Perspective: Fundamentals of coalescence-related dislocations, applied to selective-area growth and other epitaxial films

Cite as: APL Mater. 6, 120903 (2018); https://doi.org/10.1063/1.5047945
Submitted: 10 July 2018. Accepted: 26 September 2018. Published Online: 19 December 2018

William E. McMahon, Michelle Vaisman, Jeramy D. Zimmerman, Adele C. Tamboli, and Emily L. Warren

COLLECTIONS

This paper was selected as Featured

ARTICLES YOU MAY BE INTERESTED IN

Perspective: Uniform switching of artificial synapses for large-scale neuromorphic arrays
APL Materials 6, 120901 (2018); https://doi.org/10.1063/1.5049137

Improving optoelectronic materials by using virtual dislocations to visualize and understand these defects in thin films
Scilight 2018, 510005 (2018); https://doi.org/10.1063/1.5085399

Research Update: Beyond graphene—Synthesis of functionalized quantum dots of 2D materials and their applications
APL Materials 6, 120701 (2018); https://doi.org/10.1063/1.5067250
Perspective: Fundamentals of coalescence-related dislocations, applied to selective-area growth and other epitaxial films

William E. McMahon,1 Michelle Vaisman,1,2 Jeramy D. Zimmerman,1,3 Adele C. Tamboli,1 and Emily L. Warren1

1National Renewable Energy Laboratory, Golden, Colorado 80401, USA
2Department of Electrical Engineering, Yale University, New Haven, Connecticut 06520, USA
3Department of Physics, Colorado School of Mines, Golden, Colorado 80401, USA

(Received 10 July 2018; accepted 26 September 2018; published online 19 December 2018)

Although selective area growth (SAG) and coalesced SAG (cSAG) have been utilized extensively for many years to moderate the material quality of lattice-mismatched films, the geometrical factors controlling dislocations in coalesced films are difficult to visualize, and some confusion regarding the topology of dislocations for cSAG still persists. In this paper, we describe the topology of dislocations during island coalescence for cSAG, framed in terms of fundamental dislocation properties and virtual dislocations, which are used as a helpful visualization tool. We also show how the results and methods are generally applicable to coalescence of any other epitaxial film. © 2018 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/). https://doi.org/10.1063/1.5047945

INTRODUCTION

High crystalline quality semiconductor materials are critical for many electronic and optoelectronic applications. As the need grows for material bandgaps beyond those of semiconductors commercially available as substrates, research has expanded toward lattice mismatched (LMM) materials. Lattice mismatch introduces strain and defects into these materials, which are often detrimental to device performance. Many techniques have arisen over the years to mitigate high dislocation densities associated with lattice mismatch: metamorphic graded buffer growth,1,2 thermal cycle annealing,3 dislocation filters,4 and selective-area growth (SAG). While each technique has its merits and uses, in this manuscript we focus on better understanding the dislocation mitigation mechanisms enabled by coalesced selective-area growth (cSAG) of lattice mismatched films. Nonetheless, the ideas and conclusions are critical for understanding the coalescence of any epitaxial material system, including porous materials, island growth [e.g. Volmer-Weber (VW)], heteroepitaxy with coincident site matching or interfacial misfit (IMF), and forms of dislocation “blocking” such as by voids, pores, second phases, and strained layers.

Selective-area growth is well established as a method for growing small areas of dislocation-free LMM material.5,6 The concept of SAG was originally proposed by growing separate islands of material,7 which has been demonstrated to eliminate some of the strain-related issues with LMM epitaxy.5–10 This concept has also been extended to epitaxial growth of material within a via between regions of masking material and enables defect reduction by aspect ratio trapping (ART)11 and/or lateral epitaxial overgrowth (LEO).12

SAG and cSAG approaches are used to fabricate a myriad of semiconductor devices including solar cells,13–18 lasers,19–21 light emitting diodes,22 and photodetectors.23 These techniques have been utilized in multiple material systems, notably GaN epitaxy on SiC,24,25 and Si,26 and Ge,27 GaSb,28 InP,29 and GaAs,19,30 epitaxy on Si, and effectively remove threading dislocations (TDs) from each individual small area of uncoalesced LMM epitaxial material.19,31,32
Although significant threading dislocation density (TDD) reductions have also been reported for LMM films grown via coalesced SAG,\textsuperscript{14,23} this should not simply be ascribed to ART. As will be described in this paper, ART cannot, in and of itself, remove TDs from LMM cSAG epilayers. Furthermore, reports in the literature often use different methods to quantify the material quality of SAG and cSAG materials, making it challenging to directly compare different patterning and growth approaches. For example, for AlGaAs and GaAs grown on Si, reported TDDs range from 10\textsuperscript{4} cm\textsuperscript{-2} (etch pit density)\textsuperscript{33} to 10\textsuperscript{9} cm\textsuperscript{-2} [cross-sectional transmission electron microscopy (XTEM)]\textsuperscript{34} and some reports show that TDD increases after coalescence of material from individual SAG vias.\textsuperscript{35}

In this paper, we address a significant and persistent source of confusion—the topology of dislocations in coalescing systems. This confusion persists largely because the geometry of coalescing systems, over the wide range of length scales relevant to dislocations, is difficult to visualize. Here we provide a detailed examination of the topological origins of threading dislocations in coalescing materials and show how virtual dislocations (VDs) can be used as a visualization tool to facilitate understanding. Although this work was motivated by LMM cSAG, the basic concepts are generally applicable to many other situations for which the dislocation topology is difficult to visualize.

**COALESCED SELECTIVE-AREA GROWTH VERSUS VOLMER-WEBER GROWTH**

In most applications of SAG [Fig. 1(a)], the substrate surface is coated with a masking material (typically an oxide) which is then patterned with an array of vias. For organometallic vapor phase epitaxy (OMVPE) on an appropriately patterned substrate with a sufficiently large aspect ratio,\textsuperscript{36} epilayer growth will occur only in the vias, and aspect ratio trapping will direct LMM-related dislocations toward the sidewalls of the material grown in each via. This enables the growth of small areas of non-contiguous dislocation-free LMM material.

