonon interactions) [16], as well as a virtually unlimited choice of output frequencies due to size effects. [21–23] Moreover relevant solutions of nanostructures may enable applications which are compatible with the existing semiconductor technology, [9,23] thus leading for example to the integration of micro- and optoelectronic components on the same silicon wafer.
Individual nanostructures may also be used for the realization of single-photon sources and detectors. Such devices are at the basis of many novel applications proposed in the fields of photonics, quantum computation and communication [24,25]. The stimulated light emission from single nanostructures may overcome the limitations of present photon sources and enable the controlled generation of highly correlated single photons [26].

In 1987 Fulton and Dolan demonstrated the first Single Electron Transistor (SET), [27] which relies on Coulomb blockade to control the charge flow from a source electrode towards a drain electrode through two or more tunnelling barriers. These tunnelling barriers enclose one or more islets, through which the charge carriers may or may not hop, depending on the bias applied to a gate electrode. In this context, QDs are proposed as ideal components for the islets and corresponding tunnelling barriers. [28,29] In addition, innovative concepts of calculations and logic operations may be performed by manipulation of the charge occupancy and transfer in suitable architectures of nanostructures. For instance QD-Cellular Automata (QCA) composed of suitable arrays of QDs such as those proposed by Lent may surpass the traditional transistor-based technology in a number of respects, including better performance with further miniaturization (the reverse being the principal drawback of current transistor technologies), parallel computation and adiabatic operation, i.e. minimal power consumption [15,30,31].

While QCA design essentially follows conventional binary logic operations, a stronger enhancement in the computational power of devices based on suitable arrays of QDs may be attained by exploiting the full quantum nature of these systems [14,32]. By encoding the logical values 0 and 1 in genuine quantum states (either charge states, or spin states, etc.) and by relying on their coherent evolution, a new type of information processing may become accessible, exponentially faster than the classical scheme. This may result into an exponential computational speed-up over classical machines and the possibility to treat unsolved classes of problems (possibly including NP-complete problems, such as e.g. the factorization of large numbers, which may be of relevance in a variety of fields like cryptography) [33].

An exhaustive list of all possible applications of QDs could span over any field of technology which may benefit from introducing novel and versatile materials, with properties tuneable down to the nanoscale domain. However, such a list would go beyond the scope of this review. The examples given above were meant to illustrate a very general notion. The fabrication of nanostructures with a quantum confinement profile is a topical issue, which is still not enough to bring to fruition their full potential for actual applications. Another important challenge is to gain control over the location of nanostructures on a substrate of choice, with respect to each other and the surrounding environment. This review begins from this consideration. Addressing this issue is essential to engineer the density of nanostructures, mutual interactions between nanostructures (hybridization of electron energy levels, charge interactions, spin interactions, etc.) and the interface with the external circuitry (electrodes, gates, wires, fiberoptics, etc.). In turn, this notion goes much beyond the realm of QDs and represents a common challenge for a broad variety of nanostructures. The issue to govern the location and functionality of nanostructures with respect to their environment is gaining relevance in fields as diverse as biomedical imaging and therapy, [34–37] catalysis for e.g. gas to liquid conversion, and even solar energy conversion [38–41]. The concept to fabricate hybrid nanoparticles with macromolecular properties is becoming an accessible solution [42–45]. QD-based devices are an instructive example. In some devices the principal concern is the layout of a huge number of QDs with respect to each other’s nearest neighbours (as e.g. in a LED or laser), whereas in other devices it is the precise location of a finite number of QDs within a complex architecture (as e.g. in an SET or QCA platform). We begin by providing a brief description of the kind of nanostructures with potential QD behaviour considered in this review. Next we mention a few factors which may drive a spontaneous correlation in the mutual locations of coexisting nanostructures. In addition these factors may be manipulated by an artificial intrusion to realize functional architectures of nanostructures. A few examples will be given to illustrate the state-of-the-art of this field, along with present limitations and future perspectives.

1.3. Semiconductor epitaxial nanostructures: principal features

There are numerous options to confine charge and spin carriers within suitable potential traps at the nanoscale. Systems under consideration in this review correspond to a specific opportunity which has received considerable attention at least over the last two decades. The basic concept consists in using the band alignment in three dimensional heterojunctions composed of a nanostructure of one semiconductor material epitaxially embedded within the bulk of another semiconductor material. When the lateral size of the nanostructures is small enough, such a band alignment may provide versatile potential barriers for quantum confinement, which stem from elemental and strain inhomogeneities. Fig. 1 displays a representation of possible configurations for band alignment in semiconductor heterojunctions.

When the band gap of the embedded material is smaller than that of the host material, the result may be either a straddling (type I) or a staggered/broken gap (type II/III) band alignment. Under type I conditions, both the conduction and the valence edges of the embedded material lie within the band gap of the host material, so that both conduction electrons and valence holes are trapped within the embedded material (by the quantum boxes defined by the conduction and valence edge profiles respectively, see Fig. 1A). Under type II conditions on the other hand, either the conduction or the valence edges of the embedded material fall outside the band gap of the host material, so that confinement within the embedded material may occur for one kind of charge carrier only (holes and electrons, respectively; Fig. 1B illustrates for instance the case for the valence holes). However, under these conditions confinement of the complementary charge carriers may still occur, e.g. in the host material in the immediate proximity of the embedded material, for instance due to the effect of the local strain field. This configuration is represented in Fig. 1C. In both cases 1A and 1C the system may provide for the desired excited pairs localization, and so act as a quantum confined dot. Examples of the first kind of alignment include several III–V compound heterostructures like InGaNAsAs/InP (with x close to 0.5) [46], or InGaAs/InP/AlAs/InGaAs (x, y close to 0.5) [46]. Important model systems of the second kind are InP/InAlAs (x close to 0.5) [47].

![Fig. 1](image-url) Schematic representation of possible configurations for the band alignment in semiconductor heterojunctions: A: type I alignment; B: type II alignment; C: modified type II alignment. The full and open circles represent conduction electrons and valence holes confined in the nanostructure volume.
and Ge/Si, [48] where valence band holes and conduction band electrons are confined within the islands and close to the islands respectively as mentioned above [49–51]. The position of the discrete energy levels in the nanostructures depends on the size and shape of the intervening potential barriers, i.e. on the elemental composition and strain fields in the heterojunctions. This is an essential issue, which goes well beyond the scope of this review and will therefore not be discussed any further.

2. Self-organization of semiconductor epitaxial nanostructures

2.1. The Stranski–Krastanov and Volmer–Weber growth modes

Among the many different approaches to fabricate functional arrays of nanostructures with potential QD behaviour, we treat specific aspects of a particular bottom–up strategy, based on the heteroepitaxy of suitable semiconductor pairs. This is a parallel process which drives the self-organization of a high surface density of nanostructures over large areas, usually with poor long-range coherence. The self-organization dynamics result from deposition, adsorption and diffusion of atoms of one semiconductor material onto a surface of another semiconductor material with the same symmetry and crystal structure, but different lattice parameter. For instance a model system which is particularly rich in phenomenaology is germanium on silicon (with crystal orientation typically along the high symmetry (0 0 1) or (1 1 1) directions).

Here we summarize notable aspects of the fundamental thermodynamics and instabilities of these surfaces, while significant elements of their kinetics will be treated in Section 4.

The heteroepitaxy of semiconductor pairs with the same symmetry and crystal structure constrains the deposited material to reproduce the lateral geometry and lattice parameter of the substrate material. Therefore due to the lattice mismatch between the constituents, the epitaxial relationship is at the origin of excess elastic strain energy, which is stored at the heterojunction about the constituents, the epitaxial relationship is at the origin of excess elastic strain energy, which is stored at the heterojunction about the interface and the overlayer. This configuration gives rise to surface instabilities, which may trigger drastic morphological transformations. [9,52–55] In particular, the development of a transition from a flat (layer-by-layer) growth mode towards a partial release of the excess elastic strain

Fig. 2. Examples of SK (left) and VW (right) transitions. TEM micrographs of a Ge/Si(0 0 1) dome shaped island from ref [58] (left) [Reprinted figure with permission from A. Rastelli, M. Kummer, H. von Känel, Phys. Rev. Lett. 87 (2001) 256101. Copyright © 2001 by the American Physical Society]; and a AgGaAs(0 0 1) surface from ref [59] (right) [Reprinted figure with permission from E. Placidi, M. Fanfoni, F. Arciprete, F. Patella, N. Motta, A. Balzarotti, Mater. Sci. Eng. B 243 (2000) 69–70. Copyright © 2000 by the Elsevier].

discontinuous) on top of a substrate. The stability of the surface is associated with the interplay of various thermodynamic properties, such as the surface densities of free energy for formation of the film-substrate interface $\gamma_s$, of the film, and of the substrate surfaces, $\gamma_f$ and $\gamma_v$ respectively. According to the Young–Dupré equation the equilibrium condition reads $\gamma_f - \gamma_v - \gamma_s \cos(\theta) = 0$, where $\theta$ is the contact angle of the local film slope with respect to the substrate, as depicted in Fig. 3.

