ORDER AND DISORDER IN TYPE–II InAs/InAsSb SUPERLATTICES

A Dissertation

by

MATTHEW ROBERT WOOD

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Michael B. Weimer
Committee Members,	Alexey Belyanin
	Christi K. Madsen
	Joseph H. Ross Jr.
Head of Department,	Peter McIntyre

December 2017

Major Subject: Applied Physics

ABSTRACT

We employ cross–sectional scanning tunneling microscopy (STM) to examine how an *as–grown* InAs/InAsSb superlattice differs from the intended one as regards translational invariance in (001) planes perpendicular to the growth direction. This requires atomic–resolution, lateral surveys paralleling the buffer/epilayer interface for up to a micron in orthogonal (–1–10) or (1–10) cross sections, together with repeated lateral surveys at representative vertical locations (i.e., spanned superlattice repeats) within the multilayer stack.

We show that STM may be used to accurately map the period fluctuations throughout this superlattice. The concept, analogous to Bragg's law in high–resolution x–ray diffraction, relies on an analysis of the [001]–convolved reciprocal–space satellite peaks obtained from discrete Fourier transforms of individual STM images. Properly implemented, the technique enables local period measurements that reliably discriminate lateral fluctuations localized to within ~ 40 nm along <110> directions in the growth plane. While not as accurate as x–ray, the inherent, single–image measurement error associated with the method may be made as small as 0.1%, allowing the lateral period fluctuations contributing to inhomogeneous energy broadening and carrier localization in these structures to be pinpointed and quantified. The direct visualization of unexpectedly–large fluctuations on nanometer length scales is tied to a stochastic description of correlated interface roughness.

We also introduce a new technique to automatically tabulate the crystalline coordinates of previously–identified top–layer antimony atoms and construct the antimony pair–correlation functions for orthogonal cross sections. Nearest–neighbor correlations on opposing cleavage faces are inversely related, with the (–1–10) deficit at nearest–neighbor sites balanced by a compensating (1–10) surplus. The logarithm of this preference scales inversely with bulk antimony fraction. In more vivid physical terms, the preferential [110]–incorporation of nearest–neighbor antimony atoms in the bulk is traced to the inferred concentration of [110]–oriented antimony dimers at the growth surface.

to my amazing wife and our charismatic daughter

ACKNOWLEDGEMENTS

I owe much of my success in graduate school to my advisor Dr. Michael Weimer. In addition to always knowing the right set of questions to ask to get to the bottom of problems he always knew the perfect blend of when to encourage me to follow my instincts and when to bring my focus back to the bigger picture. If I have absorbed even 1% of his knowledge I'll consider myself very lucky.

Of course this lab was built on the work of previous students. Drs. George Lengel and Geoff Brown built the STM and the vacuum system in which it resides. Drs. John Harper and Jeremy Steinshnider developed the lab antimonide program into what it is today. Dr. Federico Lopez upgraded the vacuum chamber to allow for multi–week experiments, and altered best practices on how to acquire data. Dr. Kara Kanedy developed a quick method of reconstructing bulk profiles and acquired many of the datasets I am using in these analyses. I am additionally indebted to Federico and Kara for all the help and support they offered during our overlapping time in College Station.

Finally, I wish to thank family near and far who believed in me and encouraged me throughout my journey with a very special thank you to my wife whose patience was tested by the length of my graduate career, but who was loving and supportive nonetheless.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was advised by Dr. Michael Weimer and supported by a dissertation committee consisting of Dr. Alexey Belyanin and Dr. Joseph Ross Jr. of the Department of Physics and Astronomy and Dr. Christi Madsen of the Department of Electrical and Computer Engineering.

A portion of the data analyzed in Chapters III and IV was collected by Drs. Federico Lopez and Kara Kanedy. Additionally the analysis depicted in the Chapter III section titled “Analysis of Experimental Data” was conducted in part by Dr. Kanedy, and the analysis depicted in the Chapter III section titled “Direct Measurement of Interface Roughness in Quantum Cascade Materials” was conducted in part by Dr. Lopez. All other work conducted for this dissertation was completed independently.

Samples used in these studies were provided by Dr. John Klem of Sandia National Laboratories, Dr. Catherine Caneau of Corning Inc., and Dr. Stefan Svensson of the Army Research Labs.

Funding Sources

Graduate study was supported by a merit fellowship and the C.F. Squire fellowship from Texas A&M University as well as by grants from NSF (MIRTHE NSF–ERC Cooperative Agreement EEC-0540832), Sandia National Laboratories, and ARO (STIR W911NF-14-1-0645, MURI W911NF-10-1-0524).

TABLE OF CONTENTS

CHAPTER Page

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Technological Motivation

Compact, efficient lasers and detectors that operate in the mid– to long–wave infrared regions of the electromagnetic spectrum have potential applications ranging from thermal imaging to molecular identification. This regime covers two transparent atmospheric windows [1], facilitating military and security technologies, which rely upon long–range detection of heat from human bodies in the darkness of night[1]. Furthermore, many environmentally [2,3] and biologically [2,4,5] important molecules vibrate with frequencies in the infrared, making trace–gas sensing feasible for non–invasive medical diagnostics as well as air–quality or climate–change monitoring.

Mercury Cadmium Telluride (MCT) has historically been the detector material of choice at these wavelengths [6], due in part to the high level of tunability (1–30 μm) with changing composition, as well as favorable electrical properties, such as high mobility and low dielectric constant. However work to replace this material has received considerable attention for two principle reasons: the bandgap (and therefore the cutoff wavelength) is highly sensitive to uniformity of the alloy composition and a weak Hg–Te bond results in structural instabilities during epitaxial growth.

A number of compound semiconductors involving group–III and group–V elements are viable candidates to replace MCT over different portions of the infrared

---

[1] The dominant wavelength in blackbody radiation from a human body is around 10 μm.

spectrum (Fig. 1.1, [7]); it was recognized some time back that type–II superlattices (Fig. 1.2) in particular might hold distinct advantages over MCT [8] since the superlattice energy gap [9] might be made smaller than that in either constituent material[2].

Despite an early proposal for a simpler InAs / InAsSb superlattice [10], most research over the past 20 years has been devoted to the InAs / GaSb system. Recent demonstrations that the Ga–free superlattices exhibit minority carrier lifetimes up to two–orders of magnitude greater [11,12] than their Ga–containing counterparts [13,14] has renewed interest in the prospects for competitive devices based on InAs / InAsSb. In addition reports of detectivities in InAs / InAsSb comparable with those of MCT [15] have further elevated the hopes for this material system to become the cornerstone of next–generation infrared detectors.

Rapid, industrial–scale production of semiconductor devices typically relies on Metal–Organic Chemical Vapor Deposition (MOCVD), however in research settings the preferred growth method tends to be Molecular Beam Epitaxy (MBE) where an ultra–high vacuum (UHV) environment (Fig. 1.3, [16]) and slow growth rates routinely result in better quality epitaxial layers.

The low background pressures (typically 1E–8 torr or better) employed in MBE naturally limit the concentration of residual gas atoms that can potentially contaminate epitaxial layers over the many hours it takes to grow a device. Just as importantly, the mean–free path of atoms at these pressures is enormous (~ 5 km) [17]; so elemental sources, once vaporized in effusion cells attached to the vacuum system (Fig. 1.3), form

---

[2] This allows detection of longer wavelengths than otherwise possible with a bulk alloy, but comes at the expense of spatially–indirect electronic transitions.

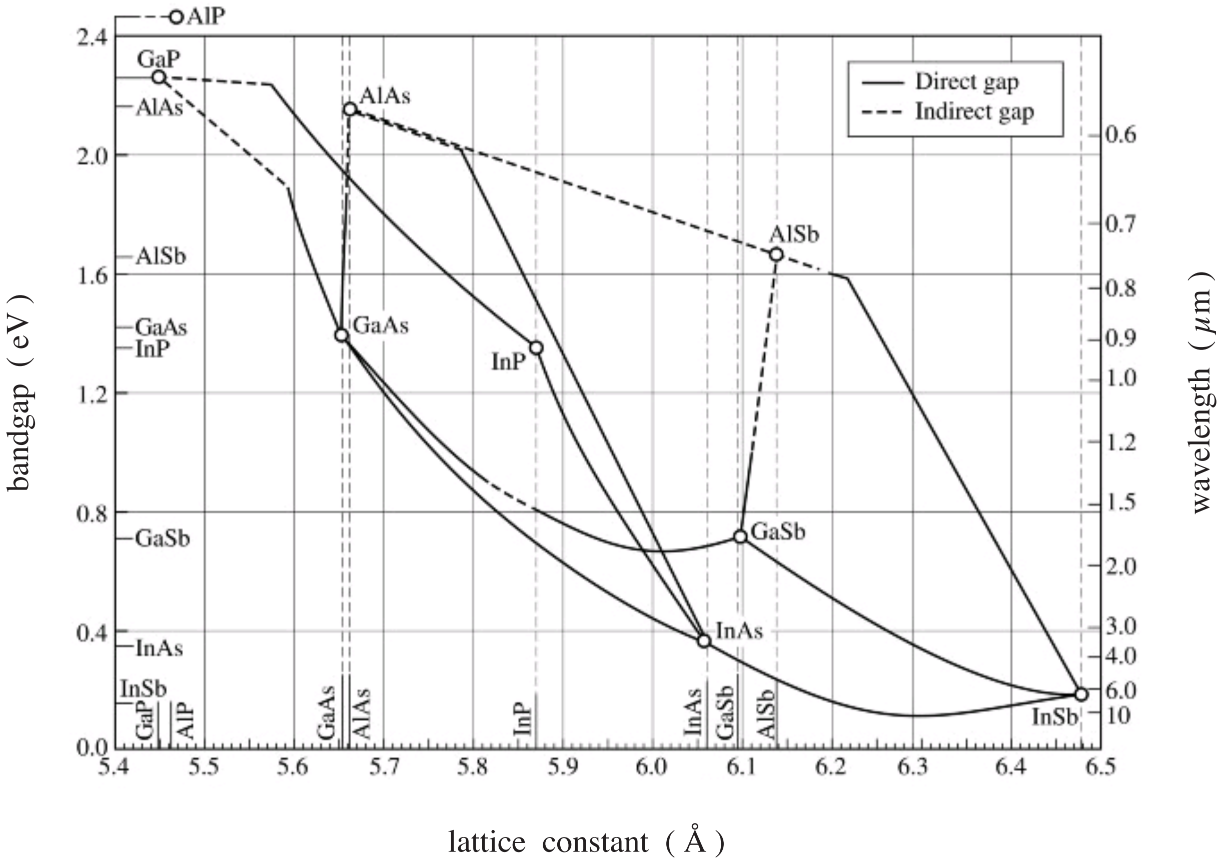


FIGURE 1.1. Bandgap versus atomic spacing for common III / V semiconductor alloys. By appropriate selection of the alloy fraction and layer thicknesses, superlattices fabricated from InAs, InSb and GaSb can function as infrared detectors for wavelengths anywhere between 5–10μm. Reprinted from the internet [7].

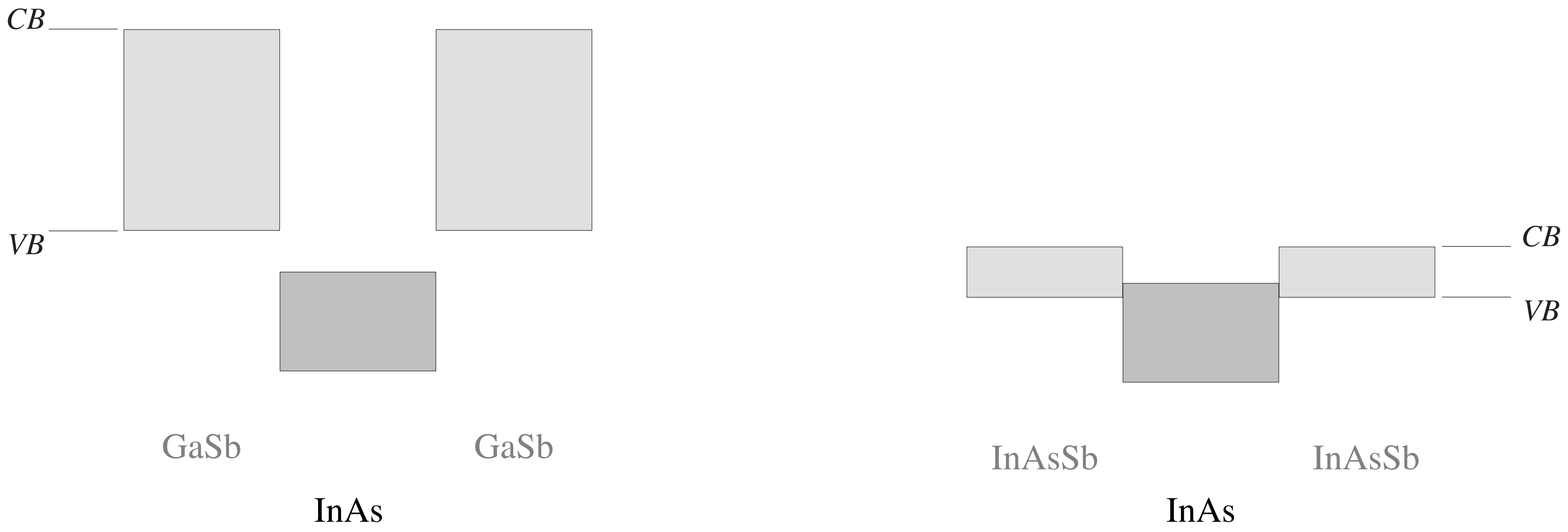


FIGURE 1.2. Representative energy–band diagrams for type–II band alignments that respectively confine electrons and holes in neighboring superlattice layers. The optical transition energy is tailored by adjusting either the thicknesses of the respective superlattice layers or the band offsets (alloy composition) subject to the constraint that the strains in each half of the superlattice compensate one another so the total strain per period sums to zero.

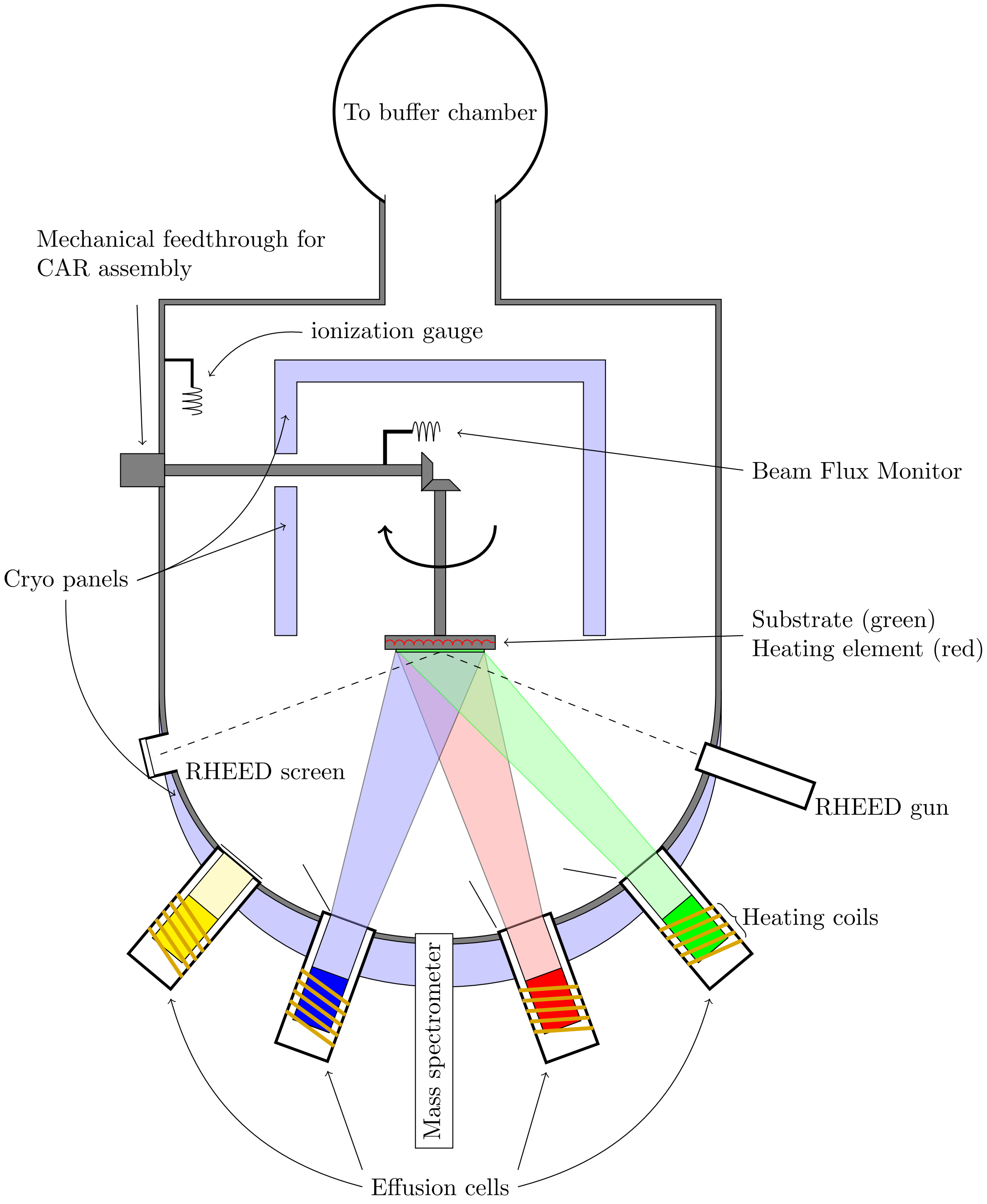


FIGURE 1.3. Schematic illustrating the major components in a typical ultra–high vacuum molecular beam epitaxy system. Reprinted from the internet [16].

beams of anions and cations which travel from their source to the (heated and continuously rotated) growth surface without experiencing collisions. Independently controlled effusion cell temperatures determine the flux from each source. Under usual growth conditions, it is the cation flux that controls the growth rate. The combined flux of anions is typically kept several times greater to promote two–dimensional growth by preventing cations from otherwise coalescing in metallic pools on the surface [18]. Once deposited, anions and cations diffuse across the growth surface until they encounter other adatoms[3] and nucleate in two–dimensional islands [19].