For some materials and applications, the volume of material created by SAG is adequate, but for large-area devices such as solar cells, the epilayer should ideally cover the entire substrate. To achieve such a large area, the area of each “epipillar” can be expanded by the LEO of “epiwings” [terms defined in Fig. 1(a)]. Coalesced SAG accomplishes this by coalescing adjacent epiwings into a contiguous “cSAG epilayer” [Fig. 1(b)]. However, this coalescence typically results in a large TDD at the coalescence boundaries.\textsuperscript{35}

One of the main themes of this paper will be the similarities between cSAG and Volmer-Weber (VW) growth. In VW, individual islands nucleate [Fig. 1(c)] and then coalesce into a continuous VW epilayer [Fig. 1(d)]. For both LMM cSAG and LMM VW growth, subtle misalignments during coalescence can create TDs in the resulting epilayer. The goal of this paper is to facilitate the visualization of these coalescence-related misalignments and their consequences.

![Fig. 1. Terminology for (top) coalesced selective-area growth “cSAG” and (bottom) Volmer-Weber “VW” growth: (left) before coalescence and (right) after coalescence. In cSAG growth, the epilayer material nucleates on regions of the substrate exposed within the vias in the mask material and then grows upwards as an epipillar and then laterally as epiwings to finally coalesce into a cSAG epilayer. In VW growth, islands nucleate on an unmasked substrate and then coalesce into a contiguous epilayer.](image-url)
COALESCENCE-RELATED DISLOCATIONS

Here we will describe specifically how the dislocation content of a cSAG film evolves as it coalesces. After coalescence, the dislocation content of a LMM cSAG epilayer will continue to evolve as it does with any LMM epilayer, but these processes have been (and continue to be) extensively studied elsewhere. Our discussion of cSAG coalescence will be based on the following three fundamental dislocation properties:

1. Topologically, all dislocations must form complete loops. [As will be discussed extensively in this paper, virtual dislocation (VD) segments can be used to complete a loop wherever a dislocation exits through the surface of a material.]

2. Pairs of threading dislocation (TD) segments can annihilate and connect two misfit dislocations if, and only if, their parent misfit dislocations (MDs) have the same Burgers vector and line sense.

3. A Burgers circuit with a non-zero closure error encircles a dislocation.

These properties and associated terminology are well established and thoroughly described elsewhere and will not be re-derived here. The following discussion will be based upon hypothetical examples which apply these dislocation properties to cSAG coalescence. Most of the discussion is directed toward LMM cSAG, but, toward the end, implications and opportunities for lattice-matched (LM) cSAG and related growth geometries will be considered.

Continuity of dislocations

To begin, Fig. 2 illustrates why aspect ratio trapping cannot, in and of itself, prevent threading dislocations from continuing upwards into a LMM cSAG epilayer. Figure 2(a) illustrates the initial phase of cSAG, in which a single dislocation has formed in a LMM epipillar. Before coalescence, this dislocation can exit the side surfaces of the epipillar via aspect ratio trapping. This creates a dislocation-free top surface upon which additional uncoalesced dislocation-free material can be grown. However, isolated dislocations such as the one shown in Fig. 2(a) will reappear in the overlying material if adjacent epipillars coalesce into a continuous film, as shown in the “allowed” configuration [Fig. 2(b)]. The “forbidden” configuration [Fig. 2(e)] is not allowed. To understand why, the “allowed” and “forbidden” cases from Fig. 2 are extended via a thought experiment in which the crystalline material replaces the mask material. In the allowed case [Figs. 2(b)–2(d)], the resulting dislocation terminates at an exterior surface [Fig. 2(d)]. In the forbidden case [Figs. 2(e)–2(g)], the two ends of the dislocation terminate at an interior surface. The topological impossibility of this becomes clear when the mask material is completely replaced with epilayer material, and the dislocation terminates within the solid [Fig. 2(g)].

FIG. 2. During cSAG, an isolated dislocation (red dashed line) can be removed from an individual uncoalesced “epipillar” (a), but will continue into the overlying epilayer to reach an exterior surface after coalescence (b). The configuration shown in (e) is not allowed because the dislocation terminates at an interior surface. To understand this, some additional hypothetical cases have been added (inside the dashed box). In (c) and (d), mask material in (b) is gradually replaced with epilayer material to obtain (d), which is allowable. In (f) and (g), mask material in (e) is replaced with epilayer material to obtain (g), which is strictly forbidden.
FIG. 3. A reproduction of Fig. 2, with virtual dislocations (VDs) added as thin red dotted lines. As growth proceeds, the VDs in (a) indicate the need for TDs to pass through the coalesced epilayer in (b). In the context of this paper, VD segments are used to schematically complete dislocation loops. If its connectivity is not changed, the exact path of each VD segment is arbitrary. By adding VD segments like these, dislocation loops can be created for every allowed dislocation configuration in this paper (and in general).

To avoid the creation of schematic diagrams with forbidden configurations such as those shown in Figs. 2(e)–2(g), “virtual dislocations” (VDs) can be used as a helpful visualization tool. Figure 3(a) is identical to Fig. 2(a), but a VD has been added to form a complete loop (thin red dotted line). Figure 3(b) is identical to Fig. 2(b), but, once again, segments of VD have been added outside of the semiconductor material (i.e., within the masking material and outside the epilayer) to form a loop. Using the same methods, loops can be formed for Figs. 2(a)–2(d) but not for Figs. 2(e)–2(g).

In the context of this paper, the sole purpose of VDs is to ensure that all topological rules for dislocations are followed, so the exact path taken by a VD is not important and should not be used to compute any elastic energies. The essential point is that for any allowable configuration of dislocations, it must always be possible to form complete dislocation loops by using VD segments to connect together points where dislocations exit through a surface of a material. When this is done, the Burgers vector associated with each VD must match the dislocation to which it is connected. By extension, VDs can be combined, but only if the associated rules for combining Burgers vectors are followed.