The values of $\gamma_f$ and $\gamma_v$ depend on the crystal orientation of the surface, which is a well-defined parameter. In contrast, $\gamma_s$ depends on the detailed morphology of the film, and in particular on $\theta$ which may be not known a priori. Hence the description provided by the equation above is of immediate fruition only for simple isotropic film materials (i.e. whenever $\gamma_s$ is a unique parameter independent of $\theta$). For instance, to a very good approximation, this condition is fulfilled for liquid metal droplets on ceramic surfaces [63]. However, in the context of semiconductor heteroepitaxy, these requisites are not met, and the Young–Dupré equation is to be restated in a more general form as $\gamma_f - \gamma_v - \gamma_s (\theta) \cos(\theta) = 0$. Hence the overall theoretical description needs to be devised in such a way that $\gamma_s (\theta)$ and the equilibrium morphology are calculated self-consistently. Nonetheless one may still refer to the simple formalism above for an immediate description of the instability behind the formation of three dimensional islands. By straightforward considerations on the Young–Dupré equation, complete wetting may be predicted as long as $\gamma_f < \gamma_v - \gamma_s (0)$; whereas three dimensional islands may be expected to nucleate as soon as $\gamma_f > \gamma_v - \gamma_s (0)$ (and so $\theta$ becomes finite in the equilibrium morphology).

Here we focus in particular on SK processes, which are the most common dynamics observed in systems of technological interest, including e.g. Ge/Si [9,52,54] and III/V heterostructures [64,65]. SK transitions are distinctive of the heteroepitaxy between moderately lattice-mismatched semiconductor pairs. Noticeable examples of

Fig. 3. Ordinary representation of the geometrical and thermodynamic quantities involved in the Young–Dupré equation.
film/substrate couples and relevant misfits \( \Delta \text{misfit} = \frac{|a_i - a_0|}{a_0} \) and \( a_i \), being the film and substrate lattice parameters respectively) are: Ge/Si \( \sim 4\% \); InAs/GaAs \( \sim 7\% \); GaP/GaAs \( \sim 4\% \); GaAs/GaSb \( \sim 7\% \); AlP/InP \( \sim 7\% \); PbSe/PbTe \( \sim 5\% \), etc., and intermediate alloys thereof. In this class of surfaces, the thermodynamic balance switches from wetting \( (\gamma_i^B \leq \gamma_i^H(0)) \) to three dimensional islanding \( (\gamma_i^B > \gamma_i^H(0)) \) conditions, at a specific coverage, which is called the critical WL thickness. Above this thickness the flat morphology is no longer preserved, thus resulting into nucleation and growth of three dimensional islands.

About the roughening transition the outermost surface (i.e., where the nucleation of three dimensional islands actually occurs) consists of a strained homojunction, such as that schematically represented in Fig. 4. Thus the layer-by-layer wetting and three dimensional islanding conditions (for the growth of layer \( i+1 \) on top of layer \( i \)) read \( \gamma_i^H(f(i)) \leq \gamma_i^H(0) - \gamma_i^{H(i+1)(0)} \) and \( \gamma_i^{H(0)} > \gamma_i^{H(i+1)(0)} \) respectively \cite{54,66}. The roughening transition occurs because above a critical threshold the energy required to form the strained interface becomes larger than the difference of the energies required to form consecutive strained surfaces. Therefore the transformation is driven by the accumulation of elastic strain energy in the flat WL. Thereafter the energy required to form facets with high local film slope is more than compensated by the partial relaxation of the elastic strain energy stored in the WL. While the former is essentially independent of surface coverage, the latter grows with overlayer thickness.

### 2.3. Strain accumulation in heteroepitaxial systems

The accumulation of elastic strain energy in the flat WL is associated with the elastic properties of the film (i.e., with Poisson’s number \( \nu \) and shear modulus \( \mu \)); with the lattice mismatch between WL and substrate materials, \( \Delta \text{misfit} \) defined above; and with the thickness of the overlayer, \( t \). Under the simplest circumstances of a flat and homogeneous film, the epitaxial process entails a biaxial stress composition. This generally results in a tetragonal distortion of the crystal unit cell. By definition of the set of coordinates \( x, y \) (in plane), \( z \) (out of plane), the strain components within the coherent film are described by the set of Eqs. (2.1) \cite{54}:

\[
\begin{align*}
\varepsilon_x &= -f \\
\varepsilon_y &= -f \\
\varepsilon_z &= f \cdot \frac{2\nu}{1-\nu}
\end{align*}
\]

(2.1)

The corresponding elastic strain energy per unit surface for a flat heteroepitaxial film of thickness \( t \) reads \cite{67}:

\[
E = 2\mu f^2 \cdot \frac{1+\nu}{1-\nu}
\]

(2.2)

Thus in this approximation the surface density of elastic strain energy increases linearly with the WL thickness, which affects the interface and surface formation energies, and ultimately drives the roughening transition.

### 2.4. Limitations of the Young–Dupré framework

The essential thermodynamic picture given by the Young–Dupré inequalities captures the basics of SK processes. Nevertheless, a few words of caution are warranted. The discussion above is meant to describe the equilibrium configuration of a coherently strained surface, which is thought to exhibit a chemically abrupt interface. As such, it is very incomplete. Under more realistic conditions, three dimensional islands form within a kinetically limited framework, where factors such as the substrate temperature, deposition rate, etc. play a key role and determine the substrate and film interdiffusion profile, the effective lattice mismatch, along with the onset of a number of possible growth instabilities \cite{52,54}. Since most of the system properties are fixed in the deposition process, \cite{54} the principal characteristics of the three dimensional islands are defined under the non-equilibrium framework imposed by various kinetic constraints (including e.g. the adatoms super-saturation and diffusivity through the WL) \cite{68,69}. Besides the essential thermodynamic considerations above, a complete treatment of SK transitions requires additional factors, including at least the possibility of kinetic roughening \cite{52,70,71}, the presence of diffusion barriers throughout the WL, \cite{72,73} a variety of possible crystal defects and the occurrence of preferential nucleation at specific surface sites such as step edges and kinks, \cite{74} where the relaxation of the elastic strain energy may become more efficient. The thermal activation of synergistic dynamics generally accounts for a delicate balance between thermodynamic factors and kinetic limitations, \cite{54,75,77} which is beyond the scope of this review.

SK-grown three dimensional islands of small enough lateral dimensions may be promising candidates for QD fabrication. They may be embedded within the host matrix in a subsequent step by over-deposition of the substrate material, which may result in either of the prototypical heterojunctions described in Section 1. This QD fabrication approach features several advantages. In addition to the aforementioned parallelism of three dimensional islands nucleation, this paradigm involves very low density of crystal defects, which is a fundamental requirement for QD performance \cite{9,78}. Moreover in principle, QDs of very small sizes may be obtained, hopefully resulting in quantum features and devices operational at room temperature.

### 3. Mutual position and size distribution of self-organized islands

In this section we introduce the discussion on the mutual position of arrays of three dimensional islands as well as their size and shape statistics, which are interrelated. Both issues are critical in view of exploiting nanostructures as active materials in novel devices. SK transitions may be used for the gamut of solid state based devices already mentioned in Section 1, and could be easily integrated into conventional semiconductor processing. Ordered growth with precise positioning of nucleation sites in arrays of nanostructures may be in principle unnecessary for applications which rely on the properties of individual nanostructures. Among these we mention light absorbing and emitting devices, such as LEDs, lasers and sensors, whose function rests on the photo-luminescence behaviour of individual QDs, i.e. on the exciton pairs localization in the neighbourhood of the three dimensional islands \cite{11,12,79,80}. However also in this context the ability to influence the relative locations of nanostructures may be still desirable so as
to achieve high surface densities of active elements, and possibly take advantage of molecular features of coupled nanostructures [7,81]. For instance Fig. 5 displays the ground state energy of a pair of coupled GaAs QDs as a function of their absolute size and separation, which was calculated by Delanty et al. [81] by solving the relevant Hamiltonian equation. The QDs interactions result into significant molecular features as far as the QDs separations lie in the order of a few QDs radii, which may be exploited in ordered arrays of three dimensional islands [82]. In turn a random distribution of nearest neighbour distances may translate into a broad distribution of photoluminescence behaviours, which may be detrimental for e.g. LEDs and lasers applications.

On the other hand the ability to control the mutual locations of three dimensional islands becomes a critical issue when device operation and performance intimately depend upon the interactions between coexisting nanostructures and the surrounding environment. For instance, this may be needed in the design of QCA arrays [15,30,31,83] according to the original architecture proposed by Lent et al. in 1993 [84]. Briefly these devices may be implemented to run binary calculations. The logical values 0 and 1 may become encoded in the occupancy pattern of two electrons injected in a QCA cell composed of four QDs placed at the vertices of a square as illustrated in Fig. 6. When allowed to tunnel between adjacent QDs within the QCA cell, the two electrons repel each other and occupy either of the two possible diagonal configurations in the square, which are conventionally labelled as 0 and 1. Similar interactions govern the occupancy pattern of neighbouring cells, which are electrically coupled due to their close proximity. These interactions may be exploited to create all functional logical connections and perform all elementary logical operations. Fig. 6 gives an example of a possible approach to realize a wire and a majority gate with coupled QCA cells. When the height of the tunnelling barriers within the QCA cells is manipulated and clocked by external gate electrodes, this platform may be used to realize computation with the wealth of advantages stated above.

Similar conditions may be met e.g. in the fabrication of an SET. Fig. 7 displays the fundamental elements required to fabricate an SET with a single QD. Electrons can hop from the source electrode to the QD and from there to the drain electrode only when the bias applied to the gate electrode, which is capacitively coupled to the QD, overcomes the Coulomb blockade potential. Here the QCA and SET technologies have been mentioned to represent the complexity of the QD architectures which may be required in feasible applications, with severe requirements on the QD locations with respect to each other and the surrounding environment.