The elemental effusion cells in Fig. 1.3 are universally maintained at fixed temperatures throughout superlattice growth (to preserve flux stability and uniform growth rates) and a spatially–modulated structure formed by actuating shutters that alternately open and close, selectively illuminating the growth surface with one or more vapor streams. Shutter cycle timing sets the superlattice well–to–barrier ratio whereas source fluxes (and therefore growth rate) set the actual layer thicknesses, each of which bears on the resulting electron and hole energy levels in the device.

Materials Issues

Despite the abrupt change in vapor flux that follows actuation of a source shutter, physical processes occurring at the growth surface prevent a similarly abrupt change in superlattice composition (Fig. 1.4, [20]). A residual vapor background in the MBE

---

[3] Alternate models suggest that adatoms diffuse until they encounter a step, which then expands across the growth surface to form terraces, this tends to occur more often on substrates with a large vicinality where steps are closely spaced.

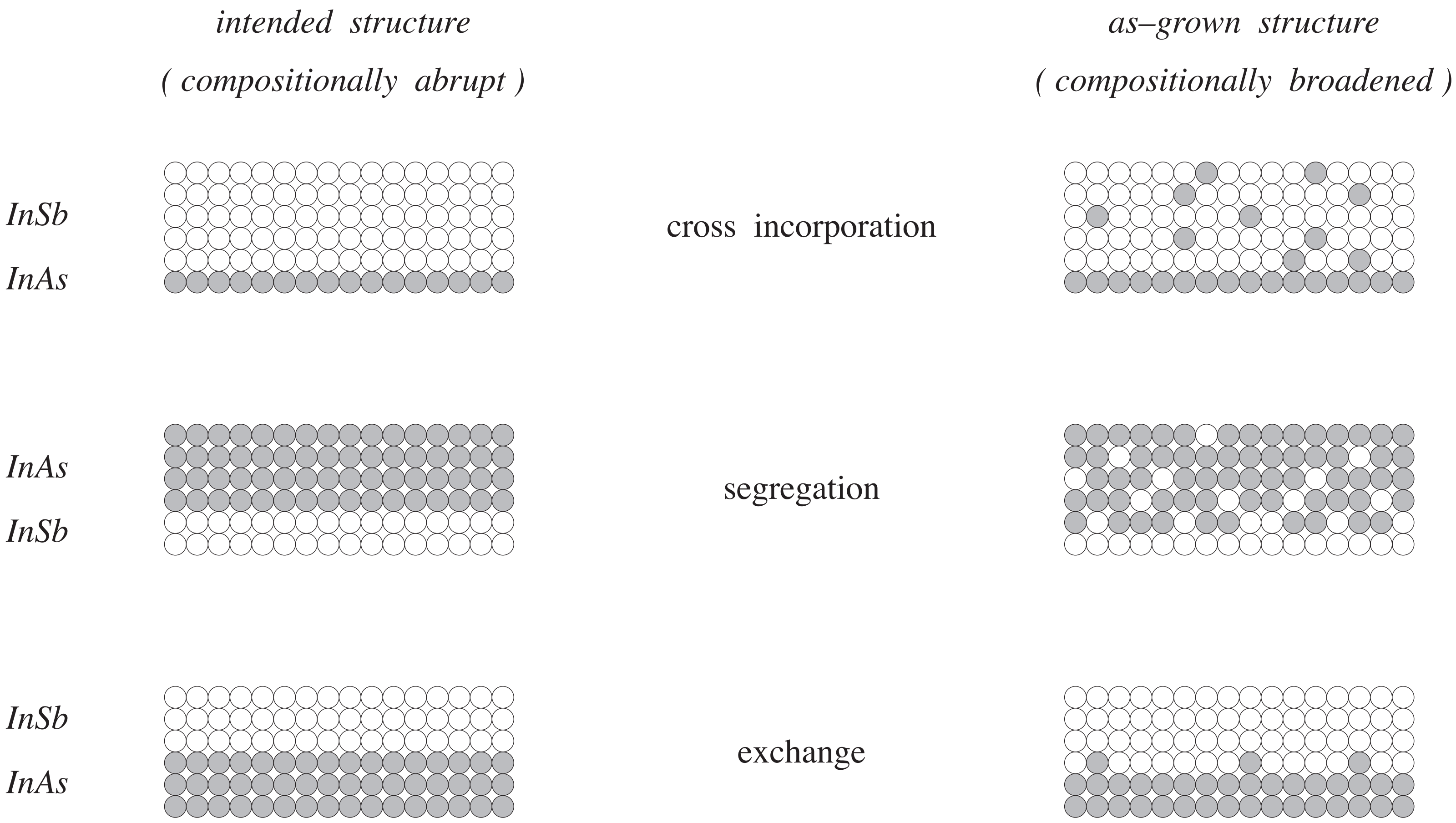


FIGURE 1.4. Schematic illustrating key physical processes that contribute to compositional grading during MBE growth. Adapted from [20].

chamber following source closure remains available to incorporate in subsequent epitaxial layers (cross incorporation) [21], potentially compromising the compositional differences between superlattice constituents. In addition, whenever an arsenide is grown atop an antimonide, as is the case in an InAs/InAsSb superlattice, antimony tends to collect on the surface (floating layer) [22,23] before a fixed fraction (segregation coefficient) is subsequently incorporated in each monolayer. This segregation leads to an exponentially–graded antimony fraction that extends well into the InAs layer. Anion exchange also generally occurs at III–V heterojunctions [21], but is especially difficult (if not impossible) to pinpoint in common–atom superlattices such as InAs/InAsSb.

We explore two further ways an *as–grown* superlattice may significantly differ from design intentions in this manuscript. Spontaneous atomic order, typically attributed to "dimer–induced subsurface stresses" when one atom in an alloy is significantly larger than the other [24], occurs in many alloyed III–V materials grown on (001) substrates[4]. Copper platinum (CuPt) order, with alternating {111} planes of arsenic and antimony atoms (Fig. 1.5) [24], has previously been observed in InAsSb [25,26]. The superimposed <111> translational symmetry in CuPt–ordered materials reduces the bandgap by an amount that depends both on the range and magnitude of the order present [27]. Indeed, efforts are now being made to deliberately engineer smaller bandgaps in bulk InAsSb alloys [28] via strain–induced long–range order; transport and band structure models for InAs/InAsSb superlattices would similarly benefit from a thorough characterization of order in these materials.

---

[4] A notable exception is AlGaAs, where aluminum and gallium are very similar in size.

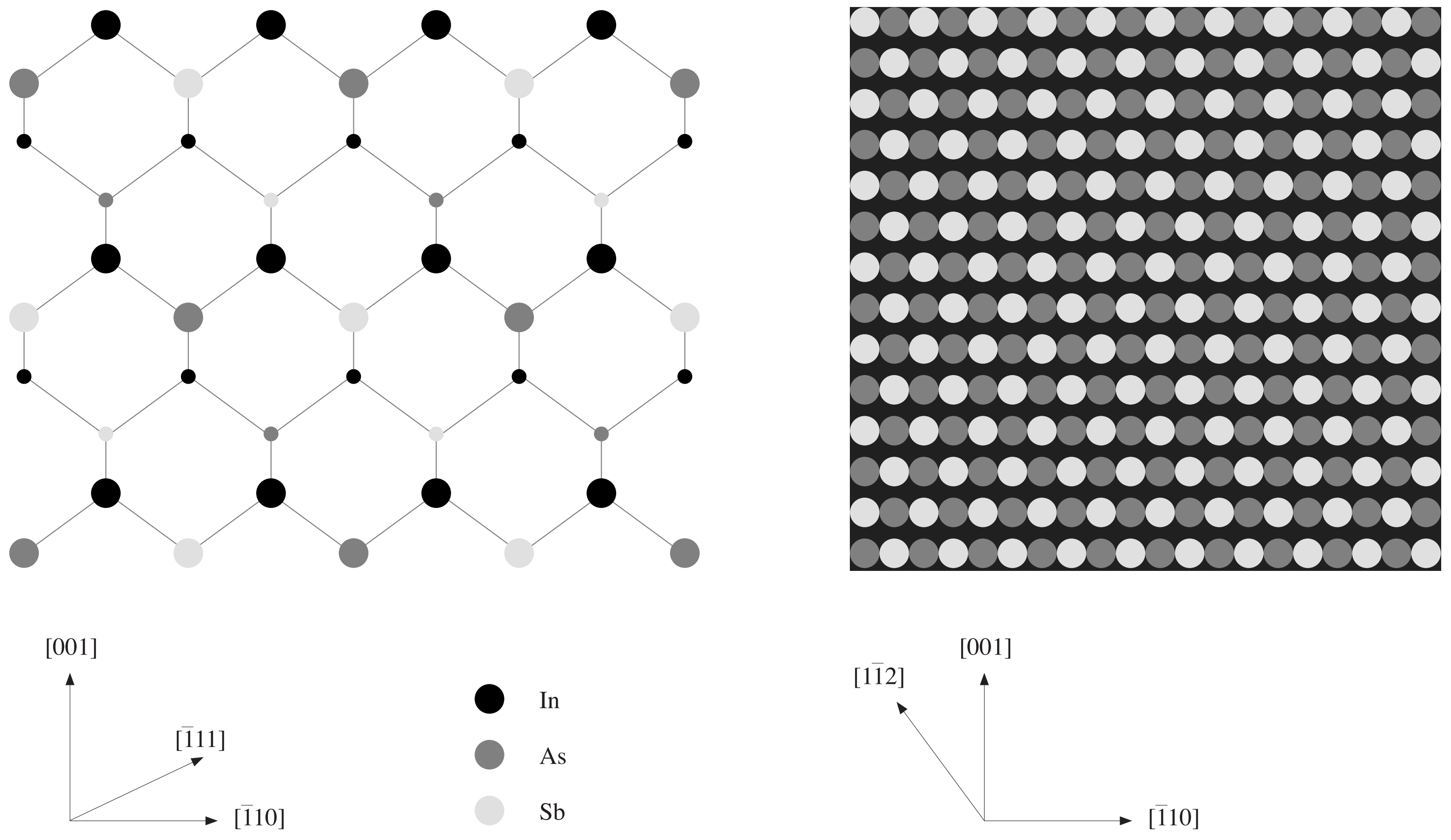


FIGURE 1.5. Ideal CuPt ordering in the bulk (left) manifests as alternating (–111) planes of arsenic and antimony atoms. The intersection of these idealy–ordered (–111) planes with a (110) surface (right) produces alternating [1–12] rows of arsenic and antimony atoms.

Period fluctuations are a poorly understood source of larger–scale structural disorder since they are difficult to appropriately characterize. Islands formed during deposition (Fig. 1.6, left) that approach complete monolayers minimally perturb a device, but those of intermediate size will result in (detectably) rough heterojunctions between superlattice constituents (Fig. 1.6, right). If this vertical roughness is normally distributed with a hypothetical standard deviation of ~ 1 ML, the periods bracketed by any pair of like heterojunctions will be within ± 1.4 ML of their mean value 70% of the time (assuming the fluctuations at successive interfaces are statistically independent), but fluctuations exceeding 1.4 ML will nevertheless occur 30% of the time.

On the other hand, correlations in the growth plane that suppress the statistical independence of fluctuations along each interface will correspondingly amplify the lateral period fluctuations between vertically independent interfaces. These lateral variations in superlattice period break translational invariance in the growth–plane for carriers with comparable wavelengths, resulting in inhomogeneous energy–level broadening. This differs from the standard picture of inhomogeneous broadening that accompanies a (laterally–uniform) vertical drift in superlattice period, which preserves lateral translation invariance. Recent work attributes the occurrence of carrier localization in InAs / InAsSb superlattices [29] to either (laterally invariant) layer–to–layer thickness fluctuations or (vertically invariant) lateral thickness fluctuations similar to those just described, but no connection between lateral period fluctuations and interface roughness is made. An explicit understanding of this connection seems

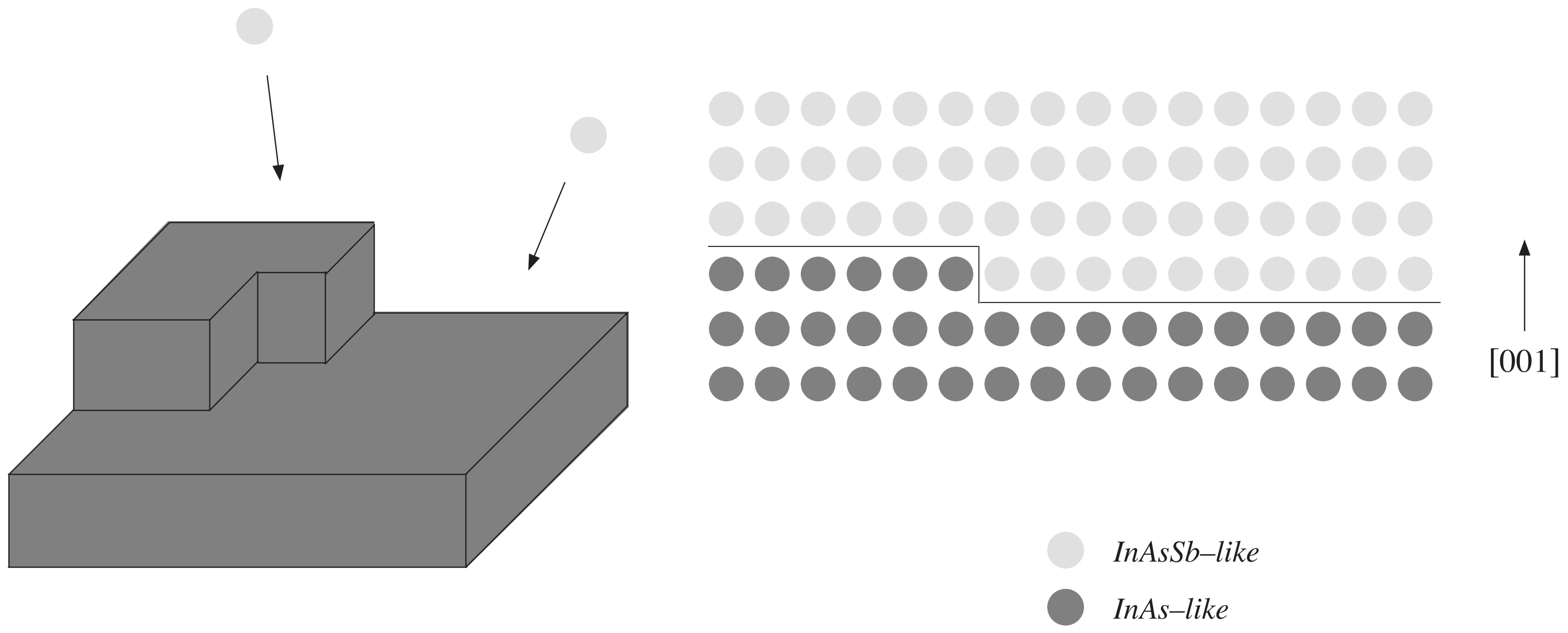


FIGURE 1.6. Incomplete InAs monolayers at the (001) growth surface typically serve as the template (left) for InAsSb growth once the antimony shutter is actuated. In the bulk crystal this roughness shows up as an abrupt interface “kink” (right) assuming other material issues (Fig. 1.4) negligibly influence the subsequent incorporation of antimony atoms.

important, both in its own right and in light of the long–recognized role that roughness plays in suppressing carrier mobility due to increased interface scattering [30].

Characterization Techniques

The engineering of MBE–grown devices is almost universally based on an assumed structural integrity and several characterization techniques are available to gauge the validity of these assumptions from the quality of *as–grown* heterostructures. *In–situ* reflection high–energy electron diffraction (RHEED) is sometimes used (as shown in Fig. 1.3) to calibrate growth rates, assess surface quality, and monitor surface reconstruction during growth. Layer stoichiometry, which is tied to source fluxes and shutter cycles, is almost always calibrated after the fact by the substrate mismatch observed by high resolution x–ray diffraction (HRXRD); in this sense *in–situ* source flux and temperature monitoring (Fig. 1.3) principally serve to establish reproducible conditions. More exotic *in–situ* probes rely on growth surface curvature, thin–film stress, or anion desorption [31,32] to assess alloy composition during the growth.

Ex–situ techniques in addition to HRXRD include transmission electron microscopy (TEM) and scanning tunneling microscopy (STM). HRXRD is the gold standard for determining superlattice periodicity and mismatch [33] over an entire multilayer stack, and at times can even yield information about vertical period fluctuations [34]. Cross–sectional TEM offers atomic–resolution information about a structure [35,36] that could in principle facilitate the measurement of lateral superlattice period fluctuations on nm length scales [37], however much like HRXRD, TEM averages over a significant sample depth, making it difficult (if not impossible) to identify order and disorder that present on short length scales; TEM has, however, been successfully used to identify vertical period fluctuations [38].

With cross–sectional STM one images and identifies individual atoms within a single plane representative of bulk III–V material. The natural cleavage of zincblende semiconductors along orthogonal (110) and (1–10) surfaces (Fig. 1.7, [39]) then enables the subsequent reconstruction of complimentary, monolayer–by–monolayer records of the growth plane's evolution across any chosen subset of superlattice repeats.