Figure 2 also helps illustrate one potential source of confusion produced by XTEM images of threading dislocations. Because the volume of material imaged in an XTEM foil is thin, it rarely includes the entire length of a dislocation. Therefore, when images similar to Figs. 2(e)–2(g) are obtained, it does not mean that a dislocation terminates within the epilayer or at an internal surface, but rather that the dislocation that originally reached the top surface, as shown in Figs. 2(b)–2(d), has exited out of the front or back of the thin slab of material used for XTEM. (An experimental example can be found in Fig. 1 of Ref. 32.)

For this reason, the threading dislocation content of an epilayer is more accurately assessed via a plan-view geometry using techniques such as planar transmission electron microscopy (TEM), electron channeling contrast imaging, or electron beam-induced current mapping. Furthermore, an area appropriate for the TDD of the epilayer should be measured. For consideration, a TDD of $10^8 \text{ cm}^{-2}$ corresponds to one TD per square micron, which is much larger than the area typically included in an XTEM sample.

It should also be mentioned that some common characterization methods such as defect selective etching (DSE) and x-ray diffraction (XRD) can lead to inaccurate TDD values. For example, if defect selective etching conditions are not properly optimized (due to over- or under-etching, for example), there will not be a one-to-one correspondence between TDs and etch pits. When characterizing material quality with XRD, both the optics and sample geometry must be carefully considered. Non-optimal optics can lead to peak broadening, and different sample geometries can lead to complexities such as anisotropy within the sample. To unambiguously measure dislocation densities, multiple quantitative characterization techniques should be employed.

**Pairwise annihilation of threading dislocation segments**

Next, we discuss the second dislocation property—that pairs of TD segments can annihilate and connect two misfit dislocations if, and only if, their parent MDs have the same Burgers vector and line sense. However, this rule simply addresses the possibility that two TD segments could annihilate. Whether or not they actually annihilate also depends upon energetics and kinetics, and these factors can become very complex. Here we merely show that voids in a coalescing material can,
in principle, influence how sets of dislocations interact during coalescence. It should be emphasized that our intention here is not to provide a predictive theory but rather to provide some representative dislocation-interaction examples for the cSAG and VW growth geometries, in order to illustrate some subtle but important differences between the two geometries.

Figure 4 illustrates a few representative limiting cases for which all dislocations are glissile. (The line direction and Burgers vector of a glissile dislocation lie within the same glide plane. Glissile TD segments can, under the right circumstances, glide together to annihilate. The line direction and Burgers vector of a sessile dislocation are not coplanar with a glide plane. Sessile dislocations can only move via mass transport/climb.)

In Fig. 4, each dislocation has been assigned a Burgers vector representing (for example) one of the four strain-relieving glissile \(\langle 110 \rangle\) Burgers vector directions for a cubic epilayer material growing with a \{100\} surface and \{111\} glide planes. Although this example is constrained to glissile dislocations, the concepts are general and can be applied to any mixture of dislocations, regardless of whether they are sessile or glissile.

The shading in Fig. 4 represents, very schematically, the displacement field at the surface associated with each underlying dislocation line. In the following discussion, dislocations will only combine...
if these displacement fields overlap. Otherwise, they will behave as isolated dislocations, essentially
unaffected by the displacement fields of other dislocations.

Figure 4(a), left, shows two hypothetical islands as they are just beginning to coalesce during
Volmer-Weber growth. Each island is populated with MDs with the Burgers vectors indicated in the
top view [Fig. 4(a), lower left], and each of these MDs is connected to a VD segment which extends
upwards through the hole in the material formed by the two points of coalescence. All four of these
VD segments must be accounted for after coalescence.

Figure 4(a), right, shows one possible post-coalescence outcome in which the two \( b_1 \) VD seg-
ments annihilate pairwise, joining the \( b_1 \) MDs together. As a result, there are no \( b_1 \) TDs after
coalescence. By contrast, the two \( b_2 \) VDs become \( b_2 \) TD segments. During coalescence, this pair of
\( b_2 \) TD segments could have annihilated pairwise but did not. The reason they did not is that prior
annihilation of the \( b_1 \) TD segments has already created one long \( b_1 \) MD, and the associated local
over-relaxation of epilayer stress can “block” the motion (and thereby annihilation) of the two \( b_2 \) TD
segments. Eventually, if epilayer growth continues, it will become easier for the \( b_2 \) TD segments to
annihilate somewhere further away from the LMM interface, spatially separated from the localized
stress associated with the \( b_1 \) MD.

Figure 4(b), left, shows a pre-coalescence configuration for cSAG, in which there is some mask
material between the cSAG epipillars, but everything else is identical to the VW case shown in
Fig. 4(a), left. However, the post-coalescence dislocation configuration shown for cSAG is different
[Fig. 4(b), right]. The possibility for different outcomes for VW and cSAG has two origins. One
difference is that the energetics and kinetics are different. This is because dislocation interactions are
governed by energy associated with the strain fields surrounding the dislocations. In cSAG, some of
this material has been removed and replaced by the mask material, which alters the elastic forces
between interacting dislocations.

The second difference is that the relevant displacement fields are different. For VW growth,
coalescence occurs at or near the epilayer/substrate interface, where the displacement fields of
the dislocations are highly localized near their respective dislocation cores. For cSAG epilay-
ers, coalescence occurs much further from the substrate and thus much further from the relevant
underlying MD cores at the epilayer/substrate interface. As the distance from a dislocation core
increases, its associated displacement field broadens and becomes superimposed with the dis-
placement fields of other dislocations. Very schematically, the four virtual dislocation segments
in Fig. 4(b), left, annihilate in Fig. 4(b), right, because the atomic bonding during coalescence is
guided by the superimposed displacement fields, which, far from the dislocation cores, are
insensitive to the exact arrangement of the underlying MDs, so coalescence can proceed without
TDs.