The self-organization of islands through an SK transition typically induces little if any self-ordering, [85,86] unless an artificial intervention is designed. This concept will be expanded in Sections 4 and 5. Here we mention how, depending on the application of interest, self-ordering may or may not be desirable, and generally requires thorough characterization [87]. The relationships between the interactions at play on the surface (both between the substrate and the islands and between coexisting islands) and the geometry and amount of correlation expectable from the self-organization should be analysed in depth, in order to understand their utility and compatibility with the architecture of the final device.

The mutual positions of the nucleation sites also affect the size and thereafter the shape uniformity in ensembles of nanostructures through the kind of dynamics summarized in Section 4. The ability to engineer the size and shape of individual three dimensional islands is reflected in the physical and chemical properties of the nanostructures, and is therefore important for any application. In particular, a broad distribution of sizes and shapes results into a broad distribution of spectral behaviours. [21,23,88] In so far as the functionality of the surface relies on the spectral properties of the islands, a well-defined size and shape distribution is often required. For instance, LEDs and lasers require a highly
uniform separation between spectral lines in a dense array of QDs [81]. A variety of more complex device concepts often insist on the uniformity of the morphological features of individual QDs, [30] whereas specific applications, like e.g. in the field of quantum computation, may involve the controlled growth of clusters and pillars of asymmetric QDs. [89] The general requirement is, in any case, the ability to gain control over the size and shape of the nanostructures.

4. Self-ordering of three dimensional nuclei

4.1. A Brownian picture of nucleation and capture: island positions

In Section 2 we have summarized the basic thermodynamics on the origin of roughening transitions in VW and SK heteroepitaxial processes, which provides a zero-order description for the formation of three dimensional islands. Here we treat the kinetics of the nucleation and growth processes, which may induce correlation in the mutual positions of coexisting nanostructures. The analytical formulation and computational framework developed to understand these dynamics is described in excellent reviews, such as the extensive work by Fanfoni and Tomellini (see e.g. ref [90]). Here we propose a qualitative picture of the sequence of events which occur during deposition near the roughening transition. The formation of three dimensional islands occurs via the interplay of competitive nucleation and captures dynamics, which are mediated through the diffusion of a non homogeneous density of adatoms from the deposition source [91,92].

When the thickness of the WL exceeds a critical value (which is null in the case of a VW process and finite, typically a few monatomic layers, in that of a SK process), the layer-by-layer growth becomes thermodynamically unfavourable, due to the excess elastic strain energy stored in the flat overlayer. Both the step-flow dynamics and the formation of two dimensional islands are hindered, and so the deposition flux results into an increasing density of diffusing adatoms which hop across the surface with high mobility. At some point these adatoms begin to condensate into three dimensional clusters, which are fit to partially release the epitaxial constraint by a local deformation of lattice parameters, thus lowering the elastic strain energy. Nonetheless, the transition from two to three dimensional growth requires a transient super-saturation of diffusing adatoms on the surface. Indeed the evolution of three dimensional islands originates from the growth of primitive stable nuclei, whose formation is mediated by collision events among a number of diffusing adatoms higher than a critical threshold i [93]. The critical size i is an attribute specific of every single epitaxial process and material pair. For instance typical values for i were estimated as 4 and 9 for Ge/Si(0 0 1) [94] and Ge/Si(1 1 1) [95] respectively.

The probability of a successful collision event increases with the local density of diffusing adatoms, and becomes significant only wherever certain super-saturation conditions are met. In the simplest picture where diffusion is governed by Brownian motion, the local probability to form a stable nucleus becomes proportional to \( \rho(\bar{r})^{i-1} \), \( \rho(\bar{r}) \) being the density of diffusing adatoms at position \( \bar{r} \). Prior to the onset of nucleation, the adatoms density is essentially uniform throughout the surface. However, as soon as a stable nucleus appears the adatoms density becomes modified in its surroundings. Upon formation, the stable nuclei begin to expand at the expense of the density of diffusing adatoms in their proximities through a capture process [96,97]. Thus mobile adatoms may reach the stable nuclei (e.g. simply via Brownian motion) and attach onto their enlarging edges. When the detachment probability is negligible, the stable nuclei behave as perfect sinks for mobile adatoms. This in turn affects the spatial uniformity of the adatoms density itself. In the stationary regime and under conditions of cylindrical symmetry, Fick’s second law of diffusion around a stable nucleus reads as follows:

\[
\frac{d^2}{dr^2} \rho(r) + \frac{\partial}{r} \left( \frac{\partial \rho(r)}{r} \right) = \frac{F}{D} \tag{4.1}
\]

Where \( F \) is the deposition flux and \( D \) the surface adatoms diffusion coefficient. If we assume the surface density of mobile adatoms to be null at the edges \( r = \ell \) of the stable nucleus and to take the unperturbed value \( \rho_\infty \), beyond a certain distance \( r > \ell, L \) being the so-called adatoms diffusion length, then the solution of Eq. (4.1) for \( \ell < r < L \) is:

\[
\rho(r) = \rho_\infty \left( 1 - \ln \left( \frac{L}{r} \right) \right) - \frac{F L^2}{4 D} \left( (r/L)^2 - 1 - (\ell/L)^2 - 1 \right) \ln \left( \frac{r/L}{\ell/L} \right) \tag{4.2}
\]

The solution given in Eq. (4.2) is represented in Fig. 8 as a reference. The super-saturation conditions locally vanish around the stable nucleus [98]. Owing to the nucleation and capture dynamics, the creation of a stable nucleus and its immediate enlargement induce a depletion in the density of diffusing adatoms (solid line in Fig. 8), and consequently in the probability of nucleation of another three dimensional island (broken line in Fig. 8). Therefore the stable nuclei are expected to become surrounded by a so-called island denuded zone, whose lateral size correlates with the adatoms diffusion length and deposition flux, i.e. the principal kinetic components in the surface heteroepitaxy. Under typical growth temperatures, characteristic values for the adatoms diffusion length are of the order of several tens to several hundreds of nm with activation energies of the order of one eV (see e.g. ref [54] and references therein). The spatial arrangement of the three dimensional islands departs from a Poissonian distribution of points [90,99].

Several authors have reported statistical distributions for parameters indicative of the local particles arrangement inconsistent with a Poissonian distribution of points, including the nearest neighbour distances, [81,100–103] the pair radial distribution functions (density of particles vs. separation from a reference particle), [104,105] the pair angular orientation functions (density of particles vs. orientation around a reference particle), [86,104] and the cell areas in the particles’ Voronoi tessellation (area of the locus closer to a reference particle centre than to any other), [94,103,106] and attributed these variations to the Brownian dynamics described above. Another noteworthy feature reported in the scientific literature is a tendency towards local crystalline order, along either the substrate pattern symmetry, [102] or the isotropic hexagonal symmetry [101,103]. Relevant examples include Ge/GaAs(0 1 1), [100] InAs/GaAs(0 0 1), [101,106] and Ge/Si both with (0 0 1) [94,104] and (1 1 1) [103] substrate orientations.

For instance successful approaches to represent the distributions for the nearest neighbour distances \( d \) and the Voronoi cell areas \( S \) are by use of an empirical Weibull function \( f_{\text{W}}(d) = \alpha A d^\alpha e^{-\alpha d^\beta} \), and corresponding nucleation probability \( P \) in the case of a critical nucleus size \( i = 5 \) (broken line).

![Fig. 8.](image)
with parameter $\alpha$ and $A(\alpha) = \Gamma(1 + \alpha^{-1})$, and a semi-empirical gamma function $f_{VCA}(S) = \beta B(\beta S) e^{-\beta S}$, with parameter $\beta$ and $B(\beta) = \Gamma(\beta)^{-1}$ respectively [94,99,106,107]. In a Poissonian distribution of dots, the parameters $\alpha$ and $\beta$ are found to approach 2.0 and 3.5 respectively, [99] while the values extracted from experimental distributions of three dimensional islands may prove significantly larger. Fig. 9 demonstrates how the experimental distributions may lie between the limit configurations of a Poissonian distribution of dots and an array of dots with crystalline perfection.

Some authors proposed a model of these results based on a rigid exclusion zone, [90,92,102] with size related to the adatoms diffusion length. In this approach each three dimensional nucleus is thought to lie at the centre of a rigid disc. Nucleation is assumed to be inhibited and unperturbed within and outside this disc respectively. The radius of this disc may be equated to the adatoms diffusion length. In contrast, other authors introduced the concept of effective repulsions between three dimensional nuclei, which may induce partial crystalline order (local effects with poor long-range correlation, and so the surface may bear resemblance to an amorphous solid) [99,100,103]. For instance, Ratto et al. developed an approach to quantify the efficiency of this effective repulsion in the definition of local order [99,103]. Local features of the statistical distributions of Ge/Si(1 1 1) three dimensional islands were compared with those generated by the distortion of a hexagonal lattice (Monte Carlo simulations), which was taken as the limit order configuration. The distortion was operated by allowing the various lattice sites to move a distance apart about the respective crystalline lattice sites. The distance distribution was modelled as an isotropic Gaussian function with identical width $s_{\text{diff}}$ for all lattice sites, which was expressed in units of the nearest neighbour distances. With these assumptions the width $s_{\text{diff}}$ is a unique parameter which describes the lattice distortion, and whose adaptation may be included in a recursive routine to fit the

![Fig. 9](https://example.com/fig9.png)  
**Fig. 9.** Left panel: low energy electron micrograph of a typical Ge/Si(1 1 1) surface with relevant Voronoi tessellation. Right panel: relevant experimental distributions for the nearest neighbour distances and Voronoi cell areas, compared with those expected from the limit configurations of a Poissonian distribution of dots and an array of dots with crystalline perfection (from ref [99]) (Reprinted figure with permission from F. Ratto, T.W. Johnston, S. Heun, F. Rosei, Surf. Sci. 602 (2008) 249. Copyright © 2008 by the Elsevier).