Scanning Tunneling Microscopy

STM relies on the quantum mechanical tunneling of electrons through a thin, vacuum barrier separating a freshly cleaved semiconductor surface from a sharp metal tip (Fig. 1.8). What constitutes an "atomically–sharp" tip is a matter of conjecture, but theoretical considerations suggest that ~ 1 Å lateral resolution is consistent with a single, metal *d*–shell orbital [40]. Achieving this resolution in practice demands that extreme care be taken to reduce ambient vibrations well below this number.

The tunneling current, which is fundamental to all STM measurements, is typically approximated by

$$I \propto \int_0^{eV} \rho_s(E, eV)\, \exp(-\,2\kappa s)\, dE \tag{1.1}$$

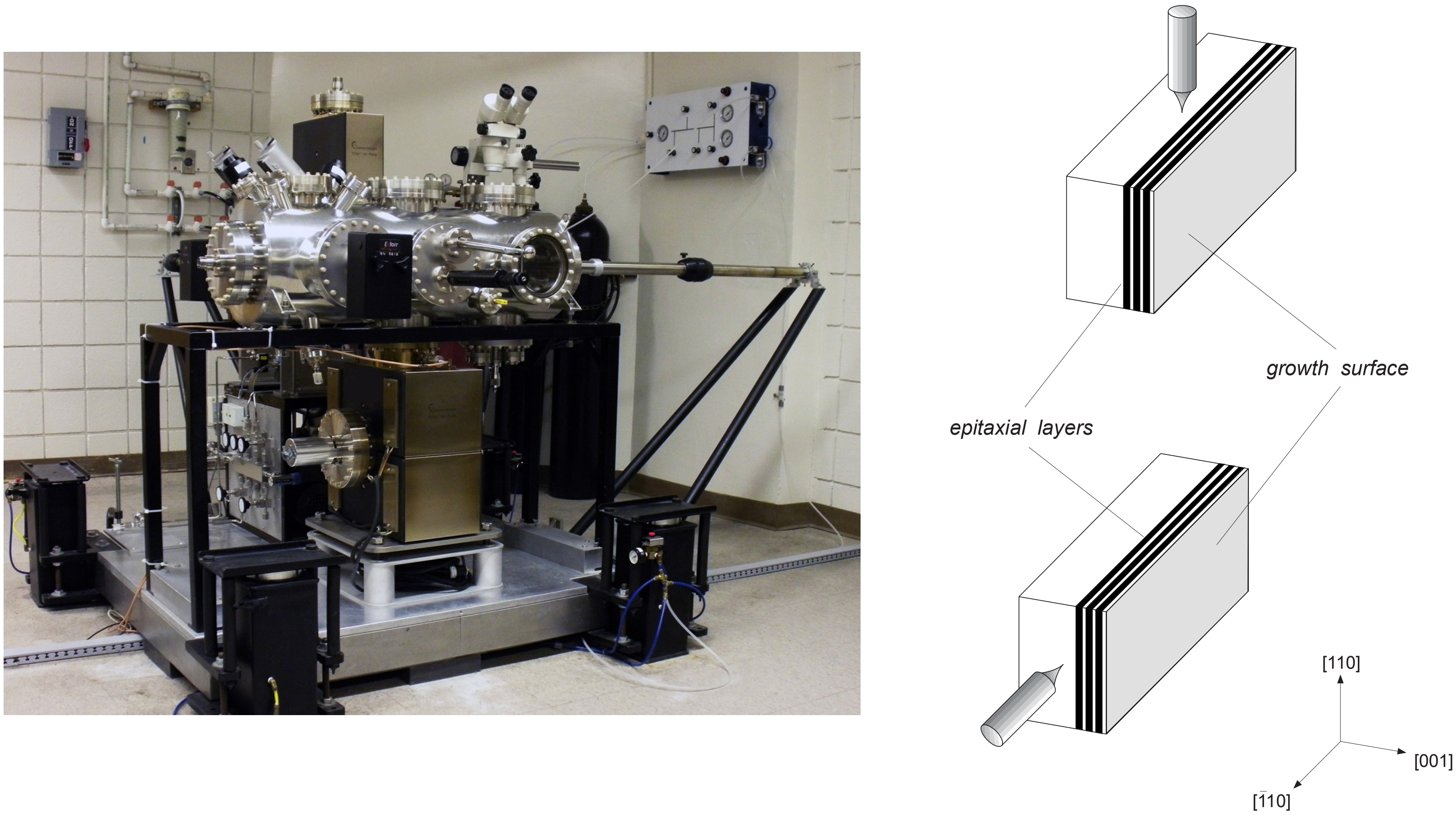


FIGURE 1.7. Dedicated UHV chamber (left) which houses the scanning tunneling microscope alongside schmatic illustrating the two orthogonal cleavage cross sections accessible with STM. Adapted from [39].

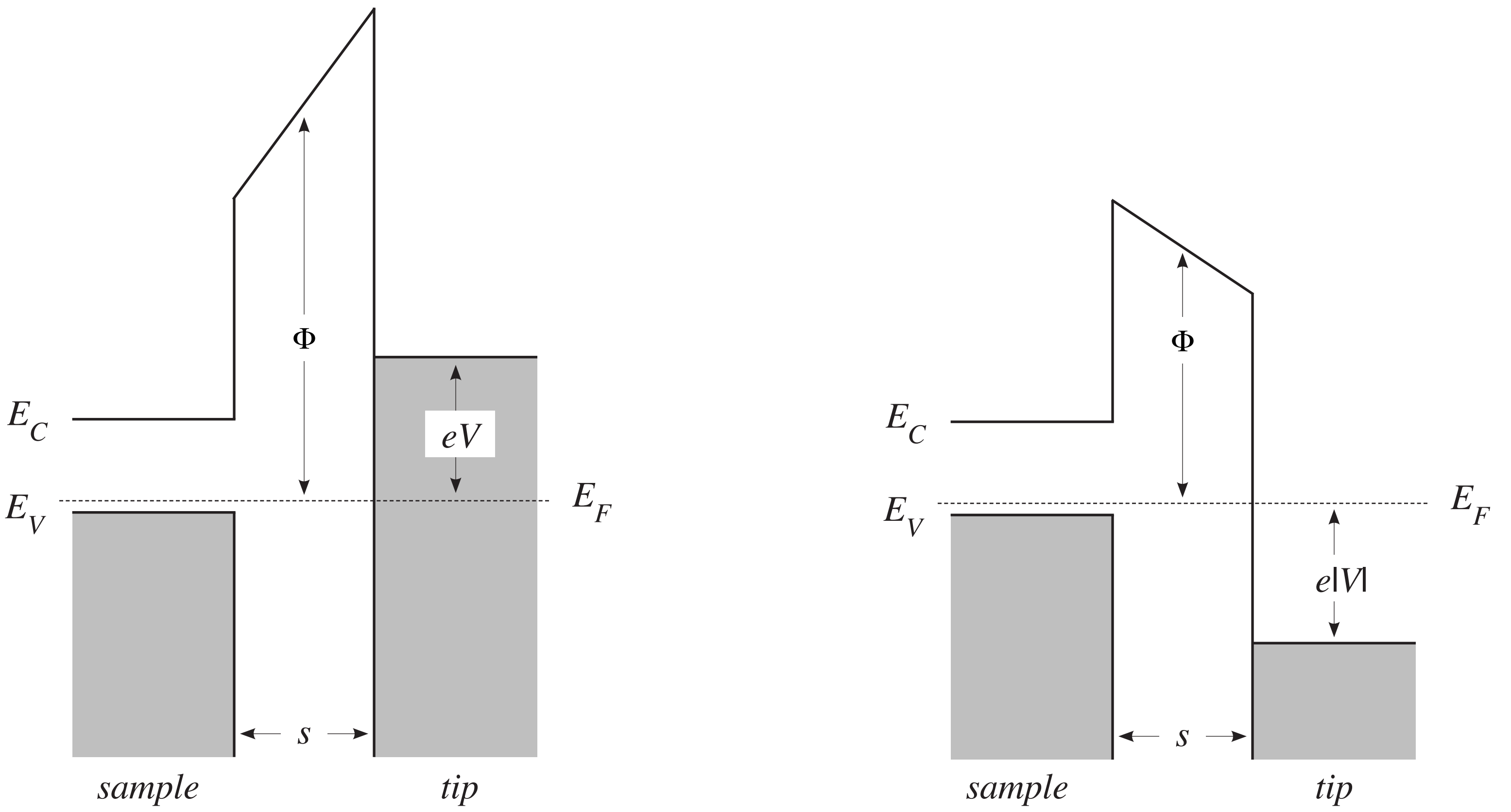


FIGURE 1.8. Schematic electron energy diagrams illustrating trapezoidal tunnel barrier separating idealized, planar tip–sample junctions for a positively– (left) and negatively–biased (right) sample. The magnitude of the elementary charge is denoted by e, the thickness of the vacuum barrier by s, and the average barrier height by Φ. Adapted from [39].

$$\kappa = \sqrt{\frac{2m}{\hbar^2}(\Phi(eV) - E)}$$

where s is the thickness of the vacuum barrier, ρ_s the energy–dependent sample density of states, and κ an exponential decay constant[5] that reflects the average barrier height, Φ, under applied sample bias eV (Fig. 1.8) [42].

The anion (filled state) and cation (empty state) sublattices at III–V <110> cleavage surfaces can be imaged separately by changing the sign of the bias voltage (negative for anions and positive for cations) [43]. Notwithstanding this atom–selective imaging as well as (1.1), there is only a small range of acceptable voltages that, in practice, produce high–quality large–area STM images. Biases that are too small risk damaging the tip by forcing it towards the surface, whereas biases that are too large result in field emission that also damages the tip, the surface, or both.

The principal cleavage directions in zincblende crystals lie in orthogonal (110) and (1–10) planes [44], each of which contains the [001] growth direction (Fig. 1.7), but only every second bulk growth plane terminates at the cleavage surface. This, at first, appears a major drawback, but is experimentally circumvented by sampling multiple surface repeats from which the bulk repeat may be reconstructed under the assumption any changes across the multilayer stack occur slowly on the scale of a superlattice period.

---

[5] The exponential decay constant κ is ~ 1 Å^{-1} for typical semiconductor work functions [41].

Unlike other low–index faces, the <110> cleavage surfaces of zincblende semiconductors do not reconstruct, but undergo a rigid–bond relaxation (Jahn Teller effect [45]) that preserves the (in–plane) translational symmetry of the bulk. This relaxation is driven by a simultaneous re–hybridization of, and (partial) charge transfer between, cation and anion dangling bonds that spontaneously lowers the energy of the system. As a result, the filled–states predominantly associated with anions are projected outward, into the vacuum, whereas the empty–states predominantly associated with cations remain more nearly planar, and thus more difficult to couple to the tip. Most images obtained in this lab, and all images contained in this dissertation, are of the filled–state anion sublattice.

The circumstances described above for a homogeneous semiconductor or lattice–matched superlattice become more complex in strain–balanced systems where the respective constituents are under compensating tensile and compressive strains. Considering each superlattice component as an anisotropic, elastic continuum [46] (whose properties follow alloy composition), the strain release attending removal of a half–space normal to the growth direction is followed by relaxation of the tensile component into, and relaxation of the compressive component out of, the remaining half–space[6]. Thus in the case of the (tensile) InAlAs / (compressive) InGaAs superlattice (considered in Chapter III) the entire tensile constituent sinks into the surface and the entire compressive constituent bulges out of the surface to produce a surface topography that closely tracks the well–to–barrier division in the superlattice.

---

[6] [001] relaxation is also possible in this continuum model.

The continuum model described above is directly applicable to cleavage–induced strain relaxation in compositionally abrupt systems, but this is clearly not the case for InAs/InAsSb, which is compositionally graded. The contrast in these superlattice images are almost entirely dominated by the strain relaxation of individual InSb bonds [47].

Since the tunneling current in (1.1) does not explicitly depend on elemental composition, an individual atom can only be identified when all 118 possibilities comprising the periodic table can be condensed down to those handful of elemental sources attached to the MBE chamber (Fig. 1.3). These few possibilities can then often be distinguished by the differences in natural bond length that become manifest following post–cleavage, site–specific strain relaxation.

For example, there are only two possible anion–cation pairings in an InAsSb alloy: InAs, whose lattice constant is 6.05Å, and InSb, whose lattice constant is 6.47Å [44]. A coherently–strained InSb bond will thus stick out further from the cleavage surface than the corresponding InAs bond, once both are free to do so. Top–layer InSb pairings are then pinpointed by the consequent increase in tunnel current as the STM tip passes over each substitutional site (Fig. 1.9).

No matter which strain relaxation mechanism dominates at the cleavage surface, the composition within this plane is representative of that in every other like plane throughout the bulk, and this is what gives cross–sectional STM it’s unique power.

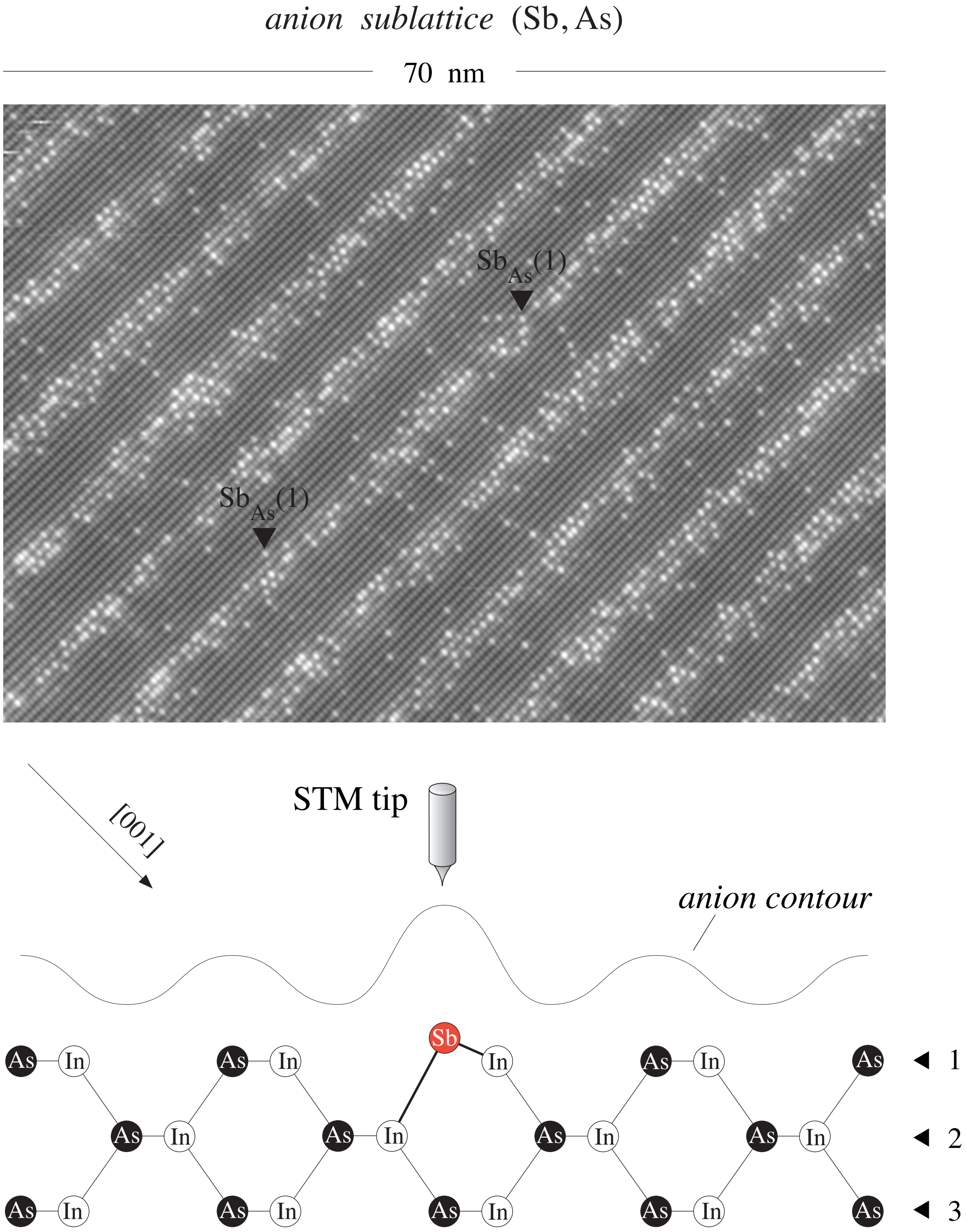


FIGURE 1.9 Atomic–resolution STM image of the anion sublattice (Sb, As) for a type–II InAs / InAsSb superlattice. Top–layer isovalent impurities are indicated by carets in the image. The lower schematic illustrates the change in current contour as the tip encounters a change in surface topography. Growth direction is from top–left to bottom–right. Adapted from [39].

Dissertation Overview

Chapter II details how atomically–flat cleavage across an InAs/InAsSb superlattice is achieved and atomic–resolution STM images acquired. Particular care is given to the discussion of STM non–idealities bearing on the accuracy and interpretation of measurements central to the scientific objective pursued here, as well as new techniques that effectively circumvent these shortcomings. Chapter III considers long–-range structural disorder, beginning with an STM measurement of the average InAs/InAsSb superlattice period; we then proceed to map the period fluctuations over nm length scales (image–by–image) throughout micron–long lateral surveys; finally, we connect these lateral fluctuations to a stochastic description of correlated interface roughness. Chapter IV refocuses on short–range atomic order. We introduce a new technique to automatically tabulate the crystalline coordinates of previously–identified top–layer antimony atoms in InAs/InAsSb and construct the antimony pair–correlation functions in both (–1–10) and (1–10) cross sections. These correlation functions are then analyzed in terms of layer strain as well as antimony fraction, and comparisons drawn between corresponding superlattice and bulk alloy experiments.