By contrast, the MDs in Fig. 4(c), left, are more widely spaced such that the superposition of the
placement fields of adjacent MDs is negligible. In this case, the dislocations can interact pairwise
as shown in Fig. 4(c), right, to create “grown-in” TDs. The identity of the resulting TDs depends on
how the VDs combine, but there is no guarantee that the resulting TDs will be glissile. (Consider,
for example, the sessile “Lomer-Cottrell” locks that can form via the combination of two dissociated
glissile dislocations in a zinc-blende crystal.\(^ {39,40} \))

It is also worth mentioning here that configurations such as in Fig. 4(b) occur preferentially
at high TDDs. As the TDD decreases, configurations like Fig. 4(c) become more common, and at
the lowest TDD, configurations with isolated TDs (Fig. 3) become most common. Therefore, any
facilitation of the annihilation of TD segments via the mechanism illustrated in Fig. 4(b) would be
most effective at large TDDs and then become less effective as the TDD decreases. This is consistent
with the existence of many TDD reduction strategies which quickly bring the TDD down to the
\( 10^7–10^8 \) cm\(^{-2} \) range, but not much further.

Geometrically necessary TDs

Finally, we will discuss the consequences of the third fundamental dislocation property—that a
Burgers circuit with a non-zero closure error must encircle a dislocation. This property implies that
an alignment error of just one atom during the coalescence of a set of adjacent islands will create a
“geometrically necessary” TD in the coalesced epilayer.\(^ {48} \)
Our discussion of geometrically necessary dislocations will be split into three sections. The first section will apply Burgers circuit analysis to the crystal lattice of coalescing islands at the atomistic level. This is helpful for basic understanding, but creating atomic-scale diagrams for every relevant coalescence event is not practical. The second section will show how the same geometrically necessary dislocations can be understood using VDs, without needing to explicitly draw the underlying crystal lattice. And finally, the third section will discuss geometrically necessary dislocations created by Burgers circuits with non-zero closure errors in the z direction (normal to the growth surface).

**Geometrically necessary TDs from coalescence: An atomistic analysis**

Figure 5 illustrates atomistically how island coalescence can lead to grown-in “geometrically necessary” dislocations. In this example, the lattice mismatch is such that 8 epilayer atoms span the same distance as 9 substrate atoms. The periodic supercell representing a fully relaxed interface is created by stacking these two adjacent layers of atoms on top of one another [Fig. 5(a)]. Misfit dislocations follow lines of atoms with the maximum displacement from ideal interfacial bonding positions, and a periodic array of supercells will create a periodic interfacial MD network. For specific favorable sample geometries, near-surface interfacial MD networks have been imaged with transmission electron microscopy (TEM) and scanning tunneling microscopy (STM). Some examples of this are Ag/Cu(111),\(^{50}\) Ge/Si(111),\(^{51,52}\) InAs/GaAs(111),\(^{53–55}\) InN/GaN(0001),\(^{56}\) and GaSb/GaAs(111).\(^{57,58}\)

When two adjacent epilayer islands coalesce, there is the potential for misalignments that can create TDs. In Fig. 5(b), two adjacent islands are placed such that they could be subsequently connected by an additional supercell to form one large undistorted island. However, this alignment is statistically rare. If two islands nucleate randomly on the surface, the probability of perfect alignment is only 1/81 for the 8/9 LMM ratio used in this example. In most cases, two adjacent islands will be misaligned [Fig. 5(c)], and the epilayer must distort during coalescence. These distortions can cancel to enable coalescence without a TD [Fig. 5(d)] or accumulate such that coalescence encircles a TD [Fig. 5(e)].

These two cases can be understood by using Burgers circuit analysis. In Fig. 5(d), a Burgers circuit has been drawn through the four coalescing islands. This Burgers circuit starts in Island 1 and...
then travels 10 atoms in each direction. On the downward leg, there is a 1/3 atom shift to the left, but this is offset by a 1/3 atom shift to the right on the upward leg. The epilayer is distorted, but there is no closure error and therefore no encircled TD. In Fig. 5(e), the registry of the islands is slightly different. In this case, the downward, leftward, and upward legs of the Burgers circuit each produce a 1/3 atom shift to the left. This produces a one-atom closure error. Prior to full coalescence, this closure error is indicated by a VD. After the islands fully coalesce, the VD becomes a “geometrically necessary” TD.

Atomistic diagrams such as those in Fig. 5 could in principle be used to illustrate any arbitrarily complicated dislocation configuration during coalescence, but this quickly becomes tedious and cumbersome. To facilitate visualization, TDs resulting from island misalignment errors can also be described using VDs, as explained in the next Section (titled Geometrically necessary TDs from coalescence: A VD analysis).

**Geometrically necessary TDs from coalescence: A VD analysis**

This section will illustrate how virtual dislocations (VDs) can be used to facilitate visualization of the formation of geometrically necessary dislocations during coalescence. It will also be used as an opportunity to show how the same concepts apply equally to sessile dislocations.

Figure 4 illustrated a limiting case in which all dislocations were glissile (prior to coalescence). This would happen if the initial interface were coherent, and any strain relief were due to subsequent dislocation glide. At the other extreme, Figs. 6 and 7 illustrate a different limiting case in which all MDs are sessile “grown-in” 90° edge dislocations, as occurs for the coalescence of fully relaxed LMM islands during “interfacial misfit” (IMF) growth on a cubic {100} surface. In this example, each individual island (or epipillar) is fully relaxed through an interfacial MD array and, prior to coalescence, contains no TDs. However, as was just shown and will be shown again here, coalescence of TD-free islands does not guarantee the formation of a TD-free epilayer. Furthermore, the same underlying principles apply equally to VW growth (Fig. 6) and cSAG (Fig. 7).