![Fig. 10](https://example.com/fig10.png)  
**Fig. 10.** Experimental distributions for the nearest neighbour distances (left) and the Voronoi cell areas (right) of Ge/Si(1 1 1) islands grown at different temperatures. The experimental results were fitted with the model described in the text (from ref [99]) (Reprinted figure with permission from F. Ratto, T.W. Johnston, S. Heun, F. Rosei, Surf. Sci. 602 (2008) 249. Copyright © 2008 by the Elsevier).
experimental data (the higher $\sigma$ the less efficient the effective repulsion). The result of this operation is displayed in Fig. 10, in comparison with the data acquired at different substrate temperatures. The sole parameter $\sigma$ proved sufficient for a satisfactory description of the experimental data, much better than the zero-order Poissonian model (random limit in Fig. 10).

As shown in Fig. 11, the best estimates of the width $\sigma$ were found in the range 0.3–0.5, \[99\] practically independent of the Si(1 1 1) substrate temperature in the range 400–600 °C, in spite of an Arrhenius decrease of the nuclei’s number density by a few orders of magnitude. These values are inconsistent with a Poissonian distribution of points (represented by $\sigma \rightarrow \infty$, or at least $\sigma \gg 1$), although the distortion is significant enough to limit the effects of local order to the very first neighbours. The overall configuration was mainly attributed to the diffusion and capture dynamics presented above.

4.2. A Brownian picture of nucleation and capture: island dimensions

Brownian dynamics naturally imply a correlation between the mutual positions and dimensions of the three dimensional islands. The processes which will be mentioned below reflect the intuitive concept that the higher the number density of local dots the smaller their dimensions for a given amount of material. In the framework of genuine Brownian dynamics, the relative growth of three dimensional islands may be described by geometric models \[108,109\]. A notable example is the model devised by Mulheran and Blackman \[96,110,111\]. Here the key notion is the capture zone of an island, \[90,112,113\] i.e. the equivalent mobile adatoms capture area of one particular island in the competition among coexisting islands to collect the incoming adatoms during deposition. The model essentially predicts the instantaneous growth rate of coexisting islands to be proportional to their instantaneous capture zones.

The exact definition of capture zone is a non-trivial issue. A useful approximation may be the Voronoi cell in the standard Voronoi tessellation of the islands ensemble \[103,106\]. The Voronoi cell of one reference island represents the fraction of the surface closer to that island than to any other, and is constructed by drawing the axes of the segments joining that island’s center of mass to its neighbours’ centres of mass. One example of a standard Voronoi tessellation is given in Fig. 12a. The attractiveness of this approximation is that, under the hypothesis of simultaneous nucleation, the standard Voronoi tessellation may be regarded as a static construction, so that the island volumes may be projected to be proportional to their Voronoi cell areas at any deposition time \[94,103,106\]. However when the island size becomes comparable with the island–island separation, this approximation no longer holds. Under these conditions the detailed shape of the islands perimeters needs to be accounted for. A better approximation may be achieved by introducing the so-called edge–edge tessellation, \[114\] which is drawn by taking the axes of the segments joining one island edges to its neighbours’ edges (see Fig. 12b).

However the best approximation as devised by Bartelt and Evans requires better comprehension of the adatoms’ diffusion dynamics \[114\]. The capture zone of one island is defined as the locus where the field lines of the negative gradient of the mobile adatoms density point towards that island. Therefore the capture zones perimeters correspond to relative maxima of the diffusing adatoms density, and are characterized by the general condition $\nabla \rho(\vec{r}) = 0$. This approximation is represented in Fig. 12c.

Recent efforts comprise additional detail, \[115,116\] including e.g. realistic boundary conditions for the attachment and detachment probabilities at the individual islands’ edges.

The capture zone model was tested in several semiconductor heteroepitaxial systems, including e.g. Ge/Si(0 0 1) (see Fig. 13), \[94\] Ge/Si(1 1 1) \[103\] and InAs/GaAs(0 0 1) \[106\] (see Fig. 14) with alternate results.

The hypothesis to describe the islands growth dynamics through their genuine geometrical environment may be oversimplified, since other thermodynamic factors such as the islands stabilities and surface potentials may play a critical role \[9,52,115,116,118\]. For instance Miyamato et al. investigated the plausibility of the capture zone model in Ge/Si(0 0 1) heteroepitaxy at different coverages (Fig. 13) \[94\]. An important factor at the bottom of the disagreement with the capture zone model is the erosion of the substrate and WL, which contributes significantly to the islands volumes and chemical composition \[54\]. Since a discussion of the islands sizes and shapes lies outside the scope of this review, we do not pursue these concepts any further. However the considerations presented in this paragraph were introduced to give an idea of the depth of interactions among different critical issues in the heteroepitaxy of semiconductor pairs.

---

Fig. 11. Reciprocal of the width $\sigma$ estimated from the analysis of the data in Fig. 10.

Fig. 12. Different approximations for the capture zones of an ensemble of three dimensional islands: the Voronoi tessellation (a); the edge–edge tessellation (b); and the solution proposed by Bartelt and Evans (c) after ref \[114\], according to the variation proposed in ref \[117\] (Reprinted figure with permission from M.C. Bartelt, C.R. Stoldt, C.J. Jenks, P.A. Thiel, J.W. Evans, Phys. Rev. B 59 (1999) 3125. Copyright © 1999 by the American Physical Society).
such as the three dimensional islands mutual positions and dimensions.

4.3. Additional interactions dictating correlation among nuclei

Aside from the Brownian dynamics discussed above, self-ordering of three dimensional islands may result from competing thermodynamic factors, \cite{87,119,120,121} i.e. interactions driven by potential energy gradients, which may involve coexisting particles and the substrate. Any inhomogeneity in the elastic strain energy density, defect distribution and adatoms diffusion energy barriers through the substrate and the WL may modulate the nucleation probability of the three dimensional islands.

### 4.3.1. Island–substrate interactions

#### 4.3.1.1. The Asaro–Tiller–Grinfeld instability

The substrate uniformity is not an absolute concept, not even under ideal conditions. In particular, in the case of an SK transition, the strain energy density in the WL may become inhomogeneous, and possibly modulated in a periodic super-structure. Such a possibility is witnessed by surface instabilities leading to the appearance of wave morphologies, \cite{52,122,123} such as e.g. originally predicted by Asaro and Tiller in 1972, \cite{124} and independently by Grinfeld in 1986, \cite{125} i.e. the so-called Asaro–Tiller–Grinfeld (ATG) roughening \cite{126,130}. Very briefly, in the case e.g. of a plane film under biaxial compression, such as a germanium WL on a silicon substrate, the elastic strain energy stored in the epitaxial overlayer may be partially released by the appearance of ripples. However this process induces an increase in the surface density of free energy. The balance between these components translates into a condition for the minimum possible wavelength (critical wavelength) of the morphological perturbation \cite{53}:

\[
\lambda_c = \frac{2\pi \mu y'(0)}{(1 - v)\sigma^2}.
\]  

(4.3)

In Eq. (4.3), we have used the same notation as before for the shear modulus \(\mu\), Poisson’s number \(v\), and surface density of free energy \(y'(0)\), while \(\sigma\) is the misfit stress. In the linear elastic regime \(\sigma = Ef\) is the product of Young's modulus and misfit parameter (see Eq. (2.1)). Any perturbation with wavelength longer than the critical wavelength may propagate through the surface with no energy barriers except for mass transport across the overlayer. Therefore the surface is unstable against a continuous evolution of ripple amplitudes. In practice, ripple formation is due to gradients in the surface chemical potential associated with strain relaxation at the ripple tops and strain accumulation at the ripple valleys. These gradients drive mass migration from the valleys toward the tops, thus enhancing the roughening in the WL.

With reasonable values for \(E \sim 10^{11} \text{ Pa}\), \(y'(0) \sim 1 \text{ m}^{-2}\), \(\sigma \sim 10^5 \text{ Pa}\) (\(\gamma \sim 1\%\)), the critical wavelength becomes a few 100 nm, which is compatible with representative distances found between three dimensional islands as observed in experimental micrographs of SK transitions under typical growth conditions. However the ATG instability is not enough to explain the formation and evolution of three dimensional islands in SK transitions, which is mediated by the formation of critical nuclei as mentioned above \cite{126}. In turn, it may still induce a strain modulation in the WL \cite{53,131}. A regular modulation of the elastic strain energy density in the WL may define a correspondingly regular network of preferential nucleation sites for the three dimensional islands at the SK transition \cite{119,131,133}. When referring to the epitaxy of Ge on Si for instance, the three dimensional islands are expected to exhibit a preference to nucleate at sites of the WL where the strain is relatively more tensile. Within coherently strain modulated domains, the array of preferential nucleation sites may become coherent as well. Thus if the strain modulation propagates with long-range order, the arrangement of three dimensional nuclei may exhibit long-range order as well. In particular, regardless of the amplitude and effectiveness of the strain modulation, the isotropic symmetry of the three dimensional islands ensemble may be broken. Obviously, since different possible instabilities...
with typical wavelengths longer than the critical wavelength may emerge, propagate and compete, the coherence of the strain modulation may be lost even over the short-range and give poor correlation.

4.3.1.2. Step edges and kinks. Surface instabilities such as the ATG roughening may come into play even under ideal substrate and overlayer conditions. Another important component which may modulate the nucleation sites arrangement are surface defects, which populate the surface with an equilibrium density at any temperature.