CHAPTER II

EXPERIMENTAL DETAILS

Introduction

Our test bed for an in–depth structural study of these III–V materials is a 33% antimony InAs / InAsSb strain–layer superlattice (Fig. 2.1) – fabricated via MBE at Sandia National Laboratories [23]. This growth nominally targeted 100 repeats of 15.4 ML InAs / 5.4 ML InAsSb atop a *p*–type GaSb buffer grown on one–quarter of a 2" *n*–type GaSb substrate[1] for an intended superlattice period of 20.8 ML.

The *as–grown* structure has been shown [23] to differ measurably from the intended one in three ways, only two of which are consequential. The superlattice period measured by HRXRD is 0.2 ML shorter than design – indicating a comparatively minor ~ 1% growth–rate miscalibration – but the structure also displays antimony segregation as well as antimony cross–incorporation. The succeeding chapters describe studies performed with a cross–sectional scanning tunneling microscope (STM) that illustrate two additional ways this *as–grown* structure differs significantly from engineering design intentions. Here, we describe the scientific equipment and preparation necessary to obtain device–scale, atomic resolution surveys, as well as potential pitfalls related to the STM scanning mechanism, that must be overcome to analyze these images.

---

[1] There are two standards for semiconductor substrates. The GaSb substrate on which this sample was grown uses the European–Japanese (EJ) standard. The United States (US) standard differs in the orientation of the major and minor flats.

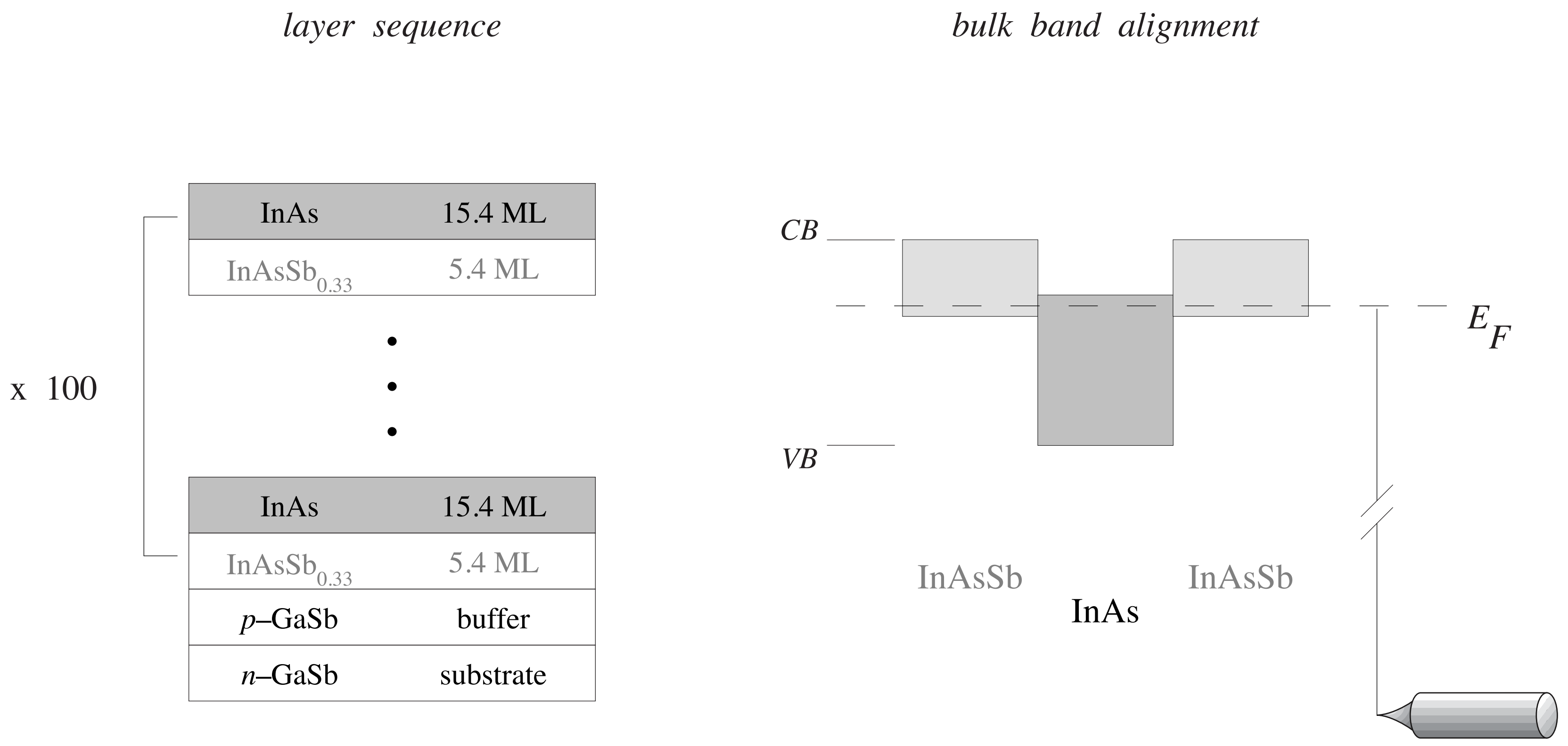


FIGURE 2.1. The sample used in this study consists of 100 repeats of a nominally 5.4 ML InAsSb$_{0.33}$ / 15.4 ML InAs superlattice grown on a GaSb buffer and substrate (left). This composition results in a staggered “type–II” bulk band alignment (right). Reprinted with permission from [20].

The STM used in this study is a commercial unit made by Omicron VakuumPhysik GmbH and housed inside an ultra–high vacuum (UHV) chamber resting atop a custom vibration isolation table. The isolation table was designed [48,49] to reduce vibrations to the pm level for all frequencies above ~ 2 Hz, through a combination of pneumatic damping of the table and *in–situ* spring suspension / magnetic levitation of the stage that supports the STM sample and tip.

The vacuum chamber that houses the STM is pictured in Fig. 2.2 with key main chamber components highlighted. The chamber is split into three stages[2], each separated by UHV compatible gate valves. Samples are introduced to vacuum in the first stage, which is pumped by a turbo pump backed by a roughing pump. To avoid vibrations during experiments these pumps are valved off and shut down after achieving base pressure with the help of a thermal bake [50]. The next stage employs a non–evaporable getter and small ion pump to reduce the pressure further before samples are introduced to the main chamber where the cleavage carousel and STM sit. Two NEGs, an ion pump, and a titanium sublimation pump bring the base pressure in the main chamber (following isolation from the first two stages) to ~ 10 picotorr for hydrogen, and more importantly, as seen in Fig. 2.3, below 0.1 picotorr[3] for all reactive molecules (XHV regime).

To obtain these impressive pressures, the vacuum system received an overhaul in 2008 after it was found that aluminum–bearing samples reacted with dilute, background

---

[2] This three stage arrangement has proven particularly stable; while the load lock is routinely cycled between atmosphere and UHV conditions, the main chamber has been under continuous XHV conditions for *seven* years now.

[3] These pressures are some of the lowest on Earth, exceeded only by accelerator beam lines [51], and are comparable to those in interplanetary space [52].

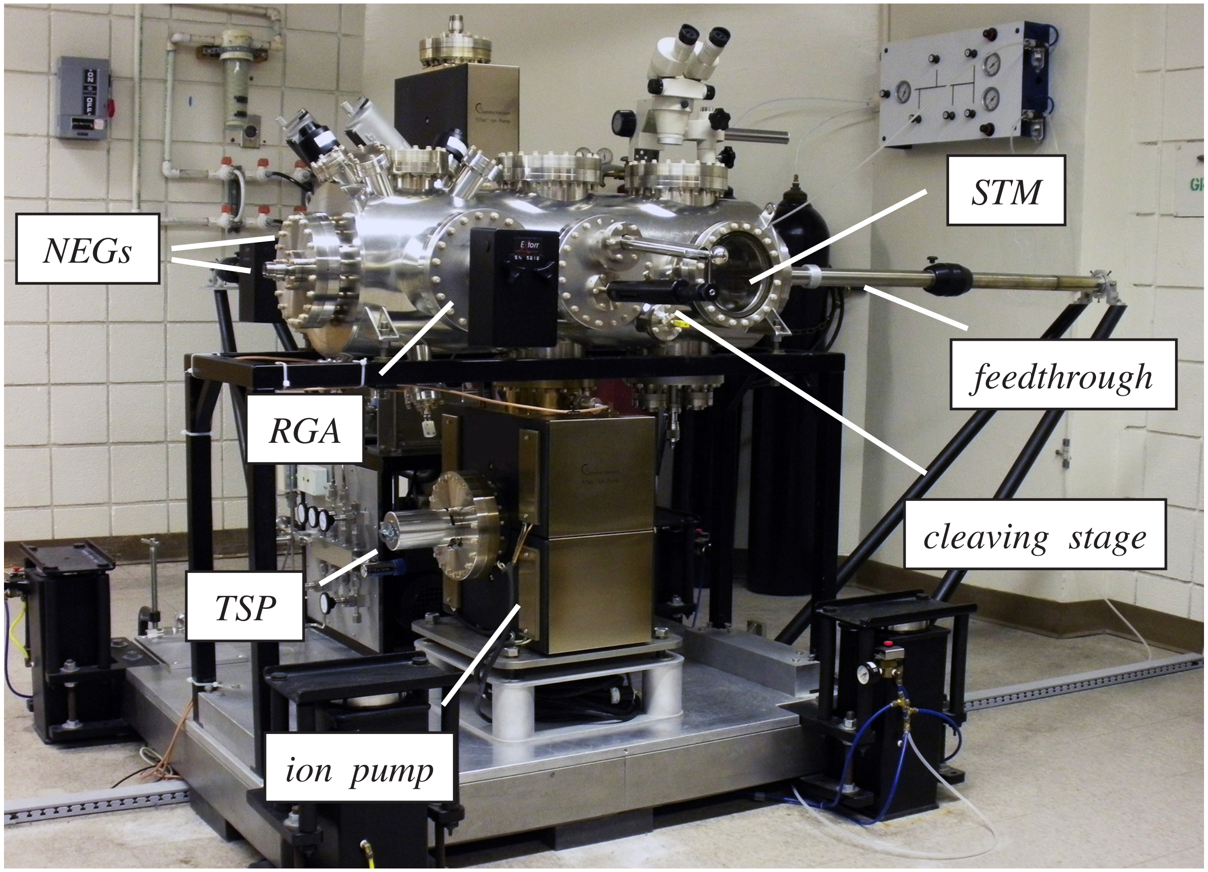


FIGURE 2.2. Main UHV chamber housing a comercial STM with major components highlighted. Pumping and motion feedthroughs were optimized for freshly cleaved Al–rich surfaces, which correspondingly allows for multi–week experiments on Al–free samples. Reprinted with permission from [39].

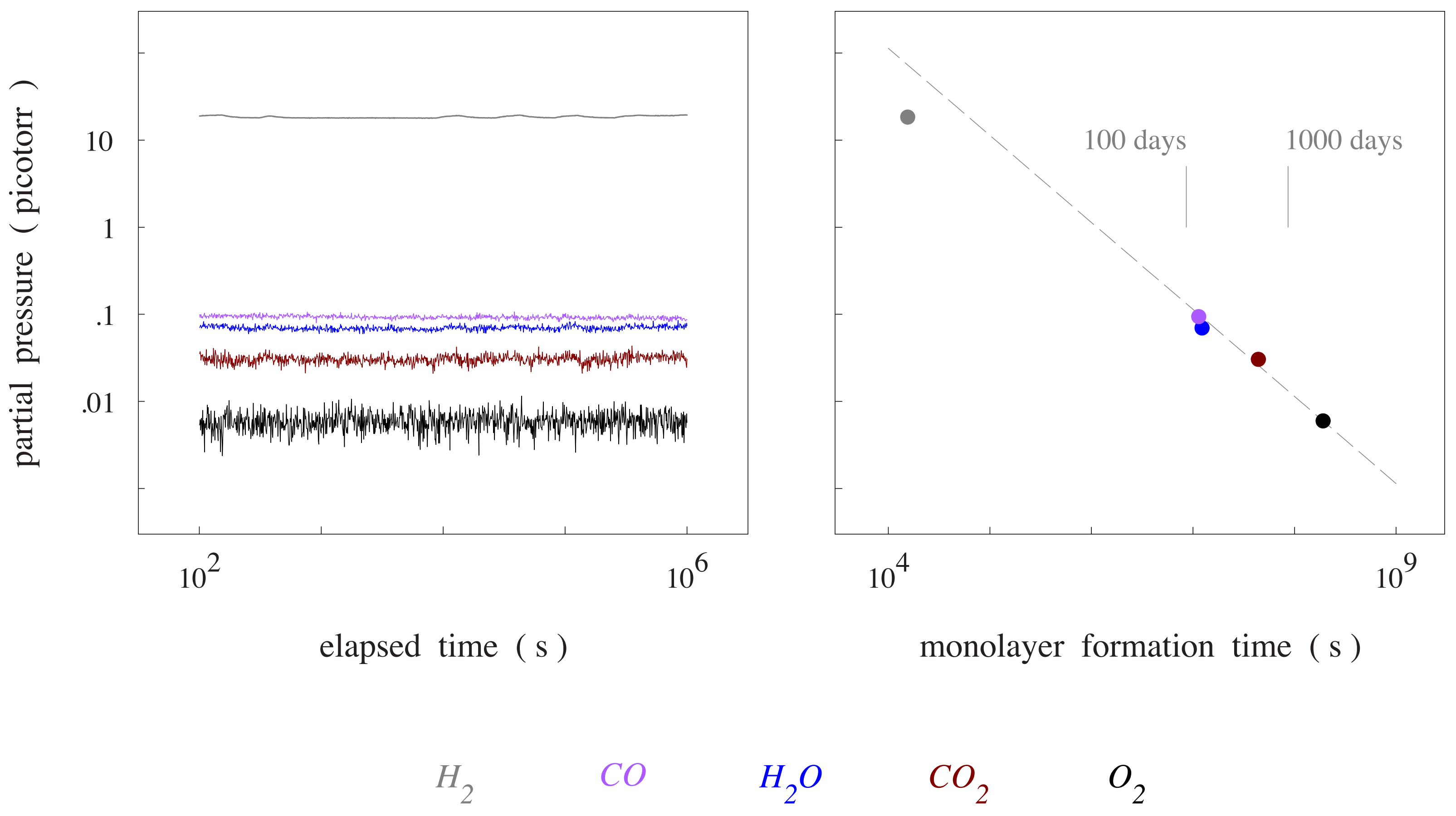


FIGURE 2.3. Steady–state XHV conditions measured by RGA with calibrated electron multiplier enabled (left) and associated (minimum) monolayer–formation times calculated for ideal circumstances (right). Adapted from [39].

levels of oxygen–containing molecules. Everything from pumping speed to sample mounting and cleaving were scrutinized and optimized [20,39]. This affords us at least a week to examine samples containing aluminum, and several months for samples such as the Al–free InAs/InAsSb superlattice, studied here.

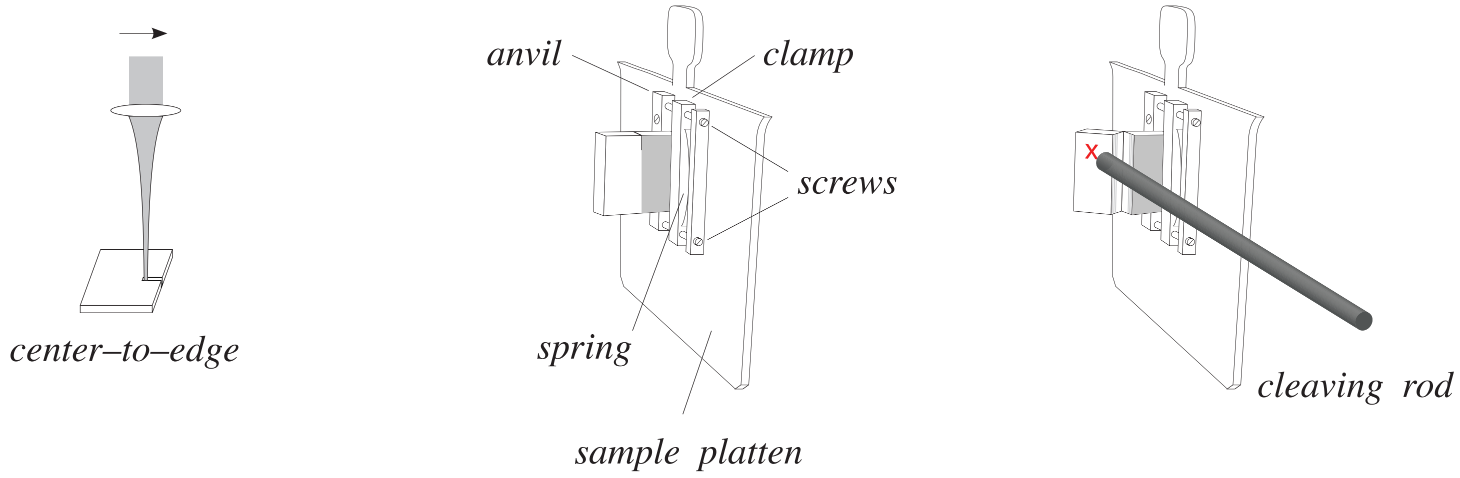
Sample Preparation

A representative history of the epitaxial growth can be reconstructed by studying orthogonal <110> cross sections with STM; access to these cross sections is facilitated by cleaving the sample along one of two principle crystal directions. In order to identify individual atoms, the freshly exposed surface needs to be atomically flat across the epitaxial layers and to navigate across the structure there shouldn't be steps or debris that stick out of the surface more than a few nm. Since the quality of the cleave is sensitive to each step of the sample preparation, a second, *ex–situ* cleaving station was constructed to allow quick turnaround for optimization of each parameter bearing on this all–important goal of atomically flat cleaves. As we explain the procedure here, we will briefly review these parameters, but more thorough write–ups of the optimizations can be found in previous theses [20,39].