Figure 6 illustrates a hypothetical island coalescence event for LMM VW growth. Figure 6(a) is a cross-sectional side view of the top view shown in Fig. 6(b). Figure 6(c) is a wider-range top view showing how a Burgers circuit can be drawn through the points of coalescence of these three islands.

**FIG. 6.** A hypothetical example for VW growth showing how small misalignments between adjacent islands before coalescence [(a)–(c)] can create a “geometrically necessary” threading dislocation (TD) after coalescence [(d)–(f)]. (a) is a side view of the cross-sectional cut-out DD′ in (b), and (b) is an enlarged view of (c), where three misaligned islands have begun to coalesce. In (d) and (e), all MDs (orange and brown lines) have joined together pairwise except for one, which must bend upwards to reach the surface as a TD. (d) is a side view of the cross-sectional cut-out EE′ in (e). (f) An enlarged schematic view of the atomic lattice near the resulting TD, whose Burgers vector \( \mathbf{b}_y \) is defined by the closure error of the Burgers circuit.
FIG. 7. A hypothetical example for cSAG growth showing how small misalignments between adjacent islands before coalescence [(a)–(c)] can create a geometrically necessary threading dislocation (TD) after coalescence [(d) and (e)]. (a) is a side view of the cross-sectional cut-out $FF'$ in (b), and (d) is a side view of the cross-sectional cut-out $GG'$ in (e). Aside from the mask/spacer layer, the initial conditions and resulting geometrically necessary TD formation are the same as for the VW example in Fig. 6. (f) For cSAG, the MD array at an epipillar/substrate interface is extended to become a virtual MD (VMD) array beneath the epiwings. These VMDs serve to indicate the registry of the epiwings with respect to the underlying substrate. If the mask/spacer material were replaced with epilayer material, the VMDs would become MDs with the same $(x, y)$ positions.

islands. Figures 6(d) and 6(e) are side and top views after coalescence, showing the resulting TD for this hypothetical example. Figure 6(f) shows the atomic lattice surrounding the resulting TD.

The lattice mismatch between the substrate and epilayer creates two orthogonal sets of regularly spaced MDs (orange and brown lines). In this example, the ratio of the epilayer/substrate lattice constants is 9/8, so the MD spacing is 8 times the primitive lattice constant of the epilayer in the plane of the interface. [This ratio of 8 epilayer atoms per 9 substrate atoms can be seen along the right edge of Fig. 6(f). Elsewhere in Fig. 6(f), this ratio changes as the strain field around the indicated TD alters the local lattice constant of the epilayer.]

In this geometry, the VD segments connected to any two parallel MDs from separate islands or epipillars can annihilate pairwise, as seen here in Figs. 6(b) and 6(e), and later in Figs. 7(b) and 7(e). This is because, for the same MD line sense, every orange MD has the same Burgers vector, so the associated VD segments can annihilate. The same is true for the orthogonal set of brown MDs. However, as will be discussed below, any small misalignments during coalescence can create unpaired MDs, which must continue upwards into the coalesced film as “geometrically necessary” threading dislocations.

To create the conditions needed to form a geometrically necessary TD, small island-to-island misalignments have been included in Fig. 6. In the $y$ direction, Island 1 is shifted down with respect to Island 3, and Island 2 has been shifted upwards with respect to Island 3. Although the underlying atomic lattice is not shown, each shift is about 1/3 of the surface atomic row spacing, as can be inferred from the relative positions of the MDs on adjacent islands. This creates a 1/3-row gap between Islands 1 and 2, which will close during coalescence, creating a TD. [These displacements match those in Fig. 5(e) but with a 90° clockwise rotation.]

The need for a geometrically necessary TD can be inferred from the pattern of MDs in Fig. 6(b). During coalescence, the lattice near the edges of each island will distort to bring MDs into alignment, enabling most MDs to be joined together pairwise by VDs. However, one unpaired MD will remain. Before coalescence [Figs. 6(a) and 6(b)], the unpaired MD is connected to a VD traveling
upwards through the hole in the coalescing film. After coalescence, this unpaired VD becomes a TD
[Figs. 6(d)–6(f)].

With regards to TD formation, the situation for LMM cSAG [Fig. 7] is essentially identical. To facilitate comparison, the misalignments in Figs. 6 and 7 are identical. The only difference between the two configurations is the mask/spacer layer. And, as was shown in Fig. 2, this mask material cannot, in and of itself, prevent TDs from reaching the epilayer surface.

For cSAG, MDs only occur at the epipillar/substrate interface, as shown in Fig. 7(f). However, as will be shown next, it is very useful to extend this array of MDs with an array of virtual MDs (VMDs). The VMDs in Fig. 7 are located where MDs would be if the spacer layer were thinned to zero thickness and eliminated, bringing the cSAG epiwings and substrate into contact without changing their lateral (x, y) registry.

The salient feature of the VMDs in Fig. 7 is that they convey information about the registry of each epiwing with respect to the substrate and can therefore be used to predict the existence or absence of threading dislocations in the cSAG epilayer after coalescence. When this is done, the one unpaired VMD must bend upwards toward the surface to form a VD before coalescence [Figs. 7(a) and 7(b)]. After coalescence, this VD becomes a TD [Figs. 7(d) and 7(e)]. In essence, Fig. 7 can be converted into Fig. 6 simply by removing the spacer layer and shortening the epipillars, until the epilayer and substrate come into contact. The resulting TDs will be the same.

Although subsequent glide, climb, and annihilation of TD segments can affect the final outcome, it is instructive to consider what TDD might be expected for a simplified representative case in which islands or epipillars nucleate with a spacing of 1 µm in both the x and y directions, with randomly chosen island misalignments. In simplified form, the problem reduces to computing the statistical probability that the random misalignments sum to zero, for which there will be no TD [as in Fig. 5(d)], or to a non-zero value, in which case a TD will be encircled [as in Fig. 5(e)].