Most point and line like defects may act as preferential nucleation sites, as will be shown in Section 5. The case of step edges and kinks is particularly ubiquitous. Step edges and kinks may act as preferential nucleation sites and induce partial alignment of the three dimensional islands, [74,87,131,134,135] due to the interplay of a variety of factors. For instance because of a lower crystal coordination, these defects are more reactive and preferentially capture the available adatoms [10]. Moreover step edges and kinks may allow for a better relaxation of the elastic strain energy [136]. Finally these defects may also modulate the diffusive dynamics of the mobile adatoms, [100] which may result into a local overload of deposited material, thereby increasing the nucleation probability.

4.3.2. Island–island interactions

Besides substrate–island interactions, other possible factors for self-ordering may be traced in island–island interactions different from those described at the beginning of this section, which may be either mediated through the substrate or WL, or possibly originate from weak forces between coexisting islands. We focus briefly onto possible island–island interactions governed by island induced strain fields, which may propagate e.g. in the WL.

For instance Raiteri et al. simulated the strain fields below Ge/Si(0 0 1) islands (with both a pyramid and a dome shape), based on an atomistic approach [54,137,138]. As for the elastic properties, this system may be regarded as a convenient model to understand general trends observed in semiconductor heteroepitaxy, the absolute value of the lattice mismatch playing the principal role. Fig. 15 displays selected results of the calculation by Raiteri et al. The lateral strain configuration below different types of islands was shown to feature a slight tensile component immediately beneath the island base, and a prominent compressive component within a corral beneath the island perimeter. The shape of this corral follows the island perimeter and extends over a few nm from the island edges. Such a strain pattern maintains an intuitive interpretation, since the lattice parameter of Ge is larger than that of Si, and the relative lattice variation relate to the ideal parameter of relaxed bulk Si.

Based on finite element calculations, Meixner et al. achieved further detail on the elastic strain energy distribution in the WL in this system [139]. Along the elastically soft (0 0 1) crystal directions, the compressive strain component decays monotonically. In contrast, along the elastically hard (0 1 1) crystal directions, there exists a local maximum of the compressive strain component located away from the island edges, despite a rapid initial decay. Consideration of the relative lattice mismatch suggests that preferential nucleation of three dimensional islands will occur under higher tensile strain, which allows for a better accommodation of Ge-rich structures [54]. Thus the compressive component around any island may provide for repulsive interactions against neighbouring islands, modify the adatoms diffusion and contribute to the suppression of nucleation at least within a distance of a few nm of the island edges [140–142]. Moreover the directional feature of the elastic strain energy distribution may modify the chemical potential landscape, thereby inducing preferential nucleation along the elastically soft (0 0 1) crystal directions of any island. This process may cause alignment of the nuclei along specific crystal directions, especially in the high number density regime [102,139].

The potential of island–island strain mediated interactions to induce self-ordering on unperturbed substrates is overall quite poor. However the effectiveness of the elastic strain energy accumulation to dictate the mutual positions in ensembles of SK islands becomes amplified in a slightly more complex context than those considered so far. When stacking multiple layers of Ge/Si islands by recursively depositing Ge above the roughening transition and then Si-capping layers of suitable thickness, islands progressively become aligned in a vertical and lateral lattice [120,143–145]. Fig. 16 (ref [120]) demonstrates the vertical and lateral alignment of nanostructures achieved after 10 layers of 7 ML Ge and 60 nm Si deposited at 750 ° C.

This phenomenon is common to the epitaxy of many semiconductor pairs, and is often attributed to the interplay of vertical and lateral strain components inside the overlayer [49,146,147]. For instance Schmidt et al. performed a systematic study on self-ordering in multiple stacks of CdSe/ZnS/Se 1 –x dots [147]. The extent of vertical and lateral organization was found to improve with the number and deteriorate with the thickness (within a range of a few nm) of the stacking overlayers. An increase in the substrate temperature promotes self-ordering, which is compatible with a thermally activated, strain driven dynamics.
Another similar example is observed during the annealing and the first stages of Si-capping of Ge/Si islands at high temperatures [86,148]. Fig. 17 shows the effect of a thin Si layer on the Ge/Si first stages of Si-capping of Ge/Si islands at high temperatures. Another similar example is observed during the annealing and the first stages of Si-capping of Ge/Si islands at high temperatures [86,148]. Fig. 17 shows the effect of a thin Si layer on the Ge/Si first stages of Si-capping of Ge/Si islands at high temperatures. Another similar example is observed during the annealing and the first stages of Si-capping of Ge/Si islands at high temperatures [86,148]. Fig. 17 shows the effect of a thin Si layer on the Ge/Si first stages of Si-capping of Ge/Si islands at high temperatures.

Under typical conditions, all the various factors mentioned above may simultaneously contribute to the definition of preferential nucleation sites. Understanding whether the tendency towards self-ordering is dominated by the Brownian dynamics behind nucleation and capture or by other thermodynamic components is rather complex. Some authors noted that the temperature dependence of the self-ordering patterns might contain precious information [99,103,147]. In this context, Ratto et al. proposed a qualitative criterion, which was applied to the growth of Ge/Si(1 1 1) and may be adapted for a variety of nanostructures [103]. The principal role of the substrate temperature in Brownian dynamics is to determine the adatoms diffusion length, which affects the number density of the three-dimensional nuclei and yet exerts poor influence on the fraction of self-ordering (self-similarity of ensembles of dots grown under different substrate temperatures). In contrast a variety of factors including the average island–island separation, the thermal activation of e.g. diffusive dynamics and entropy may determine a complex profile of the effectiveness of the other thermodynamic factors with the substrate temperature (non self-similarity). These thermodynamic processes are further complicated by simultaneous thermally activated mechanisms, such as the atomic intermixing of the substrate and deposited materials, [54] and the nucleation of misfit dislocations, [149,150] which alter the effective lattice mismatch, strain composition and chemical potential landscape of the islands, WL and substrate [54,128,151]. Therefore whether the diffusive dynamics or the other components dominate self-ordering may be reflected on whether or not surfaces prepared under different kinetic conditions display self-similarity. Based on these notions, the principal role of partial self-ordering of dots in the epitaxy of Ge/Si(1 1 1) was assigned to Brownian dynamics [103]. However the balance and hierarchy may differ from system to system, and a general approach to handle this critical issue is still missing.

5. From bottom–up back to top–down: integrated top–down approaches

In most cases of practical interest, self-ordering of semiconductor QDs is inadequate to realize applications based on the interactions between the dots and their surrounding environment. In this section, we aim to illustrate how the concepts discussed above are essential to devise possible ways forward.

Over the last decades progress in the semiconductor industry has been accompanied by the refinement of concepts to modify silicon substrates by a top–down intrusion, essentially based on conventional lithographic techniques. The roadmap to the shrinkage of the lateral and vertical dimensions of the various features imprinted in the silicon wafers is implied in the notorious Moore's law, [152] which projects an exponential increase in transistor density per microchip over time. The preservation of this roadmap is progressively leading to a point where the use of conventional concepts of integrated circuitry runs into intrinsic physical limitations [9,87,153]. Indeed this issue is among the main motivations why there exists deep interest in the alternative paradigms described in this review, based on the use of bottom–up processes. For instance the self-organization of Ge/Si nanostructures of suitable dimensions may enable the design of new generation transistors and quantum bits. The possibility to control the size and shape of these nanostructures with the substrate orientation, kinetic conditions, and coverage is a significant advantage in view of future applications [9,52,151,154–156]. However the possibility to achieve useful architectures of nanostructures within a simple bottom–up framework is controversial. Compared to the number of empirical recipes in the scientific literature to modulate the correlation and uniformity of the islands (occasionally by chemical modification of the substrate) [157,158], the pursuit of fundamental insight into underlying dynamics is a more recent challenge.

In Section 4 we have introduced several reasons why a pure bottom–up process may result into partial correlation among coexisting islands. Factors which may induce self-ordering span from island–island interactions, e.g. attributable either to the Brownian dynamics behind nucleation and capture or to elastic repulsions, to island–substrate interactions, e.g. due to strain modulations in the buffer layer and WL [121,132,133]. However the efficacy of these interactions is a critical issue. According to the Monte Carlo simulations by Larsson et al., [132,133] the role of diffusive interactions and amplitude of strain modulations are not sufficient to realize useful patterns under realistic conditions. Nonetheless, hopeful results have been achieved in more complex contexts, and particularly by the propagation of strain through vertical superstructures (e.g. by burial of the islands with substrate material, and especially by cyclic overgrowth of layers of islands and substrate materials) [49,86,120,146,147,159]. Elastic interactions between coexisting islands are believed to trigger the vertical alignment and lateral order of the nucleation sites, which also...
translates into a better size uniformity [115,160]. However despite this significant progress, the versatility in the design of nanostructures architectures is intrinsically limited by the nature of the interactions at play.

On the whole, the bottom–up fabrication of three dimensional islands with functional properties holds the promise to become a viable solution for a variety of technological issues in the near future. However, the notion to achieve arbitrary patterns of QDs through a simple bottom–up process is not realistic. The kind of interactions behind the self-organization of ensembles of QDs seem inadequate for numerous applications, ranging e.g. from SET approaches take advantage of the kind of fundamental interactions at the origin of the self-ordering dynamics described above (diffusive interactions, island–substrate interactions, island–defects interactions, island–island interactions, etc.). Here we will sketch a few noteworthy concepts through significant examples of experimental findings.