Each quarter of a standard 2" wafer is diced into individual 5mm x 5mm squares suitable for transfer to the STM[4]. This dicing is done with a ESI laser trimmer system using a Nd:YAG laser operating at 1064 nm. This same laser is then swept from the center of the die outward (Fig. 2.4 (a)) along a specified <110> direction at a rate of

---

[4] Sandia National Labs kept a portion of the InAs/InAsSb quarter wafer for optical characterization.



(a) (b) (c)

FIGURE 2.4. Overview of major steps employed to create atomically–flat cleavage surface. Sample is a) notched with a laser scribing system perpendicular to a desired crystal plane, b) mounted, and c) precisely impacted leaving a surface that is (ideally) atomically flat throughout the epitaxial layers. Mounting depicted here reflects historical orientation. Adapted from [39].

12mm/s using 0.5 W to create a notch whose depth is nearly 1/3 the wafer thickness, and whose length is ~ 0.07”. The sample is sandwiched between two stainless–steel anvils secured to a sample platen, with the notch directed upwards and held in place by pressure applied via a Be-Cu spring (Fig. 2.4 (b)); when struck during cleavage (Fig. 2.4 (c)) this notch initiates fracture along the chosen crystal axis. Both the sample and the spring are aligned as close to the bottom of the two anvils as possible to encourage high–quality cleaves near the bottom of the sample where the STM tip is subsequently positioned.

Mounted samples are placed inside a load–lock that is initially pumped down to high vacuum, then heated overnight to drive off residual water and bring the pressure to ultra–high vacuum levels. The baked–out sample platens (Fig. 2.4 (b)) are subsequently transferred to the main vacuum chamber and placed in a carousel where – after closing all valves and letting the system equilibrate – a particular die is cleaved via impulsive impact driven by pneumatic actuation of a vacuum feed through (Fig. 2.4 (c)). The speed at impact for samples grown on GaSb substrates is targeted at ~ 300 mm/s for (–1–10) cleavage and ~ 150 mm/s for (1–10) cleavage.

Sample Crystallography

To minimize the number of outgas–inducing movements post cleavage the carousel and sample platen are oriented so that the sample is cleaved facing the same direction it will face when loaded into the STM (Fig. 2.5 right). This orientation breaks

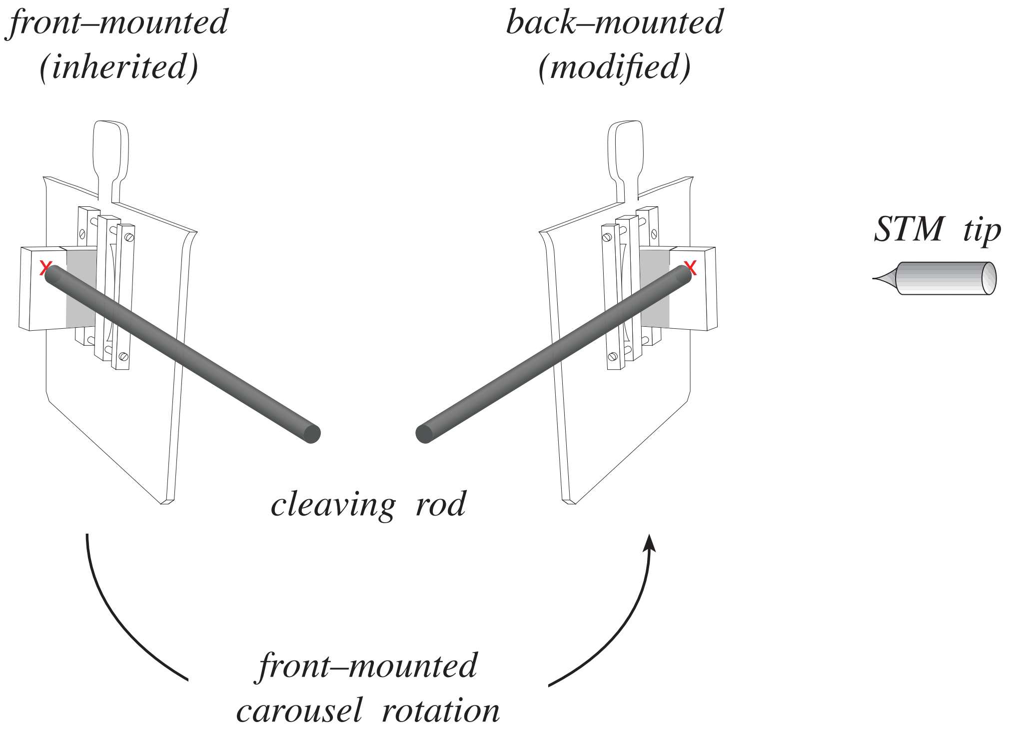


FIGURE 2.5. Traditional fastening of scored, sample dies to the front of an STM platen (left) entails carousel rotation – with its attendant pressure transient – following cleavage so the sample surface and STM tip face one another. Fastening these dies to the back of the symmetric, STM platen (right) circumvents any need for carousel rotation, since cleavage surface and STM tip are naturally aligned. This rotation changes the crystallography as seen by the STM tip. Shaded half of sample die is examined with STM. Adapted from [39].

with the historical precedent (Fig. 2.5 left) set by an earlier generation of students, so we pause here to explore its crystallographic ramifications as they relate to STM.

We begin by describing the inherited sample mounting and cleavage, illustrated schematically in Fig. 2.5 (left). The sample is always mounted notch up with the scribed (001) surface facing the cleavage tool, as this is needed for reproducible cleaves. Originally, the platen clamped the right half of the sample (as looked at from the cleaving tool) and the cleaver impacted the left half, breaking this piece off. The carousel was next rotated 180°, so that this cleaved surface then faced the tip, and the sample was subsequently transferred to the STM. This reoriented the [001] growth direction parallel to the cleaving rod and positioned the (001) growth surface to the right of (i.e. behind) the STM tip.

Cleaves are currently conducted with the sample clamped on the left half (again with the notch facing up and the [001] direction pointing towards the cleavage tool), and the right half of the die is cleaved off to expose the epitaxial layers, as shown in Fig. 2.5 (right). The sample is then transferred into the STM without rotation since the cleavage surface already faces the tip. If no other change is made, the surface examined with STM will then be opposite that historically looked at: the (110) surface on the right half of the die is the (–1–10) surface on the left half of the die, and likewise the (1–10) surface is complemented by the (–110) surface. Since these complementary planes are physically indistinguishable, the only visual change results from the 180° rotation of the (001) growth surface with respect to the STM tip.

To maintain the same (110) and (1–10) cleavage surfaces used historically, the notch direction should be reflected and the sample rotated 180° about the [001] direction. As seen in Fig. 2.6 this was done with the (1–10) cleave, where the laser scriber (Fig. 2.4 (a)) was previously swept in the [–1–10] direction but is now swept in the [110] instead. Unfortunately, the same change was not implemented for the (110) cleave, where the laser is still swept in the [–110] direction yielding a (–1–10) surface when the right half of the die is cleaved off; this oversight is of no scientific consequence, however, since both (110) and (–1–10) surfaces are physically equivalent. The currently accessible cleavage planes, (–1–10) and (1–10), are shown in relation to the STM tip in Fig. 2.7.

The 180° rotation of the [001] growth direction is better appreciated by considering the inherited and modified sample orientations from the viewpoint of the piezo–tripod in which the tip illustrated in Figs. 2.5 and 2.7 is mounted. The 180° carousel rotation needed to reorient the inherited sample mounting in Fig. 2.5 positions either (110) or (1–10) cleavage faces towards the STM tip and the [001] growth direction to the right — a linear combination of negative x and positive y — as shown in Fig. 2.8 (left). The modified sample mounting, which requires no reorientation, correspondingly positions either (–1–10) or (1–10) cleavage faces in front of the STM tip, with the [001] growth direction now pointing to the left — a linear combination of positive x and negative y— as shown in Fig. 2.8 (right).

Fig. 2.9 contrasts the resulting orientation of cleavage–exposed crystal axes accompanying inherited and modified sample mountings with respect to the x–y scan coordinates (image coordinates) in Fig. 2.8. As should be clear from Fig. 2.9, reversal of

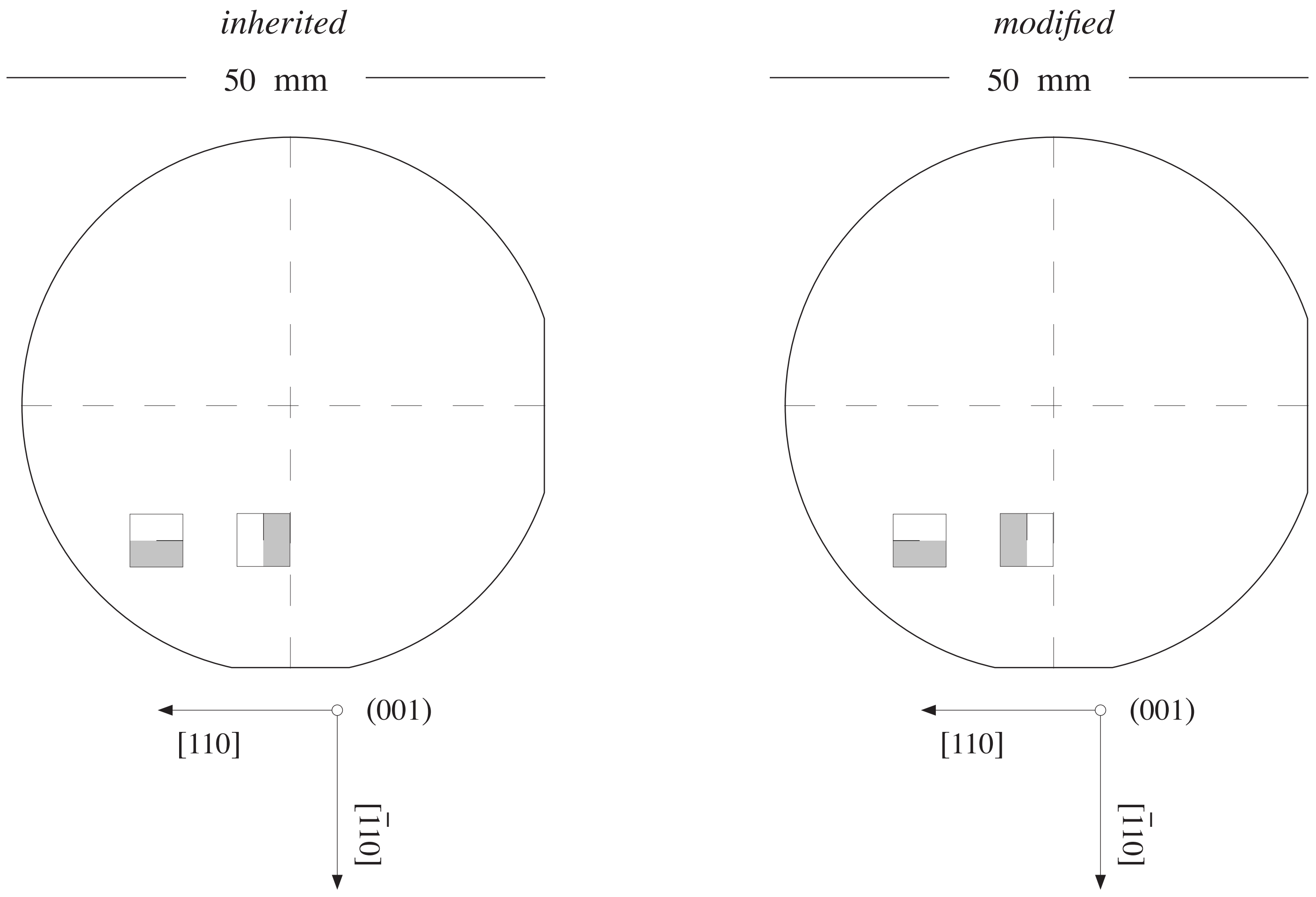


FIGURE 2.6. A two–inch wafer is diced into individual 5 mm x 5 mm STM samples with a laser scribing system. Inherited orientation aligned notches with either the [–1–10] or [1–10] directions to facilitate exposure of orthogonal (1–10) and (110) surfaces respectively following cleavage *in–situ*. Modified notch orientation mirrors [110] direction to [–1–10] to preserve a (1–10) cleavage surface for the modified sample mounting in Fig. 2.5. A corresponding change for the [1–10] notch was overlooked, resulting in a (–1–10) rather than (110) surface.

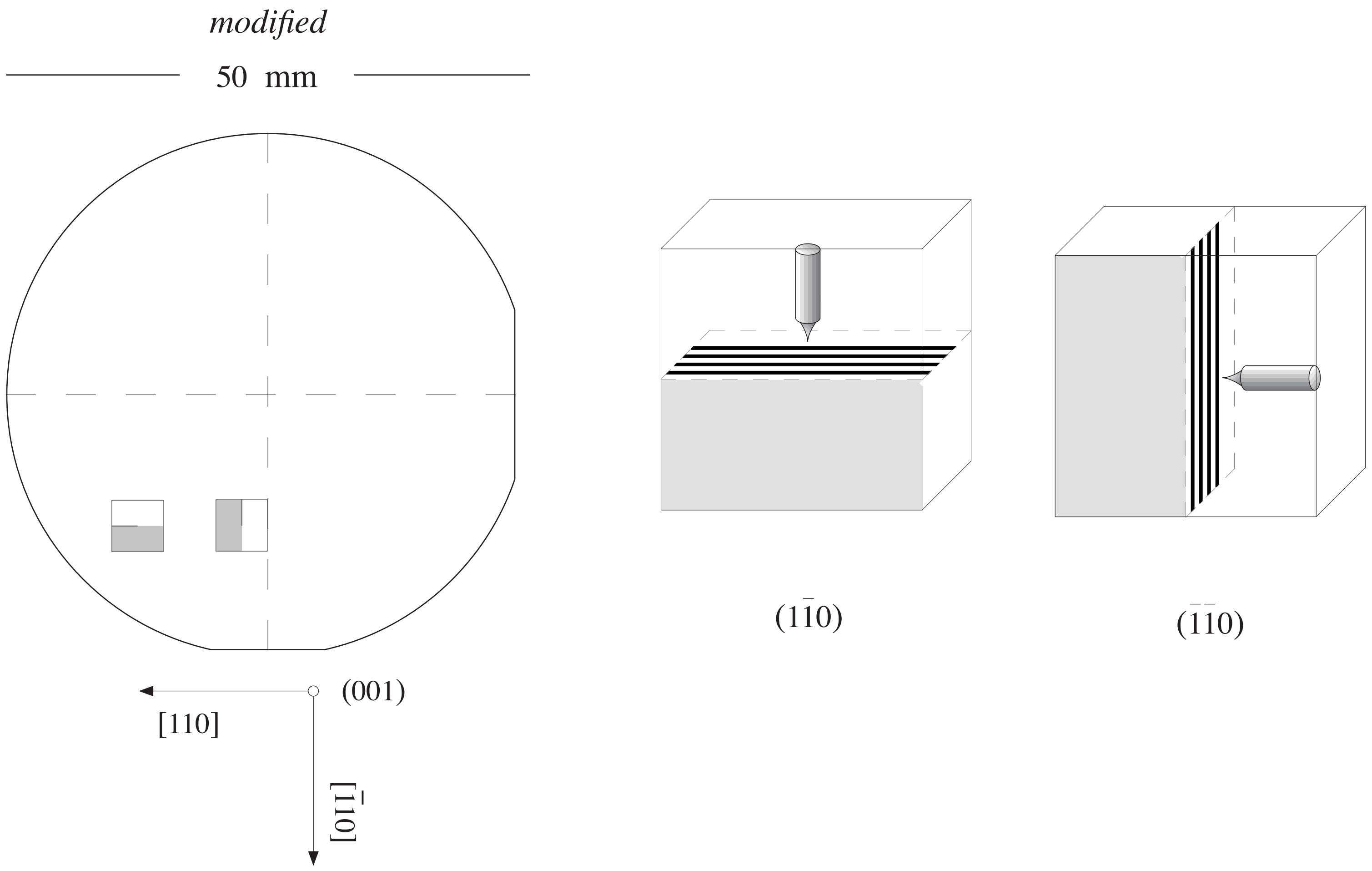


FIGURE 2.7. Modified orientation of sample dies, reproduced from Fig. 2.6 (right), and resulting perspective views of STM tip in relation to epitaxial layers for orthogonal cross sections. Growth direction is out of page, shaded half of sample die is transferred to the STM, and unshaded portion discarded. Adapted from [39].