Based on random statistics, x-direction misalignments for a Burgers loop passing through four coalescing islands (or epipillars) will sum to a non-zero value for approximately 1/3 of all coalescence events, as will y-direction misalignments. The 1/3 value is simply the probability that three randomly chosen numbers between −0.5 and 0.5 will sum to a number with a magnitude greater than 0.5, leading to a non-zero closure error such as the one shown in Fig. 5(e). [In this simplified analysis, minimization of localized lattice distortion at each point of coalescence will round any closure error to the nearest integer. If the magnitude of this integer is zero (unity), no (one) TD will be encircled.]

Ignoring z-misalignments and interactions between x and y, the resulting TDD for this example, with each island having an area of 1 µm², will be \( (1/3 + 1/3)/(10^{-4} \text{ cm}^2) = (2/3) \times 10^8 \text{ cm}^{-2} \). In other words, the TDD computed solely from geometric alignment errors will be similar to the areal density of the original island/epilayer nucleation sites. Conversely, the observed TDDs for LMM VW growth could be used to estimate the original areal density of islands/coalescence sites, and the resulting sub-micron distances are reasonable. Of course, the exact coalescence geometry can modify these numbers, and post-coalescence annihilation of TD segments further modifies the TDD (and therefore plays an important role in most TDD-reduction strategies).

Ultimately, this thought experiment indicates that the TDD due to alignment errors during the coalescence of LMM epilayers cannot (and should not) be neglected. Instead, information about coalescence-related TDs should be sought out and used to guide the development of TDD reduction techniques. Experimentally, this type of information can be extracted from images of interfacial MD networks by looking for unpaired MDs [which indicate the locations of TDs, as shown in Fig. 6(e)].

**Geometrically necessary TDs from coalescence: Δz errors**

Any closure errors in Δx and Δy are readily visualized when a Burgers circuit is viewed from above. However, small Δz errors between adjacent epiwings/islands can also accumulate to create a Burgers circuit error encircling a threading dislocation. [For Δz closure errors, the lattice will spiral upwards or downwards around the enclosed TD, as is illustrated later in Figs. 9(c) and 9(d).]

Figure 8(a) illustrates some potential sources of Δz misalignment for cSAG. Perhaps the most obvious is bending of the epiwings (Δz\text{bend}), which could be due to temperature gradients, or to
FIG. 8. (a) Examples of potential \( \Delta z \) errors for cSAG growth, which can accumulate to create a \( \Delta z \) closure error in a Burgers circuit traveling through coalescing islands, creating a TD (as shown in Fig. 9). (b) Without the mask/spacer layer, adjacent islands remain in contact with the substrate, so coalescence occurs without \( \Delta z \) errors. The white layer in (a) represents a mask layer; in both (a) and (b), the black and gray layers represent atomic layers of the substrate and LMM epilayer, respectively. Dashed circles indicate regions where the epilayer is tilted or rotated with respect to the substrate.

Forces between the laterally growing epilayer and the spacer layer such as those caused by thermal expansion differences or minute variations in the thickness of the spacer layer. On an offcut substrate, a \( \Delta z \) offset can result from nucleation of adjacent epipillars on different substrate terraces, combined with vertical lattice mismatch of the two materials. Similar \( \Delta z \) offset errors could be caused by substrate roughness or differences in substrate etching depth for different SAG vias (if substrates are etched prior to epitaxy).

Another subtlety of LMM epitaxy on offcut substrates is the “Nagai tilt,” where the epitaxial layer tilts to compensate for differences in the vertical lattice constant of the epitaxial layer and substrate. If the surface step spacing beneath adjacent epipillars is different, the non-uniform Nagai tilt can produce the “\( \Delta z_{Nagai} \)” shown in Fig. 8(a). [Another source of the epilayer tilt (i.e., epitaxial layer rotation) which can produce a “\( \Delta z_{tilt} \)” similar to \( \Delta z_{Nagai} \) is asymmetric relaxation for tilted glide planes.] Taken to an extreme, such angular misalignments could create low-angle grain boundaries, but this is outside the scope of this paper.

Understanding and control of the \( \Delta z \) errors in Fig. 8(a) is important because \( \Delta z \) errors can become an increasingly important source of TDs in LMM cSAG films as the distance between epipillars is increased (and any such TDs would increase the TDD beyond that created by \( \Delta x \) and \( \Delta y \) alignment errors).

During Volmer-Weber growth [Fig. 8(b)], many of these same epilayer tilts occur, but the contiguous epilayer-substrate interface maintains the layer-to-layer registry in the vertical direction, preventing \( \Delta z \) errors and associated TDs. This can inform the choice of cSAG mask geometry for lattice-matched applications, described in the next Section (titled Coalescence in a lattice-matched system).

Coalescence in a lattice-matched system

For LM cSAG, there is no lattice mismatch to drive \( \Delta x \) or \( \Delta y \) misalignments but TDs related to \( \Delta z \) misalignments can occur during the lateral overgrowth of the epiwings [as in Fig. 8(a)]. Figure 9 illustrates schematically how this can be affected by the mask topology.

Figures 9(a)–9(d) correspond to the cSAG geometry considered throughout all of the previous figures in this paper—namely, epipillars growing up through an array of disconnected vias in a mask and then coalescing into a continuous epilayer. Although lattice-matching will ideally prevent \( \Delta x \) and \( \Delta y \) errors, the \( \Delta z \) errors from Fig. 8(a) can accumulate in this cSAG geometry to cause TDs in the resulting epilayer as the epiwings coalesce. Nothing about this geometry suppresses the \( \Delta z \) errors which create the closure error and associated TD in Figs. 9(c) and 9(d).