5.1. Substrate patterning

5.1.1. Modulation of the step density

A widely used approach to manipulate the nucleation of three dimensional islands is the design of suitable architectures of steps, kinks, troughs, mesas, etc. on the substrate [9,87]. When the step geometries and heights become significantly larger than one monoatomic layer (e.g. on appearance of sawtooth like profiles), the notion of faceting may be more appropriate. In any case, the concepts developed in Section 4 still apply. The external intervention modifies the adatoms diffusivity, strain composition and chemical potential landscape of the surface, dictating preferential nucleation sites for the reasons introduced above. However, a proper design of the exogenous intervention may be significantly more effective than the weak interactions observed in self-organization dynamics.

5.1.1.1. The step bunching and faceting instability. A model approach which sits at the border between self-organization and top-down fabrication takes advantage of an instability which is typical of vicinal semiconductor surfaces, including in particular Si(0 0 1) and Si(1 1 1). In these substrates the equilibrium step distribution is an interesting issue [52,134]. Due to the asymmetry in the kinetic (e.g. adatoms diffusion barriers) [54,161] and thermodynamic (e.g. elastic strain energy distribution) [162] conditions at step edges, a uniform distribution of step edges becomes unstable against fluctuations [131,163,164]. In particular a surface under strain may tend to undergo step bunching, i.e. the progressive accumulation of step edges into multilayer bunches. This instability does not exhibit a typical size, [52,162] i.e. the ideal limit configuration is one giant step bunch between two semi-infinite flat terraces. However the realization of this process is impractically slow. One way to obtain step bunching is by deposition of a buffer overlayer [131,165]. Under a homoepitaxial deposition flux the step-flow kinetics and thermodynamic factors work synergistically towards the development of the instability. In these circumstances theoretical arguments predict the typical size of the step bunches to be finite and depend on kinetic parameters such as the deposition flux (shorter bunches at higher fluxes) and adatom diffusivity (higher bunches at longer diffusivities) [52]. Moreover since on average high bunches tend to emit steps and short bunches tend to capture steps, all bunches tend to converge to the same typical size. Another way to stimulate step bunching is by applying a direct flow of electrical current at temperatures above the semiconductor sublimation point (e.g. ~1220 °C for Si) [85,134,166]. The tendency towards step bunching may be controlled with the direction of the electrical current. The effect is maximized when the electrical current flows along the step-down direction and minimized along the step-up direction, as dictated by the wafer miscut. Unfortunately the realization of a regular pattern of step bunches is hindered by different kinetic limitations, which narrow the versatility of this approach to specific experimental conditions. For an excellent review on the
step bunching instability as well as other types of faceting instabilities we refer to Teichert [52].

The pattern induced in these substrates significantly modifies the SK dynamics in the heteroepitaxial growth of three dimensional islands [85,131,165,167]. For instance Motta et al. investigated the nucleation of Ge/Si(1 1 1) on surfaces prepared with different extents of step bunching [85,166]. Early nucleation sites tend to emerge and line up along the step edges. Additional deposition results into multiple rows of three dimensional islands within the terraces, parallel to the step edges. This self-ordering is ascribed to the combination of kinetic and thermodynamic factors, including Ehrlich–Schwoebel barriers, diffusive nucleation, chemical potential modulations and elastic island–island repulsions. Fig. 18 compares the patterns achieved by deposition of Ge onto different vicinal substrates where the step bunching instability was triggered by depositing a Si buffer [131]. While intriguing and instructive, spontaneous step bunching may not yield the versatility needed in the design of complex architectures of nanostructures.

5.1.1.2. Optical lithography and conventional etching. Similar effects are observed when the substrate is patterned by conventional and unconventional lithographic methods, i.e. when steps, kinks, troughs, mesas (i.e. extended plateaus), etc. are defined by artificial fabrication.

In this context the use of conventional optical lithography is attractive due to the extensive know how developed over recent decades. Several methods have been explored, based on the combination of various irradiation sources, sacrificial layers and etching agents [87,168,169]. Both the diffusive dynamics and the thermodynamic landscape of these substrates may be affected by the external deformation, which in turn modifies the nucleation of heteroepitaxial islands. Depending on the ratios of the dimensions of the artificial pattern, the strain configuration, and the kinetic conditions realized during growth, three dimensional islands preferentially nucleate either on top of the mesas, or along the edges, or within the troughs (see Fig. 19 [169] where all islands were realized at the same substrate temperature of 600 °C and the principal effect is due to the relative substrate strain). Partial alignment parallel to the edges is a common outcome, which closely resembles the configuration found in self-organized step bunches. However the artificial pattern allows for an adaptable allocation of these steps [170].

Another opportunity related to the use of conventional optical lithography is the engineering of facets with different crystal orientation [171,172]. For instance, Ferg et al. investigated the nucleation of Ge-rich dots on a (0 0 1) oriented Si substrate patterned with troughs and mesas which were bound with (1 1 1) facets of different length [173]. Some detail of this work is instructive to exemplify the degree of development and complexity of these conventional methods. The Si(0 0 1) surface was first oxidized and covered with a photoresist. Subsequently the photoresist was exposed and developed to create different rectangular windows via conventional lithography. Silicon oxide was selectively removed below these windows by a buffered oxide etch solution. Then the bare Si(0 0 1) surface inside these windows was attacked with a KOH/isopropyl alcohol mixture, which anisotropically etches (0 0 1) and (0 1 1) Si planes, thus decorating the whole perimeters of the uncovered Si rectangles with (1 1 1) facets. The duration of the attack with the KOH/isopropyl alcohol solution determines the average size of the (1 1 1) Si facets. Thereafter the sacrificial silicon oxide was completely eliminated by a buffered oxide etch solution and Ge was deposited in a chemical vapour deposition reactor. The length of the (1 1 1) facets

Fig. 18. AFM images of 8 ML Ge deposited at 600 °C onto different vicinal substrates with a 500 nm Si buffer: (a) Si(0 0 1) 1.5° off; (b) Si(1 1 1) 1.5° off towards [1 1 –2]; (c) Si(1 1 1) 1.5° off towards [1 –1 2]. Scan size is 3 × 3 μm² (from ref [131]). (Reprinted figure with permission from I. Berbezier, A. Ronda, Phys. Rev. B 75 (2007) 195407. Copyright © 2007 by the American Physical Society).

Fig. 19. AFM images of Ge islands on Si substrates patterned with stripes along the [1 1 0] direction by a combination of holographic lithography and reactive ion etching. The upper four panels display Ge islands on a bare Si surface; the lower two panels display Ge islands on a GeSi buffer (from ref [169]). (Reprinted figure with permission from Z. Zhong, A. Hallilovic, M. Mühlberger, F. Schäffler, G. Bauer, J. Appl. Phys. 93 (2003) 6258. Copyright © 2003 by the American Institute of Physics).
governs the nucleation dynamics. Within several tens of nm of the (0 0 1) plane (at 650 °C deposition temperature), the (1 1 1) facets are denuded of three dimensional nuclei. Therefore nucleation is suppressed on (1 1 1) facets below a critical length. In Fig. 20, the depletion zone is discernible on the edges of the (1 1 1) facets close to the (0 0 1) plane, while no such zones emerge on the borders between (1 1 1) facets. These results are interpreted with a net flux of Ge adatoms from the (1 1 1) facets to the (0 0 1) plane, driven by chemical potential gradients and limited by the effective adatoms diffusivity. Interestingly Schwarz–Selinger et al. achieved a qualitatively reversed pattern of preferential nucleation by creation of craters of vicinal Si(0 0 1) facets through ns laser pulses bombardment (lateral resolution in the micron length scale) [174]. Under these circumstances the chemical potential gradients drive a Ge adatom accumulation towards the vicinal Si(0 0 1) facets, where strain relaxation may be more effective.

Among the advantages of conventional lithographical methods are its parallel nature and pattern versatility. Its main disadvantage is its well-known resolution limits, which may be incompatible with the precise definition of nanostructures architectures.

In 1996 Nguyen et al. demonstrated the first MOS device (an array of transistors and capacitors) drawn by extreme ultraviolet lithography at a wavelength of 13.5 nm [175]. The exploitation of this spectral window translates into the possibility of defining features down to few tens of nm. However the use of this technology is hampered by the large absorption of optical components which are required for conventional mask lithography. Grützmacher et al. introduced the use of extreme ultraviolet interference lithography (EUVIL) and reactive ion etching to realize three dimensional crystals of nanostructures with high precision [176]. This method proved to support pattern control in the subnanometer regime and pattern periodicity below 30 nm. Subsequent to extreme ultraviolet exposure and photoresist development, a periodic pattern of square depressions was transferred into the substrate by reactive ion etching. Then the deposition of a buffer layer led to the transformation of the square depressions into sharp pits with a depth of a few nm. Upon deposition of Ge, three dimensional islands were observed to form in these sharp pits with a low density of crystalline defects, which provided a uniform array of dots. The dots are understood to fill the sharp pits prepared by EUVIL and reactive ion etching due to the microscopic analogue of capillary interactions (i.e. essentially chemical potential effects) [177]. Finally a whole three dimensional lattice of nanostructures with quantum confinement behaviour was achieved by the recursive overgrowth of substrate and island materials, as shown in Fig. 21. As in the self-organization discussed in Section 4, order progressively improves due to elastic strain energy interactions among coexisting islands. The size uniformity of the islands improves with the crystalline perfection as well. The principal concern with the application of an interference lithography approach is its intrinsic restriction to periodic patterns.