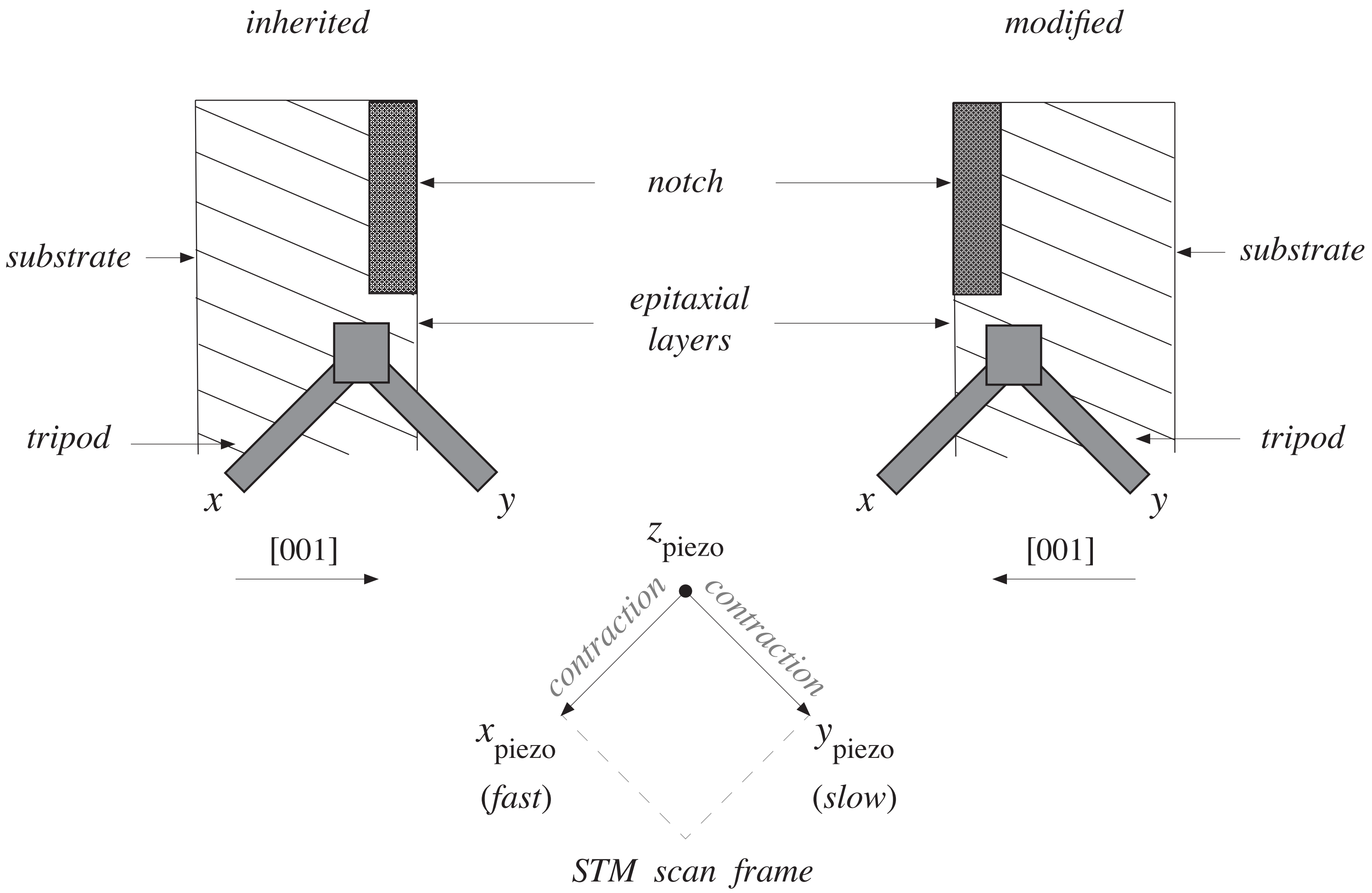


FIGURE 2.8. Schematic comparing inherited and modified sample orientation following transfer to the STM, as viewed from behind the tripod. The inherited [001] direction is a linear combination of negative *x*, positive *y*; conversely the current [001] direction is a linear combination of positive *x*, negative *y*. Positive *x* and *y* values are realized through contraction of fast– and slow–scan piezos respectively.

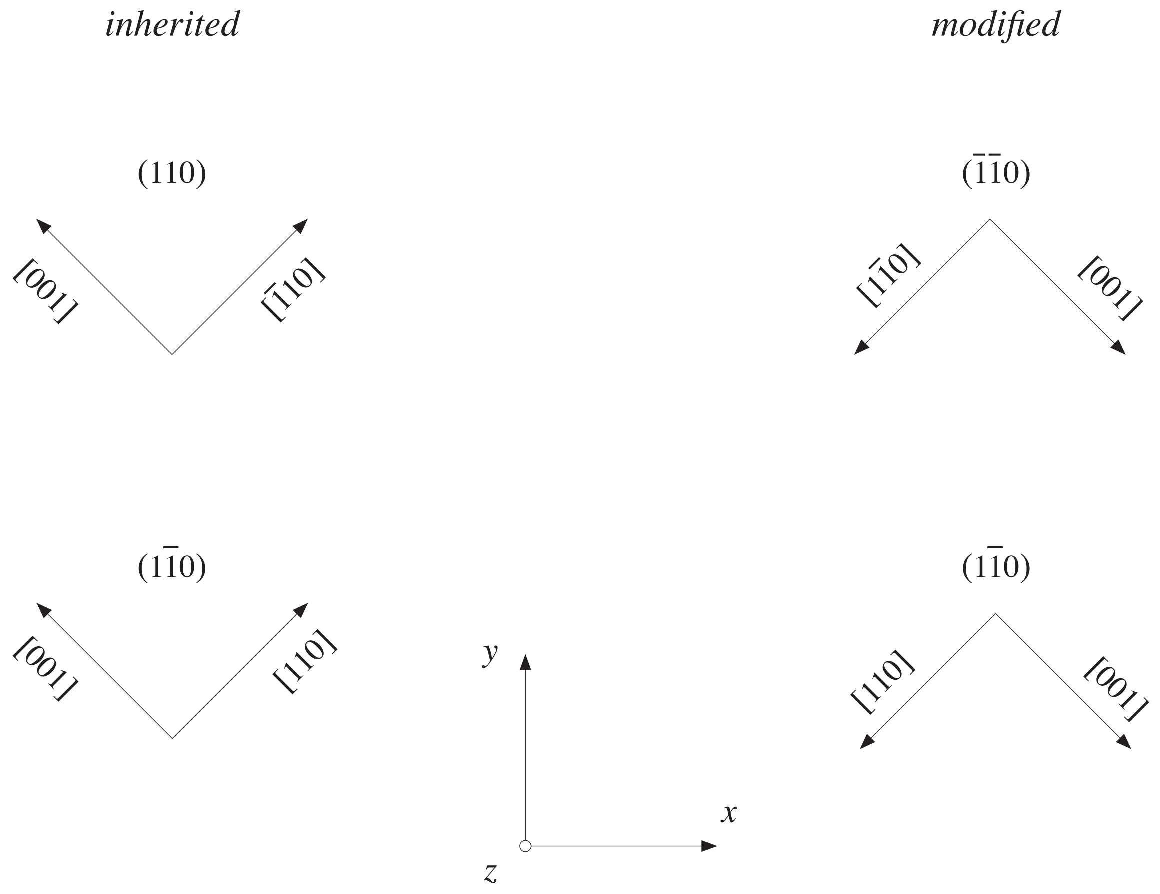


FIGURE 2.9. Inherited and modified crystal axes in relation to standard *x* and *y* image axes on accessible cleavage cross sections. Modified axes on (–1–10) surface are obtained by rotating inherited (110) axes first by 180º about the [001] direction to present a (–1–10) cleavage surface, then by 180º about the [–1–10] direction. Finally the [–110] vector is reflected to obtain the right–handed coordinate system shown. Modified axes on the (1–10) surface are obtained by rotating 180º about the [1–10] direction.

the [001] direction manifests in the STM images (Fig. 2.10) in the following way: with an inherited sample mounting [001] runs from lower right to upper left, but with the modified sample mounting it now runs from upper left to lower right.

The growth direction in an InAs / InAsSb superlattice is unambiguously identified by symmetry–breaking antimony segregation, which creates a “charging” / “discharging” cycle in the antimony fraction as a function of distance in the growth direction[5]. Although we will adhere to the rigorously correct identification of crystal directions here (Fig 2.10, black and Fig. 2.9, right) others [20,23,39] have opted for a simplified labeling (Fig 2.10, grey and Fig. 2.9, left) that amounts to the presumption of (110) versus (–1–10) cleavage.

Scan–related Image Distortion

As previously described in connection with Fig. 2.8, a sharp tip is placed at the vertex of a piezoelectric tripod scanner whose mutually–perpendicular fast– and slow–scan axes are fixed at ± 45° with respect to (001) crystal planes, or, what amount to the same thing, ± 45° with respect to <110> directions. This scanner controls the fine movement of the tip (in the first *x–y* quadrant) as needed to form an image as follows: the fast–scan piezo is incrementally contracted stepping through the first line of an image (increasing *x*, forward line scan), followed by an incremental expansion re–sampling the first line (decreasing *x*, reverse line scan), before the slow–scan piezo is incrementally contracted (increasing *y*) to move to the next line, each increment forming

---

[5] In some non–common–atom superlattices (such as InAs/GaSb) the growth direction can also be identified by pinpointing heterojunction–specific interface bonds [47].

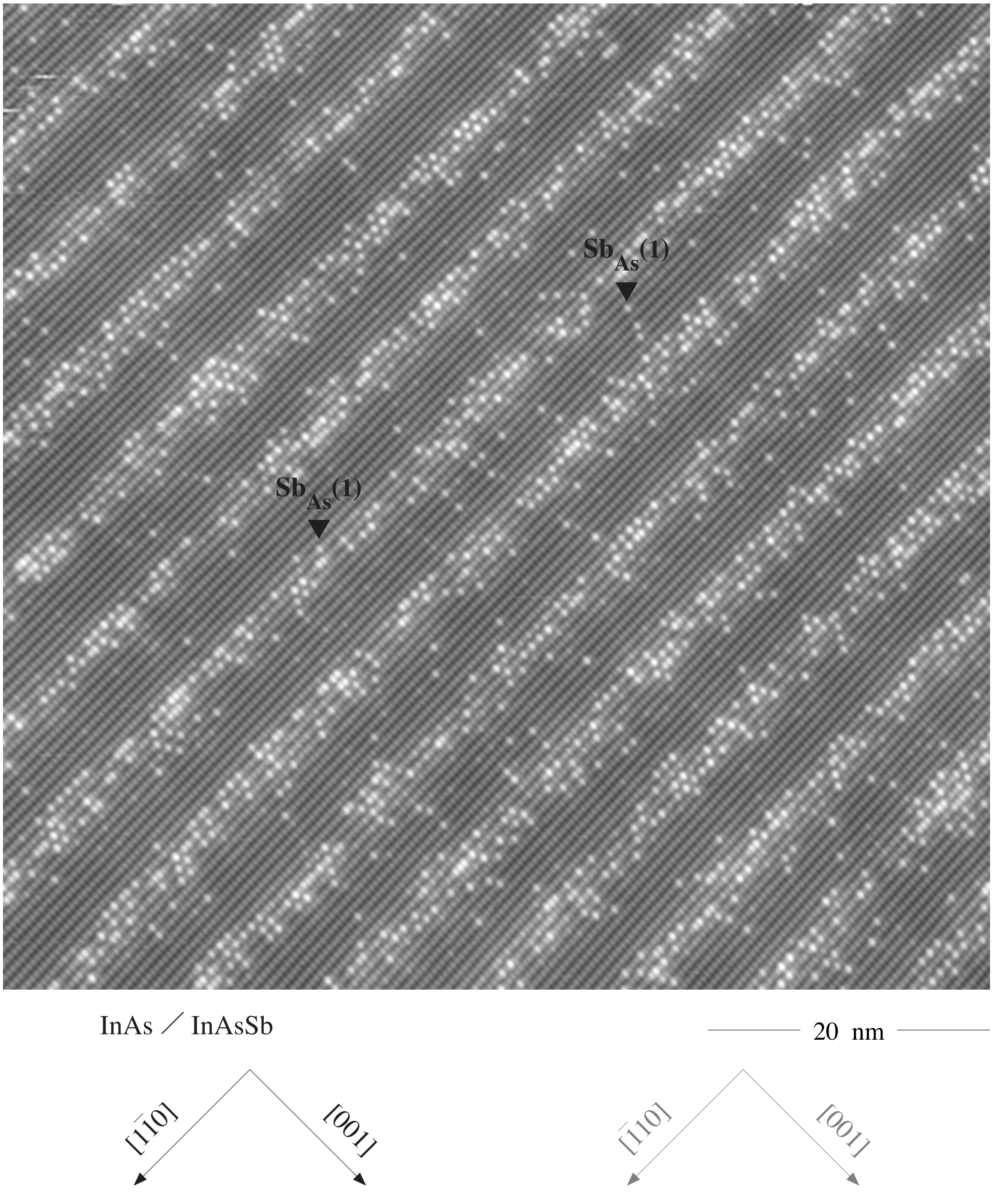


FIGURE 2.10 Atomic–resolution STM image of the anion (Sb, As) sublattice from a type–II InAs / InAsSb superlattice. Antimony–for–arsenic replacement within the cleavage–exposed plane is identified by carets. Growth direction is from top–left to bottom–right. Simplified crystal axes (grey) assume the cleavage plane is (110), while rigorous axes (black) correctly identify cleavage plane as (–1–10). Adapted from [20]

the real–space pixels visualized in the simultaneously acquired forward and reverse STM images. Images may be similarly obtained in quadrants 2–4 by inverting the rastered directions of the x and / or y piezos (Fig. 2.8).

Elongation of the piezos with each computer controlled voltage step was historically calibrated by counting the number of substrate atoms that fit into a given window size, but two extraneous circuits in the piezo–driver control board have since been disabled. The first fed a small portion of the slow–scan voltage into the fast–scan signal, which could be used to counter image skew. The second circuit allowed for arbitrary rotation of the scan frame, but was later found to have a defective bit that prevented raster scans with a truly 0° scan angle. It was decided that any subsequent recalibration would be enforced after the fact, using average lattice constants for the entire image as described below, instead of adjusting the step size of the voltage input to the piezo scanners to compensate for these modifications.

The fine movement of the tripod may also be used in another way. Offset voltages may be applied to translate the tip between images – up to ~ 1 micron in either the lateral, <110>, or vertical, [001], directions – so that successive images may be strung together to form a continuous survey. Coarse movement[6] (any distance greater than 1 micron) on the other hand, necessitates translating the entire sample with a slip–and–stick piezo movement [53]. The drawback of this type of mechanism is that the sample stage has a very large mechanical inertia and tends to continue drifting in the same direction (Fig. 2.11) for several hours after intentional movement has ceased.

---

[6] This coarse movement is used to initially position the epitaxial layers directly beneath the tip.

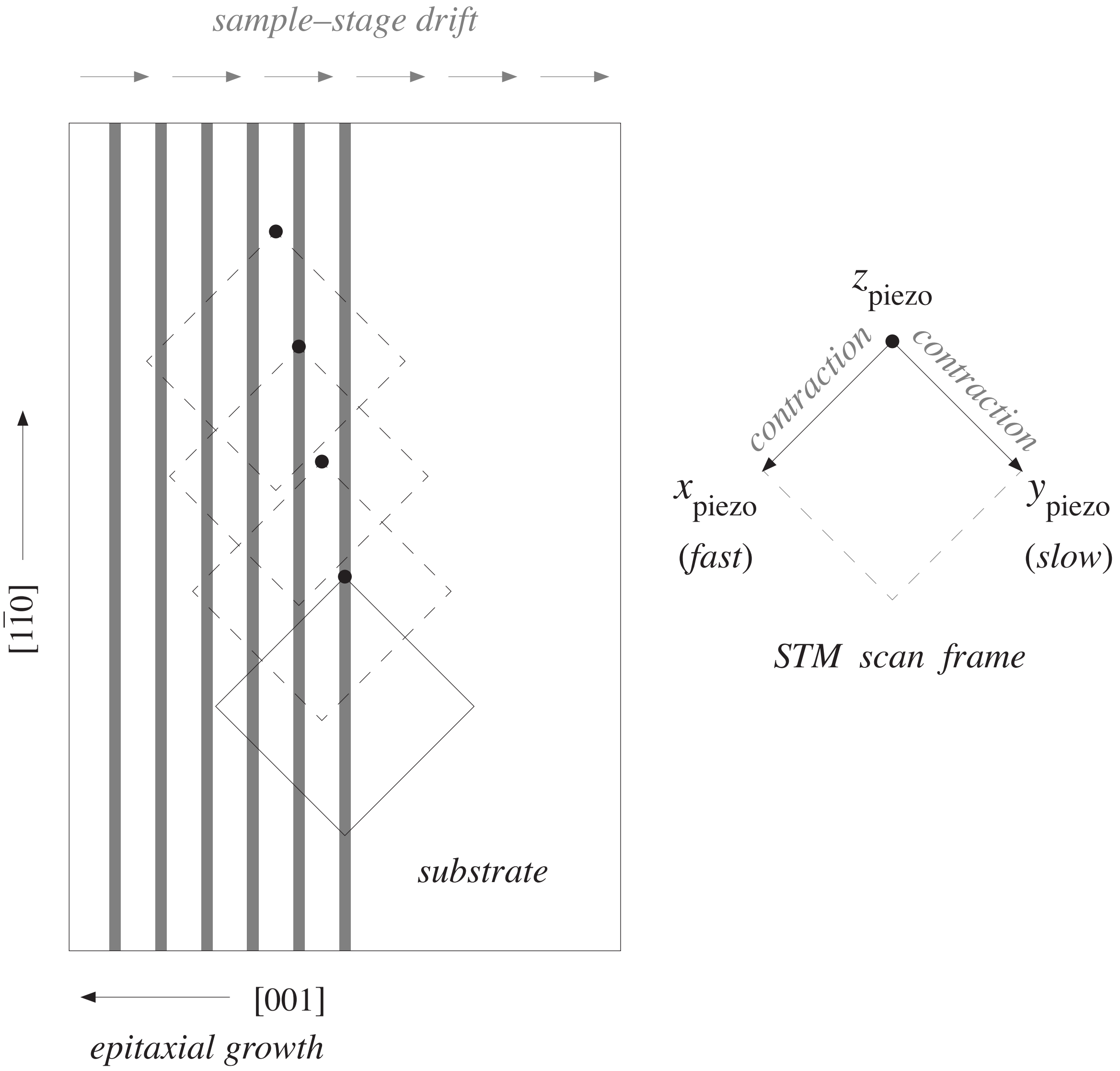


FIGURE 2.11. Schematic diagram illustrating sample drift (left) due to piezo–inertia of the sample stage as the scan frame is displaced vertically in the [1–10] direction. This stage coarsely positions the epitaxial layers under the tip, and moves either left (in growth direction), or right (opposite growth direction) with respect to the tip. The induced drift typically requires several hours to settle. Also illustrated (right), an STM image is formed via contraction of the *x*– and *y*–piezos. Cleavage plane coincides with plane of paper. Growth direction is right to left. Reprinted with permission from [39].