Figures 9(e)–9(h) illustrates a different mask geometry which could, in principle, avoid TDs related to \( \Delta z \) errors. This geometry inverts the mask geometry into an array of disconnected mask pillars, each of which is completely surrounded by exposed substrate surface prior to growth. If the initial contact between all epilayer islands occurs where the islands are in direct, contiguous contact
FIG. 9. Two examples illustrating how the mask topology could affect TD formation during lattice matched (LM) cSAG, e.g., GaAs epitaxy on a GaAs substrate. In [(a)–(d)], the epilayer grows upwards through an array of disconnected vias in the mask and then grows laterally above the mask material (lateral epitaxial overgrowth) before coalescing. Any of the $\Delta z$ errors in Fig. 8(a) can accumulate to create a geometrically necessary TD, as shown in (c) and (d). In [(e)–(h)], the mask topology has been inverted to become an array of disconnected mask pillars. Coalescence of LM islands occurs contiguously in contact with the substrate around each mask pillar [as in Fig. 8(b)] such that a Burgers circuit encircling a masked region will contain no closure errors. Therefore, no geometrically necessary dislocations are required during subsequent coalescence as the epilayer buries the mask pillar, as shown in (g) and (h). These examples also apply equally to the coalescence of lattice-matched crystalline materials on porous substrates.

with the substrate [Fig. 9(f)], then there will be no $\Delta z$ errors at the points of initial coalescence, and a Burgers circuit such as the one in Fig. 9(g) will not encircle a dislocation. Once this occurs, subsequent epilayer growth is topologically able to coalesce laterally over the mask material with no threading dislocations or other defects. Of course, any nucleation of defective material on the mask surface can act as a source of defective material in subsequently grown material, so suppression of nucleation on the mask is likely an essential requirement for good lattice-matched epitaxy for any cSAG geometry. For this reason, different results may be expected for different growth methods and conditions (OMVPE versus molecular beam epitaxy, for example).

Interestingly, Fig. 9 can be applied equally to coalescence of a porous crystalline material by simply removing the mask material from Figs. 9(b) and 9(f) to create voids. When this is done, the implication is that a porous crystalline surface created by disjoint holes [analogous to Fig. 9(f)] can coalesce more easily into a defect-free overlayer than a porous crystalline surface consisting of disjoint crystalline pillars [analogous to Fig. 9(b)].
GENERAL APPLICABILITY

To reinforce the main points of this paper, Fig. 10 shows how virtual dislocations can be applied to several other common growth geometries, besides cSAG. The diagrams on the left illustrate forbidden configurations in which dislocations terminate (unphysically) at interior surfaces, without ultimately reaching exterior surfaces. The diagrams on the right are corrected to illustrate allowed dislocation configurations, using VDs as a visualization tool.

These diagrams make clear, for example, that dislocations do not terminate at a “dislocation blocking layer”—a common tool used in a variety of applications. Instead, dislocation blocking layers lower the TDD by either facilitating the annihilation of TD segments [Fig. 10(d)] or (if and only if they are close enough to an edge) guiding them to the epilayer edge. And, as has been discussed at length in this paper, TDs which “terminate” at sidewalls during SAG are not necessarily eliminated from the overlying epilayer formed during cSAG [Fig. 10(f)]. It is therefore critical to understand the physical processes underlying the terms “dislocation blocking” and “dislocation termination”; otherwise they can become very misleading.

These diagrams emphasize another point which was made earlier in the discussion of Fig. 4, namely, that TDD reduction strategies that work by annihilation of TD segments are likely to be far more effective at high TDDs than at low TDDs. When fully understood, cSAG strategies can be used to very deliberately steer the remaining TDs in a LMM film to specific locations, but they cannot, in and of themselves, fully eliminate TDs.

Finally, it is interesting to notice that all of the growth geometries in Fig. 10 are identical from the perspective of dislocation topology. To create (d), the voids in (b) can be filled with the “dislocation blocking” material, then combined, and flattened to extend to the edges of the epilayer. To create (f), the voids in (b) can be stretched downwards to contact the substrate. For (h), the voids in (b) can be stretched upwards to reach the surface and downwards to reach the substrate and then collapsed during coalescence. In all cases, the use of VDs ensures that forbidden dislocation topologies are avoided.

While the discussion in this paper focused on planar substrate topologies for conceptual simplicity, other growth configurations may be useful for achieving high quality LMM epitaxy for different...
applications. For instance, V-grooved substrates\textsuperscript{14,30,33,67–70} and confined epitaxial lateral overgrowth (CELO)\textsuperscript{71} have been shown to be beneficial for defect control. Template-assisted selective epitaxy (TASE)\textsuperscript{72,73} can also achieve good quality nanoscale heteroepitaxy with a variety of geometries (nanowires, cross junctions, and 3D stacked nanowires). The concepts discussed in this paper are generally applicable and can be extended to help visualize the formation of coalescence-related TDs in these or any other LMM system.

CONCLUSIONS

Selective area growth is widely used to create localized volumes of dislocation-free material through aspect ratio trapping for small-area devices on lattice-mismatched substrates. For large-area devices, these small defect-free volumes would ideally be coalesced into a large defect-free epilayer. Unfortunately, for LMM growth, the resulting coalesced SAG epilayer will contain geometrically necessary threading dislocations which degrade the material quality of the cSAG epilayer, even if the epiwings of the starting uncoalesced islands are dislocation-free.

Motivated by persistent confusion in the literature about the existence and origin of these geometrically necessary TDs, we have illustrated their causes and consequences by carefully examining a series of representative examples. In each case, “virtual dislocations” can be used as a visualization tool to enforce the rules governing dislocation topology during coalescence, and in so doing, understand the source of any geometrically necessary TDs for a given configuration.

A careful comparison of LMM cSAG and LMM Volmer-Weber growth shows that the origin of geometrically necessary TDs is essentially identical for both cases, but with a few subtle differences. One difference relates to dislocation combination at voids during cSAG coalescence, which may help cSAG (or related approaches) remove TDs via dislocation combination, but only when the TDD is large. When the TDD becomes smaller, voids cannot be used to remove isolated TDs. Another difference is that, for some cSAG mask geometries, $\Delta z$ alignment errors can create additional TDs.