5.1.1.3. Electron and ion beam lithography. Extending the reach of photolithography to the nanoscale is an enormous challenge. This has prompted the development of unconventional lithographic approaches. Popular alternatives in the context of the heteroepitaxial growth of semiconductors are electron beam lithography (EBL) and ion beam lithography (IBL). The principle behind the EBL is similar to optical lithography except that the resist is exposed to

Fig. 20. AFM image of a negative Si pyramid after Ge growth. Scales are in units of micrometres (from ref [173]). (Reprinted figure with permission from S.S. Ferng, T.H. Yang, G. Luo, K.M. Yang, M.F. Hsieh, D.S. Lin, Nanotechnology 17 (2006) 5207. Copyright © 2006 by the Institute of Physics Publishing).

Fig. 21. TEM (a) and AFM (b) micrographs of 10 period stacks of Ge islands and Si spacer layer (10 nm) deposited onto an EUVIL patterned substrate. The inset in panel (a) shows a close-up of a dot in the first layer. The histogram in panel (c) illustrates the narrow size distribution in diameter and height of the Ge dots at the top surface of the three dimensional nanostructures crystal (from ref [176]). (Reprinted figure with permission from D. Grützmacher, T. Fromherz, C. Dais, J. Stangl, E. Müller, Y. Ekinci, H.H. Solak, H. Sigg, R.T. Lechner, E. Wintersberger, S. Birner, V. Holý, G. Bauer, Nano Lett. 7 (2007) 3150. Copyright © 2007 by the American Chemical Society).
a focused electron beam, which is modulated and scanned through
the surface (normally with no mask). While the electron beam may
be focussed down to a few nm, the resolution of this technique is
limited to a few tens of nm, mainly due to secondary electron
scattering through the resist [178]. For instance EBL was used to
pattern GaAs(0 0 1) substrates prior to InAs deposition [82,179].
Arrays of sharp pits and shallow holes may result into different
effects with preferential nucleation within the sharp pits and
between the shallow holes respectively. Kiravittaya et al. used EBL
to realize three dimensional crystals of nanostructures with
enhanced size uniformity and specific photoluminescence finger-
prints (see Fig. 22) [82].

IBL may be exploited in different ways. One way is by
modification of a resist essentially as in the EBL, with a better
resolution associated with a lower energy of secondary electrons
[178]. Another way is by direct sputtering of the semiconductor
substrate, which may yield even better resolution and yet suffer
from further issues including redeposition and surface contami-
nation [180]. The lateral resolution attainable by IBL is a few tens of
nm. The notion of modulating nucleation dynamics by exploiting
arrays of holes defined by IBL is well established [105,180–183].
Depending on kinetic conditions such as substrate temperature
and growth rates, three dimensional islands may preferentially
nucleate either at the hole centres or at the hole perimeters. The
profound modification induced by these methods may give
additional opportunities. For instance Portavoce et al. demonstrat-
ed a selective attenuation of the WL critical thickness in areas with
high hole densities in the Ge/Si(0 0 1) system. This may be
exploited to suppress island nucleation away from the holes much
beyond the adatoms diffusion length [182]. These authors realized
complex architectures of dots e.g. to simulate a QCA adder circuit
embedded in a flat environment (with several hundreds dots
accurately positioned in an artificial macromolecule and less than
10% defects in this case, see Fig. 23).

Despite the remarkable progress in these technologies, the
choice of kinetic conditions to achieve effective control over QD
positions is critical, and the success rate is typically 50% or less
[179,180]. Another notable drawback of EBL and IBL is the serial
nature of the process, which does not appear as a sustainable
alternative to pattern large wafers. So far the principal technologi-
cal contribution expected from these approaches is in the
production of masks e.g. for optical lithography and other
unconventional methods.

5.1.1.4. Nano imprint lithography. A topical challenge in the design
of unconventional lithography methods is the combination of the
versatility of serial approaches and the sustainability of parallel
approaches. In 1996 Chou et al. proposed a new paradigm defined
as nano imprint lithography (NIL), [184] based on the transfer of an
arbitrary pattern embossed into a solid mold onto an arbitrary
substrate. In the original proposal, the technique rested on the
compression of a solid mold to create a depth contrast pattern in a
thin resist film deposited on a substrate, and then on a reactive ion
etching and lift-off procedure to transfer the pattern to the
substrate through the entire resist thickness. The fabrication of
metal patterns with a feature size of 25 nm and a period of 70 nm
was demonstrated. Since 1996 a diversity of variants of NIL have
been conceived, based on alternative concepts to imprint the
substrate, e.g. by use of thermoplastic, photo-sensitive and
electrochemically active polymers. Already in 2002 Chou demon-
strated the feasibility to realize NIL in silicon substrates with no
need for additional resists and chemical treatments, which was
referred to as laser-assisted direct imprint [185]. In this approach a

![Fig. 22. Height histograms from the outermost layer of InAs/GaAs QDs grown on patterns with 210 nm periodicity (a and b), 160 nm periodicity (c), and no pattern (d). Insets show corresponding 1 x 1 μm² AFM images (from ref [82]). (Reprinted figure with permission from S. Kiravittaya, A. Rastelli, O.G. Schmidt, Appl. Phys. Lett. 88 (2006) 043112. Copyright © 2006 by the American Institute of Physics).](image1)

![Fig. 23. AFM images of arrays of Ge islands grown on FIB-patterned Si(0 0 1) substrates: (a) square arrays with different periods; (b) simulation of a QCA adder circuit. In both cases, note the ability to maintain large regions free of islands (from ref [182]). (Reprinted figure with permission from A. Portavoce, M. Kammler, R. Hull, M.C. Reuter, F.M. Ross, Nanotechnology 17 (2006) 4451. Copyright © 2006 by the Institute of Physics Publishing).](image2)
conditions the tilt angle facilitates the possibility for these dynamics [195]. Under these expected periodicity becomes too small (<8 nm), which annihilates the possibility for these dynamics [196,197]. When a twist angle 20° is fabricated, a = 0.543 nm being the lattice parameter of silicon. For instance periodicities of 100 and 10 nm for twist angles of 0.15° and 1.6° respectively [195].

These templates may serve to modify the nucleation of three dimensional islands, e.g. either by exploiting the strain profile associated with the interfacial dislocation network, [195,196] or by introducing a selective chemical etching step to superimpose a groove network above the interfacial dislocation network [194]. For instance on the so-modified Si(0 0 1) substrates, Ge self-organizes into three dimensional islands localized and limited within the defect network, thus mimicking its symmetry and periodicity [194,195]. Ordered arrays of ultra small Ge hut clusters may be obtained, with typical size as small as a few nanometres, when the ratio of silicon wafer thickness to interface dislocation periodicity is small. Fig. 24 is taken from ref [195] and proves the possibility to modulate the degree of self-ordering with the thickness of the Si(0 0 1) wafer.

On account of its relative cost-effectiveness, efficiency in the definition of the spatial arrangement of the three dimensional islands and flexibility in the design of the periodicity of the dislocation network, this method may become of interest for applications where a dense array of small and uniform nanostructures is needed, e.g. for lasers and biosensors. Nonetheless in this architecture relevant device performances require verification due to the presence of a dense dislocation network in close proximity to the nanostructures. Dislocations, similarly to other kinds of crystal defects are known to modify the optoelectronic properties of semiconductor materials, the associated intragap levels possibly acting as traps for charge carriers, thus affecting the overall hole-electron generation and recombination mechanisms [197,198].

### 5.1.2. Design of dislocation networks

The use of step edges as in the examples above to engineer the kinetic framework and enthalpic landscape of the substrate is a popular choice. However other approaches to modify the substrate may be feasible as well. Another interesting example which exploits the design of artificial templates of defects in the substrate to dictate preferential nucleation sites originates from the fine manipulation of dislocation networks, which is a recent achievement [188]. This option may be implemented e.g. by the technology to cleave (e.g. by ion implantation) and bond thin silicon wafers on solid silicon substrates [189–191]. When a Si(0 0 1) wafer is bonded onto a Si(0 0 1) substrate under controlled conditions, a regular pattern of interfacial dislocations may emerge, the spacing and orientation of the super lattice depending on the mutual orientation between wafer and substrate [192,193]. When a twist angle ϕ around the surface normal is used, a network of twist dislocations with periodicity \( \lambda(\phi) = a/2\sqrt{2}\sin(\phi/2) \) is fabricated, \( a = 0.543 \text{ nm} \) being the lattice parameter of silicon. For instance periodicities of 100 and 10 nm are achieved with twist angles of 0.22° and 2.2° respectively [194]. In contrast when the twist angle is too large (e.g. >20°) the expected periodicity becomes too small (<1.1 nm), which annihilates the possibility for these dynamics [195]. Under these conditions the tilt angle \( \theta \) along the surface orientation may play the principal role and define a network of tilt dislocations with periodicity \( \lambda(\theta) = a/4 \sin(\theta/2) \). This equation corresponds to

![AFM images of arrays of germanium dots grown on silicon-bonded films: (a) 15 nm thick, and (b) 30 nm thick (from ref [195]).](Image)
repulsive barriers for the diffusing Ge adatoms. Depending on the ratio of the diffusion length of the Ge adatoms to the separation between the Au pads, a range of different morphologies may be realized, including in particular regular and uniform arrays of three dimensional islands (see Fig. 25) [201].

On the other hand this approach holds the potential to allow for arbitrarily complex QD motifs, via a proper arrangement of the apertures realized in the shadow nanomask used to create the chemical inhomogeneities. The example mentioned above does not belong to the conventional kind of chemical modifications of semiconductor epitaxy, since the three dimensional islands nucleate on the bare substrate between the metal pads. However the inhomogeneous chemical pattern together with the thermal treatment before and during film deposition may cause contamination and interdiffusion, as revealed by the chemical maps in Fig. 25 (note the high gold fraction on the three dimensional islands’ surface). Finally we mention the potential to refine this approach to fabricate functional architectures of metal pads and semiconductor particles to conjugate concepts of plasmon enhancement and QDs e.g. for novel light emitting devices [204].