When planning an experiment this must be factored in to allow ample time for the sample to settle in position; once settled the sample inertia is typically no longer a concern.

The tripod itself also has inertia and will consequently display creep (where the tip needs time to adjust to changing directions). This creep appears at the bottom of every image (Fig. 2.12, left) due to a rapid return of the slow–scan piezo to the scan origin at the conclusion of every image. It also shows up in a more subtle way, when the image frame is offset parallel to the scan direction, as we discuss later in this chapter. The piezoelectric ceramics that move the tripod scanner are also notoriously problematic [54]. Under ideal expansion the change in length of the piezo material would depend only on the change in voltage applied to the material; in a realistic piezo, however, the amount the material expands depends not only on the size of the voltage step, but also on the voltage itself, making the response nonlinear (Fig. 2.12 right). To make matters worse, the expansion additionally depends on the history of the voltage applied to the material, making the response hysteretic; this hysteresis affects forward and reverse images differently. We will rely on forward images as examples in the remainder of this chapter and comment on any differences that pertain to the reverse images. The severity of the resulting distortions is also dependent on the size and history of the offset voltages applied to the piezos, so for that reason it is important to analyze survey images, whose offsets that are relatively small.

In a typical forward image the piezo nonlinearity just described results in curved, or bowed, <110> rows and [001] columns. Bowing along the row is subtle, but is

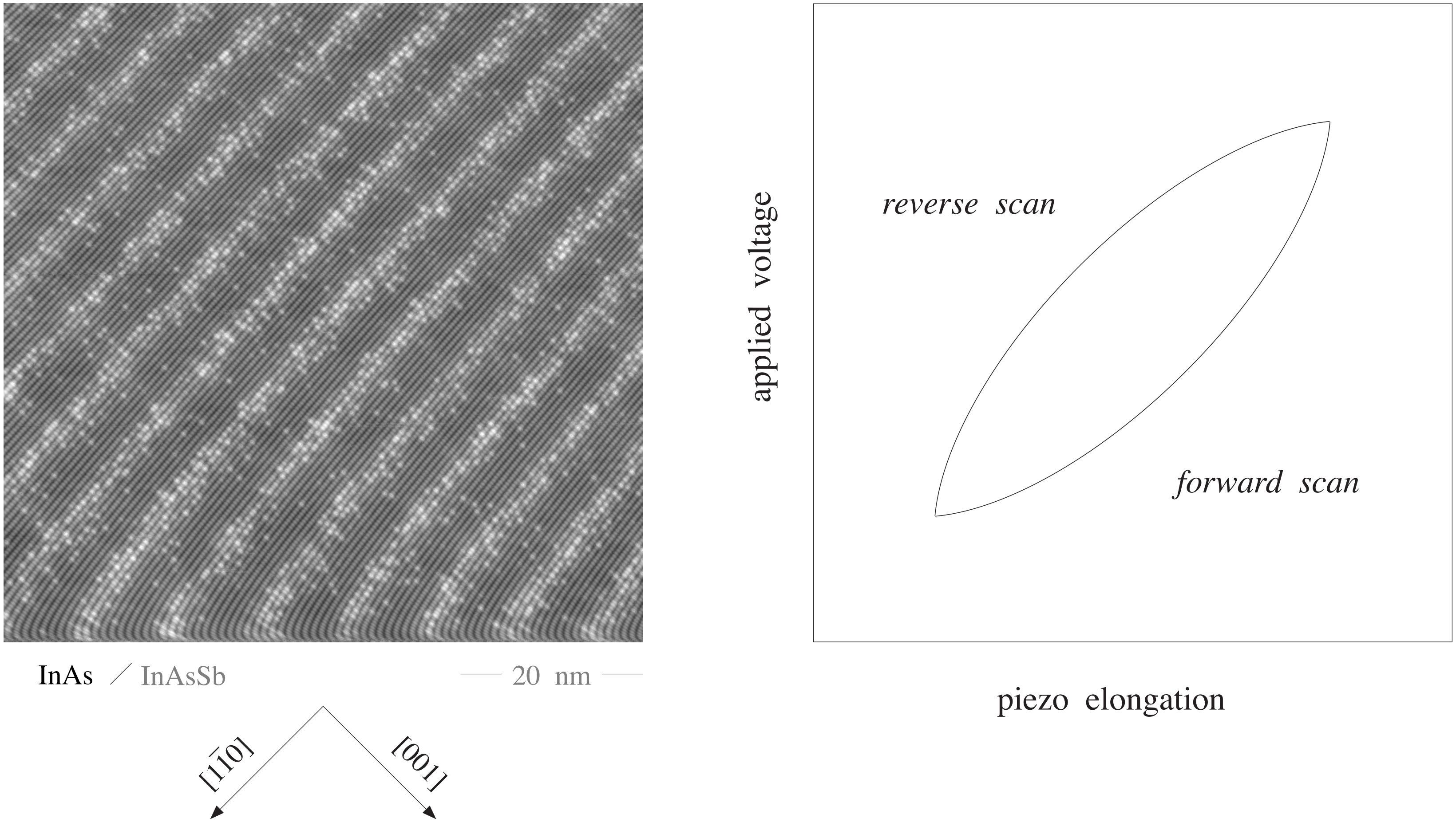


FIGURE 2.12. Piezo–scanner nonidealities include, but are not limited to, piezo–creep (left) and hysteresis (right). The whiplash present at the bottom of the image is a direct consequence of piezo–creep; a subtler, but nevertheless undesirable symptom of piezo–creep also occurs throughout the image if the y–offsets are inadvertently taken to be parallel to the slow–scan direction. Piezo non–linearity (hysteresis), on the other hand, is unavoidable; using a restricted area of the image where these piezos respond most linearly proves a suitable workaround. Adapted from [39].

highlighted when a box is aligned to a single <110> row in the lower portion of an image (Fig. 2.13, left) and then projected upward to compare with a box aligned to the same <110> row in the upper portion of the image (Fig. 2.13, right). A deflection of between one–half and one monolayers or ~ 0.5° is typical. A box likewise aligned to a single [001] column in the lower portion of an image (Fig. 2.14, left) and projected up to compare with a box aligned to the same [001] column in the upper portion of the image (Fig. 2.14, right) shows a deflection of about five <110> lattice spacings equivalent to ~ 2.7°.

Hysteresis in the piezos causes the magnitude of these row and column deflections to interchange between forward and reverse images, where the [001] columns are minimally bowed, but the <110> rows are deflected significantly. Thus <110> interface profiles extracted from reverse images are hopelessly distorted, and only forward images may be relied upon for measurements of interface roughness described in Chapter III. Likewise the reciprocal lattice vectors inferred below from forward and reverse images will differ, but the periods measurements outlined in Chapter III are designed to be insensitive to these distinctions.

Image Analysis in Reciprocal Space

The deflections above are closer to a maximum angle, rather than an average, but they serve to illustrate the challenge faced by analog correction circuitry or in our case a naïve computer algorithm employed to translate between the digitized image and the physically meaningful [001] and <110> coordinates. We instead turn to reciprocal space

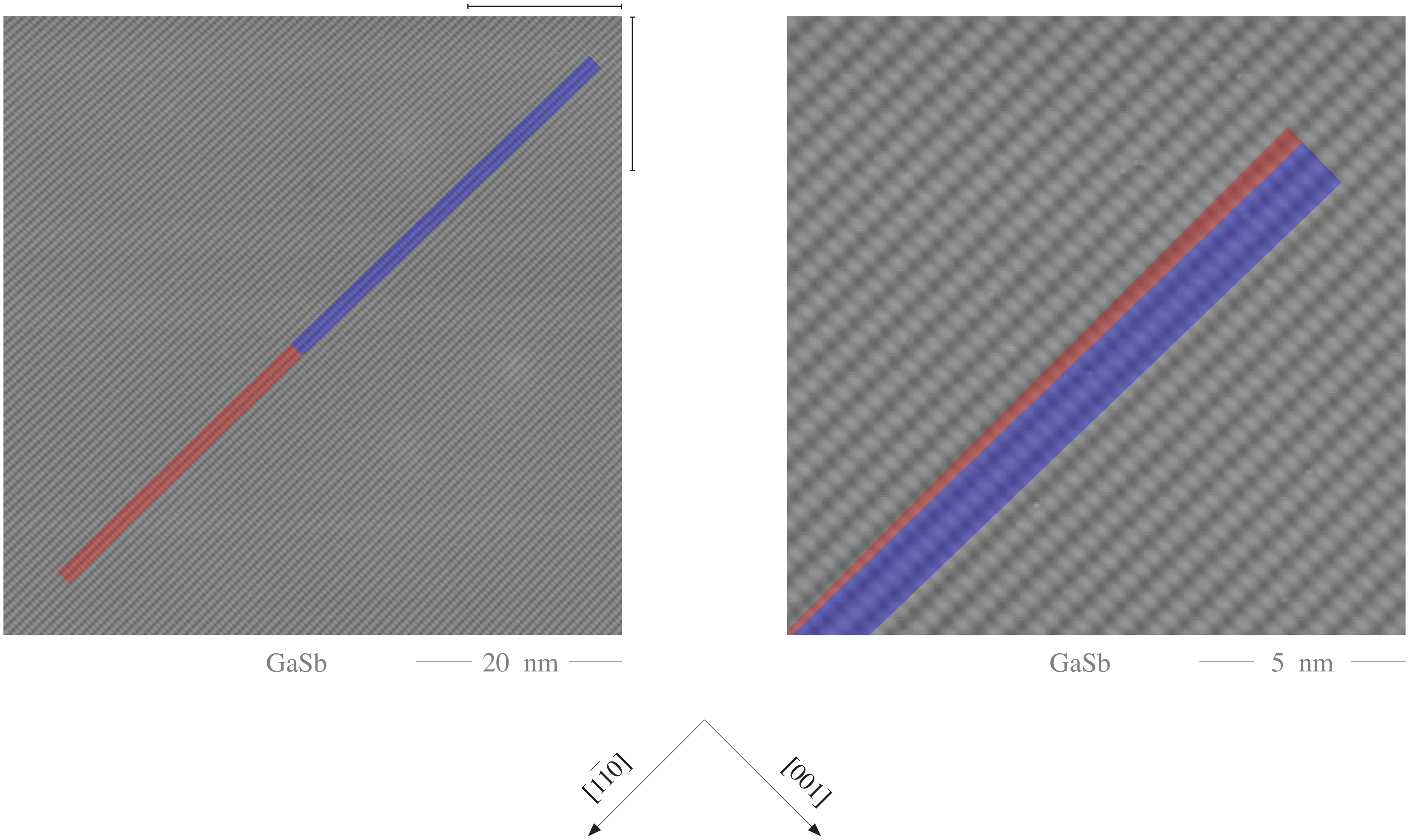


FIGURE 2.13. Atomic–resolution image of the antimony sublattice in a GaSb buffer layer (left). Single <110> row (aligned with atomic crests) is split in half with bottom–half segment projected upward to accentuate misalignment with respect to top–half segment. A close–up view (right) of the top–right corner exhibits a small misalignment (red shade) somewhere between one–half and one surface monolayer corresponding to ~ 0.5°. Growth direction is top left to bottom right.

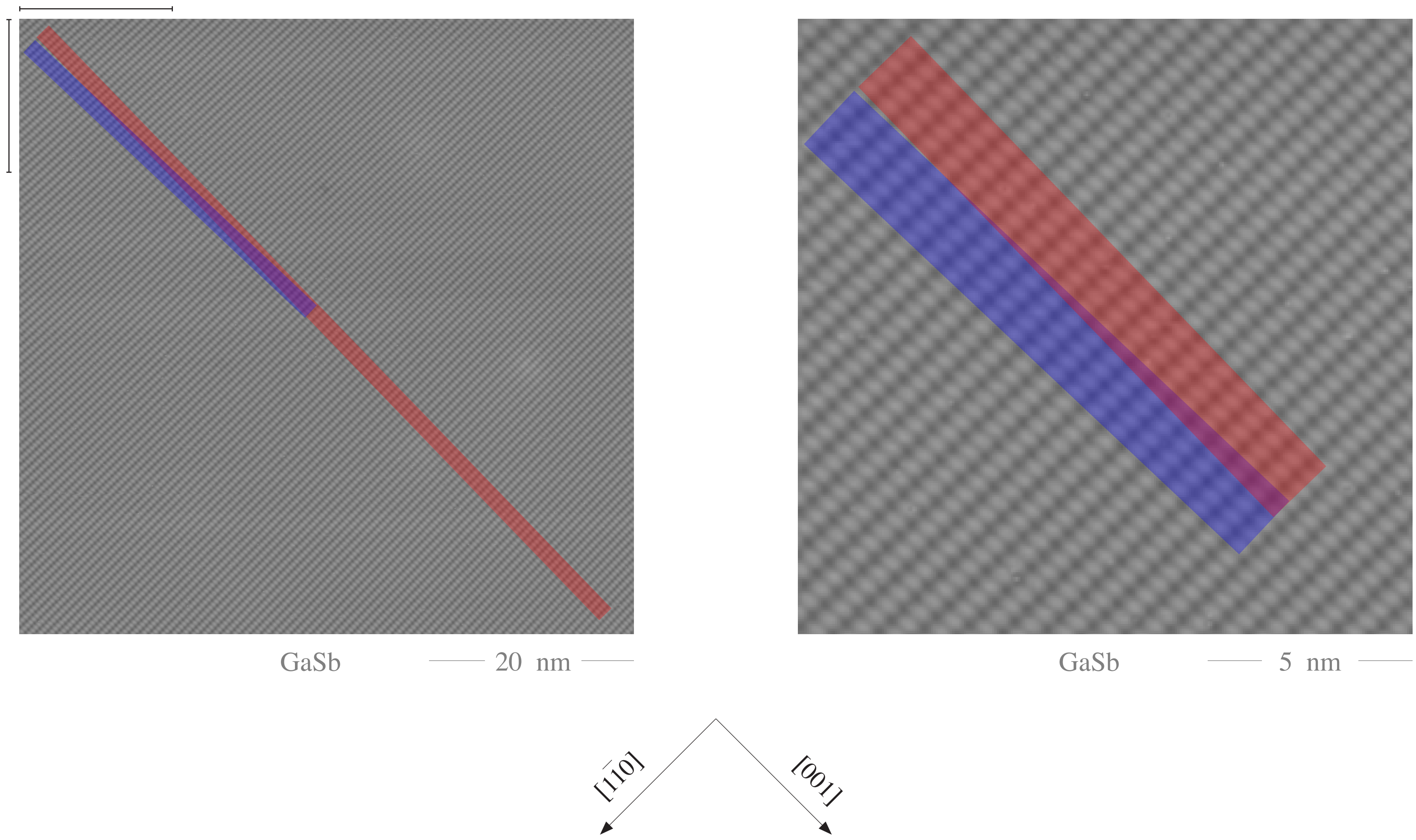


FIGURE 2.14. Atomic–resolution image of the antimony sublattice in a GaSb buffer layer (left) reproduced from Fig. 2.13. Single [001] column (aligned with atomic crests) is split in half with bottom–half segment projected upward to accentuate misalignment with respect to top–half segment. A close–up view (right) of the top–left corner exhibits a misalignment (red shade) about five <110> lattice spacings corresponding to ~ 2.7°. Growth direction is top left to bottom right.

where distortion averaged over an entire image may be judged by Fourier transforming our real–space image, plotted as a logarithmically–scaled power spectrum in Fig. 2.15. Four "peaks" (encircled in white) corresponding to the unit atomic mesh in real space dominate this map; each of these reciprocal lattice vectors (RLV) is smeared out due to the real–space bowing noted above, and close inspection reveals each peak is actually split into two distinct maxima. As we'll see in a moment, there is a certain amount of symmetric ringing in the discrete Fourier transform (DFT) due to finite window effects, however this ringing decreases rapidly with distance and therefore would not cause this peak to bifurcate. It's helpful to remember that large frequencies (such as the extent of the DFT window) represent small distances (such as the real–space resolution), since all images in this chapter have a resolution of 1 pixel / Å, every DFT will have the same extent, however each DFT has been cropped to the inner ¼ (by area) of reciprocal space to focus attention on the region inside the RLVs.