Finally, the use of virtual dislocations helps avoid common misinterpretations of cross-sectional TEM data; more accurate TD densities can be obtained using multiple quantitatively rigorous planar-view techniques. Although the included examples were chosen to facilitate the understanding of geometrically necessary dislocations for LMM cSAG, the visualization methods illustrated here can be applied to a variety of situations, and a few examples of this were discussed.

ACKNOWLEDGMENTS

We are grateful for helpful discussions with Ryan France, Andrew Norman, and Emily Makoutz and a critical reading by Minjoo Larry Lee. M.V. was supported by a NASA Space Technology Research Fellowship. This work was authored by Alliance for Sustainable Energy, LLC, and the manager and operator of the National Renewable Energy Laboratory for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding was provided by the DOE Office of Energy Efficiency and Renewable Energy (EERE) and Solar Energy Technologies Office (SETO) through Contract No. DE-EE-00028394. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work or allow others to do so, for U.S. Government purposes.

\textsuperscript{1} E. A. Fitzgerald, Mater. Sci. Rep. 7(3), 87 (1991).
\textsuperscript{2} E. A. Fitzgerald, A. Y. Kim, M. T. Currie, T. A. Langdo, G. Taraschi, and M. T. Bulsara, Mater. Sci. Eng. B 67(1), 53 (1999).
\textsuperscript{3} M. Yamaguchi, J. Mater. Res. 6, 376 (1991).
\textsuperscript{4} T. Ward, A. M. Sánchez, M. Tang, J. Wu, H. Liu, D. J. Dunstan, and R. Beanland, J. Appl. Phys. 116(6), 063508 (2014).
\textsuperscript{5} T. Riedl and J. K. N. Lindner, in Nanoscaled Films and Layers (Intech, 2017).
\textsuperscript{6} Q. Li and K. M. Lau, Prog. Cryst. Growth Charact. Mater. 63(4), 105 (2017).
\textsuperscript{7} S. Luryi and E. Suhir, Appl. Phys. Lett. 49(3), 140 (1986).
\textsuperscript{8} S. Bietti, A. Scaccabarozzi, C. Frigeri, M. Bollani, E. Bonera, C. V. Falah, H. von Känel, L. Miglio, and S. Sanguinetti, Appl. Phys. Lett. 103, 262106 (2013).
55 L. A. Zepeda-Ruiz, D. Maroudas, and W. H. Weinberg, Surf. Sci. 418(2), L68 (1998).
56 L. Zhang, W. E. McMahon, Y. Liu, Y. Cai, M. H. Xie, N. Wang, and S. B. Zhang, Surf. Sci. 606(21-22), 1728 (2012).
57 S. H. Huang, G. Balakrishnan, A. Khoshakhlagh, A. Jallipalli, L. R. Dawson, and D. L. Huffaker, Appl. Phys. Lett. 88(13), 131911 (2006).
58 S. Huang, G. Balakrishnan, and D. L. Huffaker, J. Appl. Phys. 105(10), 103104 (2009).
59 P. Sheldon, K. M. Jones, M. M. Al-Jassim, and B. G. Yacobi, J. Appl. Phys. 63(11), 5609 (1988).
60 W. E. McMahon, I. G. Batyrev, J. M. Olson, and S. B. Zhang, Phys. Rev. Lett. 89(7), 076103 (2002).
61 M. Xiao, J. Zhang, X. Duan, H. Shan, T. Yu, J. Ning, and Y. Hao, Sci. Rep. 6, 23842 (2016).
62 H. Nagai, J. Appl. Phys. 45(9), 3789 (1974).
63 J. E. Ayers, S. K. Ghandhi, and J. L. Schowalter, J. Cryst. Growth 113(3-4), 430 (1991).
64 I. Mizushima, T. Sato, S. Taniguchi, and Y. Tsunashima, Appl. Phys. Lett. 77(20), 3290 (2000).
65 A. Boucherif, G. Beaudin, V. Aimez, and R. Arès, Appl. Phys. Lett. 102(1), 011915 (2013).
66 R. Hull, J. C. Bean, G. S. Higashi, M. L. Green, L. Peticolas, D. Bahnck, and D. Brasen, Appl. Phys. Lett. 60(12), 1468 (1992).
67 W. Guo, L. Date, V. Pena, X. Bao, C. Merckling, N. Waldron, N. Collaert, M. Caymax, E. Sanchez, E. Vancoille, K. Barla, A. Thean, P. Eyben, and W. Wandervorst, Appl. Phys. Lett. 105, 062101 (2014).
68 T. Orzali, A. Vert, B. O’Brien, J. L. Herman, S. Vivekanand, R. J. W. Hill, Z. Karim, and S. S. Papa Rao, J. Appl. Phys. 118, 105307 (2015).
69 M. Paladugu, C. Merckling, R. Loo, O. Richard, H. Bender, J. Dekoster, W. Vandervorst, M. Caymax, and M. Heyns, Cryst. Growth Des. 12, 4696 (2012).
70 A. Krost, R. F. Schnabel, F. Heinrichsdorff, U. Rossow, D. Bimberg, and H. Cerva, J. Cryst. Growth 145, 314 (1994).
71 L. Czornomaz, E. Uccelli, M. Sousa, V. Deshpande, V. Djara, D. Caimi, M. D. Rossell, R. Erni, and J. Pompeyrine, in 2015 Symposium on VLSI Technology (IEEE, 2015), p. T172.
72 P. D. Kanungo, H. Schmid, M. T. Bjork, L. M. Gignac, C. Breslin, J. Bruley, C. D. Bessire, and H. Riel, Nanotechnology 24(22), 225304 (2013).
73 H. Schmid, M. Borg, K. Moselund, L. Gignac, C. M. Breslin, J. Bruley, D. Cutaia, and H. Riel, Appl. Phys. Lett. 106(23), 233101 (2015).