5.2. Deposition flux modulation

All the top–down interventions reviewed so far rest on a procedure which includes at least two steps, i.e. first the design of a pattern on the substrate and subsequently the deposition of a film material. The pattern is intended to modulate the kinetic framework and enthalpic landscape of the substrate, thereby dictating preferential sites for the nucleation of three dimensional islands, which generally occurs through the interplay or competition of kinetic and enthalpic effects. However the equations mentioned in Section 4 suggest a more immediate approach to pattern the nucleation probability by modifying the mobile adatoms density. This may be achieved by modulation of the deposition flux (e.g. from a physical or chemical vapour source), i.e. by relative enhancement of the deposition flux where nucleation is desired roughly within the mobile adatoms diffusion length. Fig. 26 displays a possible scheme to realize this concept by use of a shadow mask.

As we shall discuss below, the principal challenge in this approach, which is otherwise ubiquitous in a variety of fabrication fields, is its application to the nanoscale [87].

5.2.1. Micro or nano sphere lithography

In 1982 Deckman and Dunsmuir introduced a versatile concept to pattern substrates, which was originally referred to as natural lithography, [205–207] and which may be implemented to modulate the deposition flux intensity. To begin with this model approach, one or two layers of micro or nano spheres (typically composed of polystyrene or silica) of consistent diameter are uniformly spread onto a substrate, e.g. by spin coating [207]. These spheres naturally self-organize into a close packed hexagonal lattice. In turn the pattern of solid closures and interstitial apertures among solid closures drawn by this hexagonal lattice of spheres may be exploited for lithographic patterning.

This technology may be used in a variety of approaches, including e.g. as a mask for various etching methods to pattern the substrate before deposition, [206–208] essentially as in the cases treated above. Alternatively it may be applied for instance as a mask to etch the film after its deposition. A recent example in this context is the work by Cong et al., who combined nano sphere lithography and oxygen reactive ion etching to sculpt a sheet of graphene which was previously transferred by mechanical cleavage onto a silicon substrate coated with thermal silica [209]. The result was a hexagonal array of graphene disks with diameters and separations proportional to the nanospheres diameters (in the range of several hundreds of nm).

Another alternative opportunity to exploit micro or nano sphere lithography may be as a mask for selective deposition [207,210]. In this concept the film material is deposited directly through the interstitial apertures, with no additional substrate intrusion. This is a radical innovation with respect to the cases treated above, since the modification of the pattern of nuclei is not pursued by a preliminary modulation of the kinetic framework and enthalpic landscape of the substrate. Rather, the adatoms density is modified by a direct modulation of the deposition flux over the micron and nano length scales. To our knowledge, this approach has not been applied in the heteroepitaxy of semiconductors to order nanostructures with quantum confinement profile thus far.
Its principal limitations include a poor versatility of accessible patterns and practical difficulty to manage and lift-off the micro or nano spheres during and after deposition [207].

5.2.2. Nanomask technology

The example reported under subparagraph Design of chemical patterns also suggests another possibility to modulate the deposition flux by using the shadow nanomask technology [199,200]. This approach consists in transferring the conventional stencil technology to the nanoscale, which was mainly implemented by Brugger and co-workers [87,199,211]. In brief, a shadow mask with an arbitrary pattern of hollow apertures is brought into close proximity to the substrate, prior to deposition. Deposition results into the transfer of the negative pattern on the substrate under minimally intrusive conditions. The principal challenge is the realization of the shadow nanomask, which is to be obtained from ultra thin membranes (e.g. silicon nitride films, ~100 nm thick) to enable the replica of architectures with lateral features of the order of few tens of nm. The flexibility and effectiveness of this method has been demonstrated on different and functional material pairs [87,199,200,212–216].

For instance Cojocaru et al. combined the shadow nanomask approach with pulsed laser deposition (PLD) [217,218], to pattern Ge/Si(0 0 1) three dimensional islands into desired architectures (see Fig. 27) [212]. The WL critical thickness was exceeded only below the hollow apertures, which restrains nucleation to the desired positions. The kinetic conditions applied during and after deposition give control over the number of nuclei per hollow aperture, which exhibit excellent epitaxial relationship to the substrate.

Due to its parallel profile, versatility, compatibility with full wafer treatment and possibility to clean and recycle the shadow nanomask for multiple copies, this approach may be sustainable and feasible for a variety of industrial applications [87,199]. Among the relevant critical issues we mention problems related to shadow nanomask fabrication and use. These include: excessive strain in the ultra thin membranes, which limits the choice of possible patterns before stencil collapse and, more importantly, the overall size of the stencil (which realistically cannot be fabricated in wafer size dimensions as it is too fragile); smearing of the deposited pattern, which may result from the geometrical layout comprising the relative separations and alignment of the deposition source, shadow nanomask, substrate and interposed gap; and clogging of hollow apertures with high aspect ratio (ratio of the membrane thickness to the hollow diameter), which may develop even during deposition and deteriorate the pattern replica. Despite these issues, the nanomask technology is receiving increasing interest from the scientific and corporate communities, and we foresee exciting developments in the near future.

6. Conclusions and perspectives

In this review, we have discussed the self-organized growth of semiconductor heteroepitaxial nanostructures at surfaces, with particular interest for future quantum confined systems such as e.g. QDs. Model QDs may be realized by inclusion of chemical and strain inhomogeneities of suitable lateral dimensions in semiconductor materials of technological relevance including silicon and III/V compounds, which may be implemented in heteroepitaxial processes. We have emphasized the opportunities defined by synthesizing nanostructures with potential QD behaviour with narrow size distributions in ordered two dimensional and eventually three dimensional architectures for various types of device applications. We chose to narrow the scope of this review to the case of semiconductor nanostructures, although it should be noted that alternative systems such as semiconductor nanowires (both homoepitaxial and heteroepitaxial) and inorganic nanotubes are an emerging class of quantum confined systems with interesting optoelectronic properties apt for applications in sensing, thermoelectrics, solar energy conversion and electronic devices, partly due to their marked anisotropic shape [219–249].

The intrinsic physical limits that are being met by conventional lithographic techniques used in the semiconductor industry have called for new paradigms for device fabrication in general and surface patterning in particular. While the bottom–up approach has become very popular for scientific studies, its limitations in terms of achieving a high degree of ordering and narrow size distributions have prompted a wealth of hybrid proposals where top down and bottom–up strategies are brought to convergence [250,251]. The investigation of mere bottom–up processes reveals a rich phenomenology, which should be understood to gain control over the locations of the nanostructures with respect to each other and their environment, including kinetic and enthalpic interactions between neighbouring dots and between each dot and the substrate. In particular the correlation between specific features in the substrate and the emergence of preferential sites for the nucleation of the nanostructures gives opportunities to engineer external interventions. In this context, an emerging concept is that of ‘surface cues’, [252] i.e. the ability to pattern a surface with structures that will act as cues, thus guiding the nucleation and ordering of inorganic nanostructures (as we have described herein), or the adsorption of organic molecules, [253–258] or even the behaviour of living organisms such as cells [259–263]. This is a hopeful area of surface chemistry, with applications e.g. in
semiconductor fabrication, supra-molecular ordering in two dimensions and biomaterial implants. The use of such surface cues presents endless opportunities to control the formation of long-range ordered architectures at surfaces. In the case of semiconductor nanostructures, embodiments of surface cues of principal interest include e.g. atomic step edges, bunches and combinations thereof (encompassing a variety of architectures such as mesas, sawtooth profiles, troughs, etc.), networks of dislocations and chemical patterns, which may be dictated through an external intervention. These surface cues modulate the kinetic framework and thermodynamic landscape of the substrate, and can be integrated with intrinsic concepts of atomic diffusion lengths and capture zones to manage the precise locations of epitaxial nanostructures. The application of surface cues may be an intrusive intervention. Another possibility is to leave the substrate unperturbed and modify the local density of deposited material in the heteroepitaxial process. While this notion follows the conventional approach of mask lithography, its translation down to the nanoscale is a recent achievement which still poses severe challenges.

The proof of individual nanostructures with quantum confinement properties is a solid achievement which has attracted much interest from academic and, more recently and indicatively, corporate counterparts. In the future, the exploitation of such nanostructures will require intense capacities to control their size, shape and precise arrangement in order to engender full quantum systems, including molecules, macromolecules and crystals. In turn this objective will mark an extraordinary milestone in materials science and enable step-change technological innovation e.g. in the fields of optics, quantum computation, catalysis, photovoltaics, diagnostics, biodiagnostic, etc. We are aware that such one revolution may spring from some of the pioneering results mentioned in this contribution.

In this contribution, we have sorted a number of critical concepts in the fabrication of two dimensional and possibly three dimensional architectures of semiconductor nanostructures. We have proposed a few examples which in our opinion illustrate well the merit, exploitation, limitations and perspectives of these concepts. We have avoided the use of unnecessary mathematics, theoretical constructs and experimental details, which may be found in the various references, in an attempt to deliver a convenient guidebook of immediate fruition to the experimental scientist. We hope that our key of the reading will be of inspiration to the novice and interest to the expert, and become a useful reference in the thriving context of producing functional arrays of quantum confined systems.

Acknowledgements

F. Rosei is grateful to the Canadian Research Chairs Program for partial salary support. Work on QDs in Prof. Rosei’s laboratory has been funded by NSERC of Canada (Discovery Grants), FQRNT (Quebec), the Canada Foundation for Innovation and INRS start-up funds.

References