The frequency resolution in the power spectrum shown in Fig. 2.15 has been artificially increased by zero–padding along both the *x*–and *y*–scan axes. Much like the quantum uncertainty principle, real– and reciprocal–space resolutions cannot both be made arbitrarily small. The reciprocal–space pixel size,

$$\Delta k = \frac{1}{N\Delta x} = \frac{1}{L} \,, \tag{2.1a}$$

is set by the inverse of the total length (L) [55] given by the number of pixels (*N*) times the real–space pixel area (Δx). The power spectrum size, then, is given by

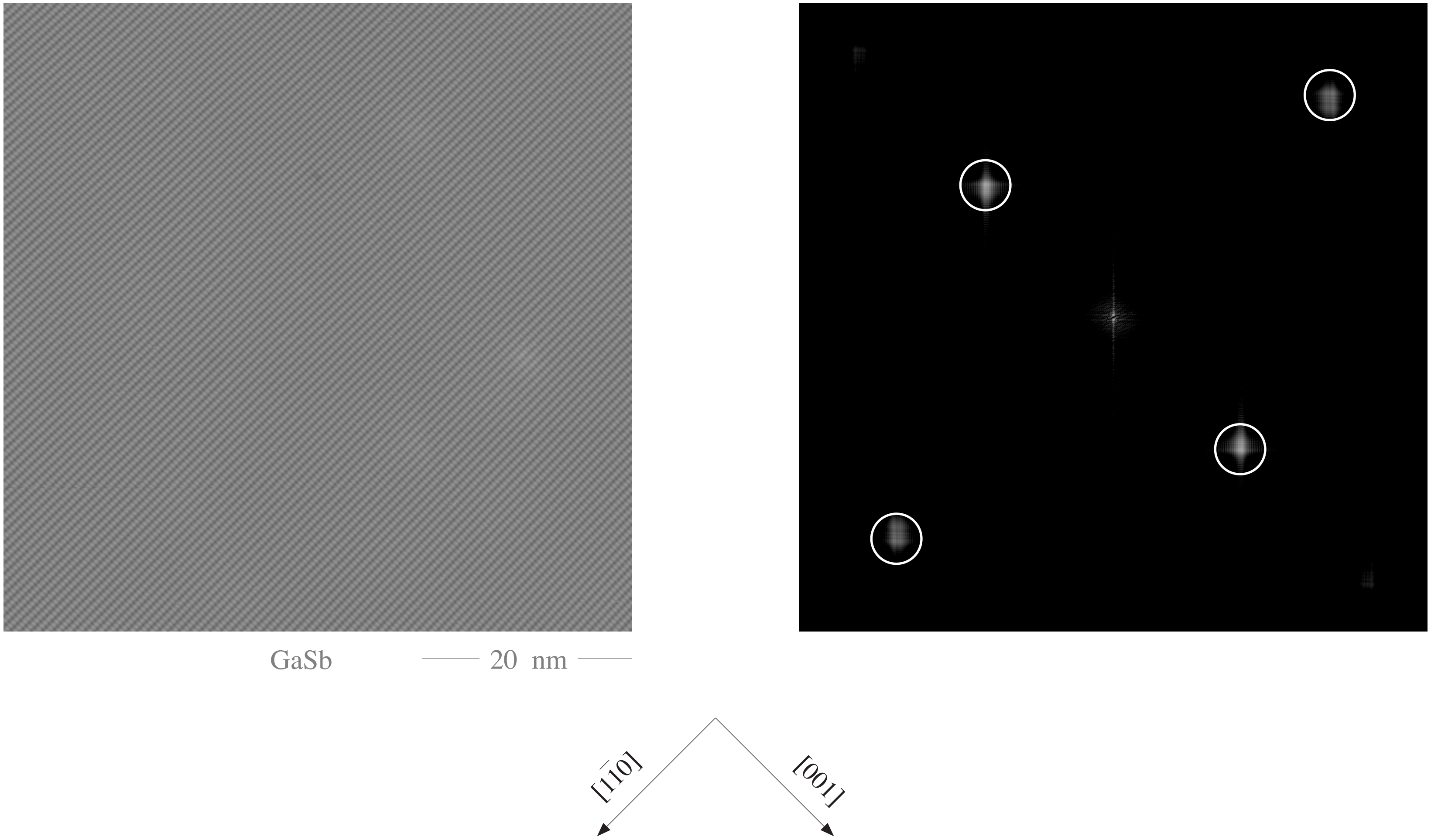


FIGURE 2.15. Atomic–resolution, STM image of the Sb sublattice over a GaSb buffer (left) and corresponding reciprocal–space map (right). Highlighted spots denote the [001] and <110> reciprocal–lattice vectors defining the two–dimensional surface mesh. Power in each reciprocal–lattice vector is distributed over multiple spots when the entire image is digitally transformed. Growth direction is from top left to bottom right.

$$N\Delta k = \frac{1}{\Delta x} \ , \tag{2.1b}$$

where the real–space pixel area sets the high frequency limit[7], which will remain constant so long as the real–space image resolution remains unchanged. To decrease the reciprocal–space pixel area we need to either obtain larger images – which are experimentally limited by the distortions just described – or, alternatively, pad the data with zeros to increase the total length (zeropadded length $L' > L$) over which the DFT is calculated.

Parseval's theorem[8],

$$\sum_{n=0}^{L-1} |f_n|^2 = L \sum_{k=0}^{L-1} |F_k|^2 \ , \tag{2.2}$$

states that the sum of the direct–space image power is the same as the summed power in the reciprocal–space spectrum. To account for zeropadding we can rewrite (2.2) in terms of zeropadded values (labeled by primes) as

---

[7] Since reciprocal space is inversion symmetric the range of k values is actually $-1/(2N\Delta x)$ to $1/(2N\Delta x)$, and the maximum frequency is known as the Nyquist frequency [55].

[8] The exact form of Parseval's theorem depends on the DFT convention utilized. The form in (2.1) corresponds to the signal processing convention [55] adopted in this manuscript.

$$\sum_{n=0}^{L'-1} |f'_n|^2 = L' \sum_{k'=0}^{L'-1} \left|F'_{k'}\right|^2 . \tag{2.3}$$

The total power in the zeropadded direct–space image – the left–hand side of Eq. 2.3 – is clearly the same as the power in the original image – the left–hand side of Eq. 2.2. In order for the right–hand sides to be equivalent, the zeropadded powers at integer *k* must be smaller than the original powers by two factors of ${}^{L'}\!/_{L}$,

$$|F_k|^2 = \left(\frac{L'}{L}\right)^2 \left|F'_{k'}\right|^2 \tag{2.4a}$$

where

$$k = k'\,\frac{L}{L'} \;\; \text{when}\; k' \; mod \left(\frac{L'}{L}\right) = 0 \tag{2.4a}$$

the first due to the scaling in (2.3) versus (2.2), and the second from computing the sum over the densely–sampled points described by (2.1a).

To help visualize this concept, consider the one–dimensional example of a rectangular box (Fig. 2.16, left), where the length has been increased 5–fold $\left({}^{L'}\!/_{L} = 5\right)$ by adding zeros to the end of the data stream; the resulting power spectrum (shifted so that the reciprocal–space origin is centered) is interpolated (Fig. 2.16, right), reducing

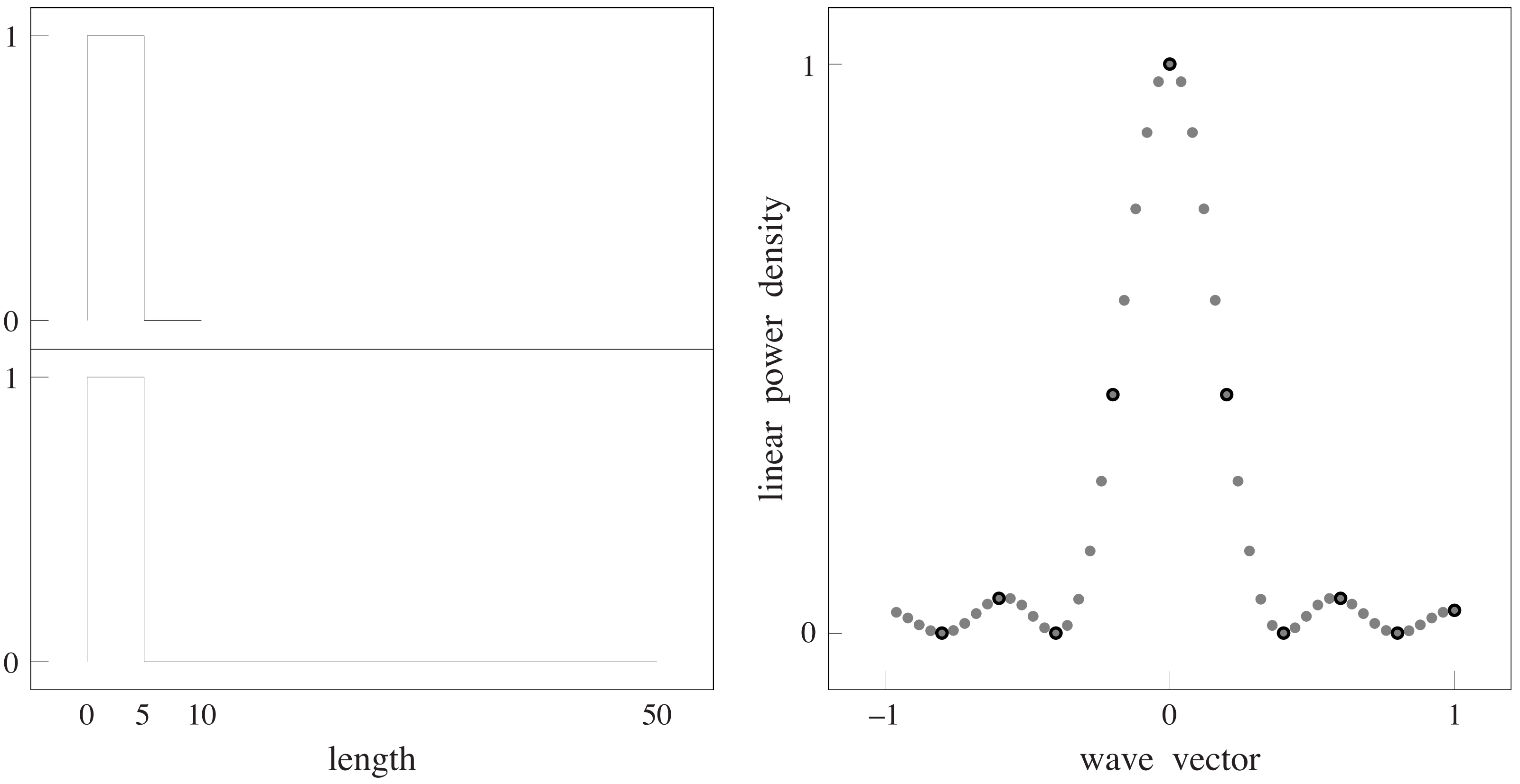


FIGURE 2.16. Rectangular box (left, top) and zeropadded box (left, bottom) illustrate the increase in reciprocal–space resolution (right) obtained by increasing the length in real space. When rescaled to unit amplitude the zeropadded spectrum (grey) reproduces the original spectrum (black) while increasing resolution by a factor of five; this rescaling masks the fact that the same total power (which is constant according to Parseval's theorem) is now spread over more reciprocal–space bins when zeropadding.

the uncertainty in the peak maximum by this same factor[9]. Rescaling both spectra to unit amplitude eliminates the factor of $\left(\frac{L}{L'}\right)^2 = \frac{1}{25}$ in Eq. (2.4).

Zeropadding is a standard technique in image processing [56] and is used for calculating correlations and convolutions as well [57]. All image power spectra that are shown in this document have been zeropadded. In this chapter, where we are focusing more on qualitative behavior, the padded size was not consistently adhered to; the graphs of RLVs have been scaled appropriately so that they can be compared regardless of the zeropad utilized; in Chapter III, where the zeropad size directly determines the uncertainty in the peak maxima and therefore the uncertainty in the measured period the zeropad size is consistently 16000x16000. We will return to the relationship between real– and reciprocal–space resolutions in Chapter III when we discuss the direct measurement of interface roughness in QCL materials where zeropadding is actually a poor idea.

The ability to convert from as–imaged (distorted) "nominal" Angstroms to either dimensionless monolayers, in the [001] direction, or <110> lattice constants, in the growth plane, is crucial for accurate measurements of both period fluctuations and alloy order. We therefore pause, here, to thoroughly examine the frequencies present in a given STM image in more detail.

---

[9] Fitting one–dimensional sections through reciprocal–space peaks along the fast–scan–direction (we'll see in the next section that this direction is preferred for our period measurements) to a standard Sinc squared form also reduces the peak uncertainty. It was experimentally established that zeropadding together with a simple search for local maxima yields the same result as fitting the sections; however there is a strong preference for zeropadding and peak finding, which can be easily automated, whereas fitting cannot.

Two series of measurements were made by cropping small–area windows out of a forward–scan image (contracting the *x* piezo, Fig. 2.8). The image is first cropped to remove all whiplash from the bottom (50 Å) and then symmetrically cropped (25 Å) to maintain a square aspect ratio, producing a 700 Å x 700 Å image from a nominal 750 Å x 750 Å scan. In the first series a 100 Å x 100 Å crop window was moved from left to right in steps of 50 Å, keeping the vertical location fixed midway between top and bottom edges of the image (Fig. 2.17, left); in the second series this crop window was moved in the same way from bottom to the keeping the horizontal location fixed midway between the left and right edges of the image (Fig. 2.18, left).

The DFTs (Fig. 2.17 and Fig. 2.18, right) resulting from these small–area crops show none of the smearing found in the full image DFTs, but instead exhibit pronounced ringing emanating from the strong atomic spots (circled in white). This ringing is a natural consequence of finite data; any experimental measurement can be thought of, mathematically, as an infinite expanse gated by a rectangular box that zeros everything outside a chosen frame[10]. This real–space multiplication results in a reciprocal–space convolution between the corresponding Fourier components found in an image and a two–dimensional Sinc function, which is the Fourier transform of a rectangular box [55], the one–dimensional version of which is shown in Fig. 2.16, right. Since the Sinc lobe spacing is inversely proportional to the window size, this ringing falls off faster and is therefore less noticeable with larger windows. Because this Sinc function is oriented

---

[10] In many ways this gating is complementary to zeropadding. By padding our data with zeros we more accurately represent the infinite expanse that has been gated by our measurement.

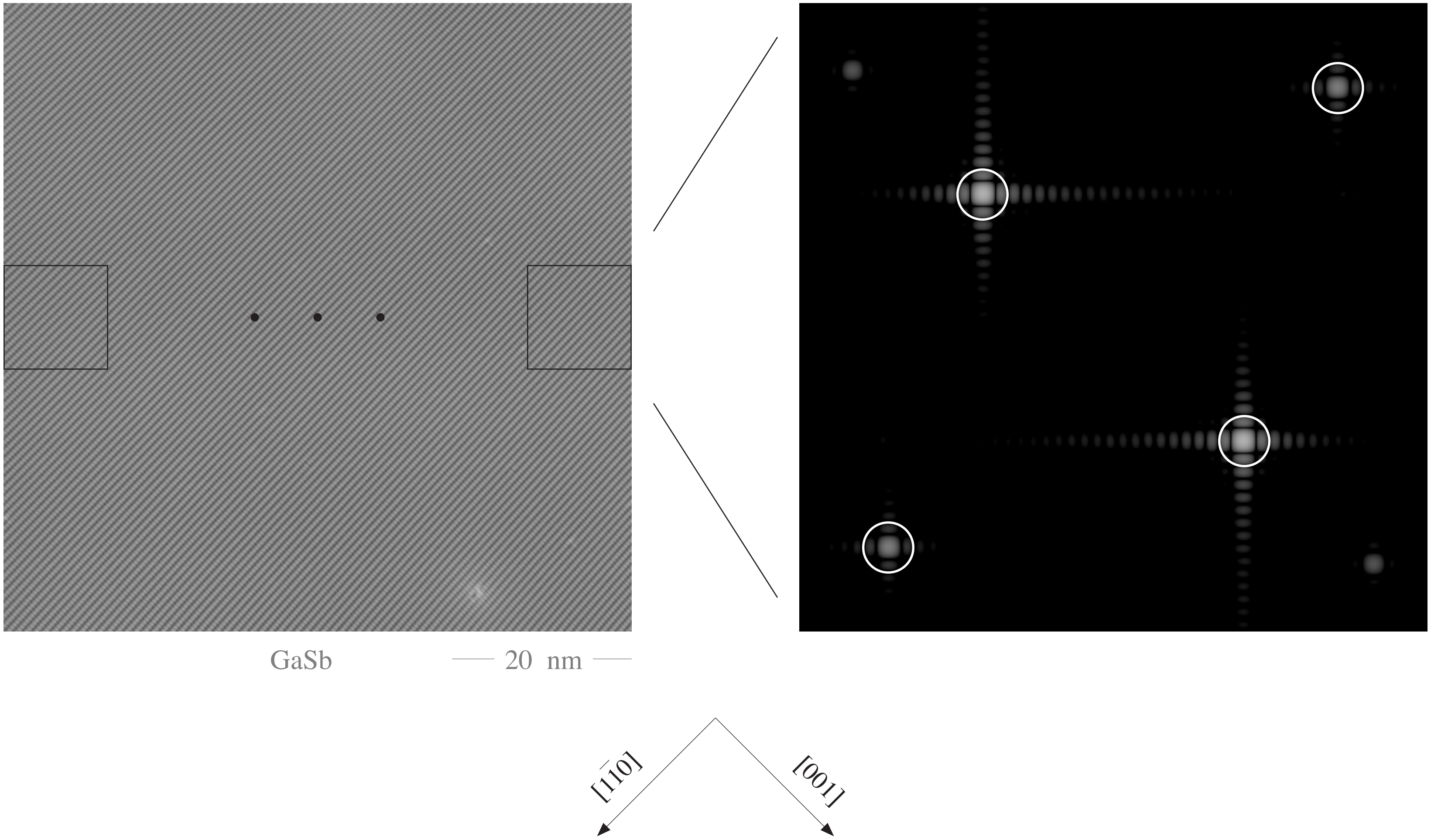


FIGURE 2.17. A series of vertically centered windows (left) were cropped from an image of the GaSb substrate, displacing each 100 Å window by 50 Å to completely cover the image (from left edge to right edge). The location of the reciprocal–lattice vectors (circled in white) within the resulting power spectra (right) varies with crop location. This progression of crop windows allows one to map out the nonlinearity of the *x* piezo.

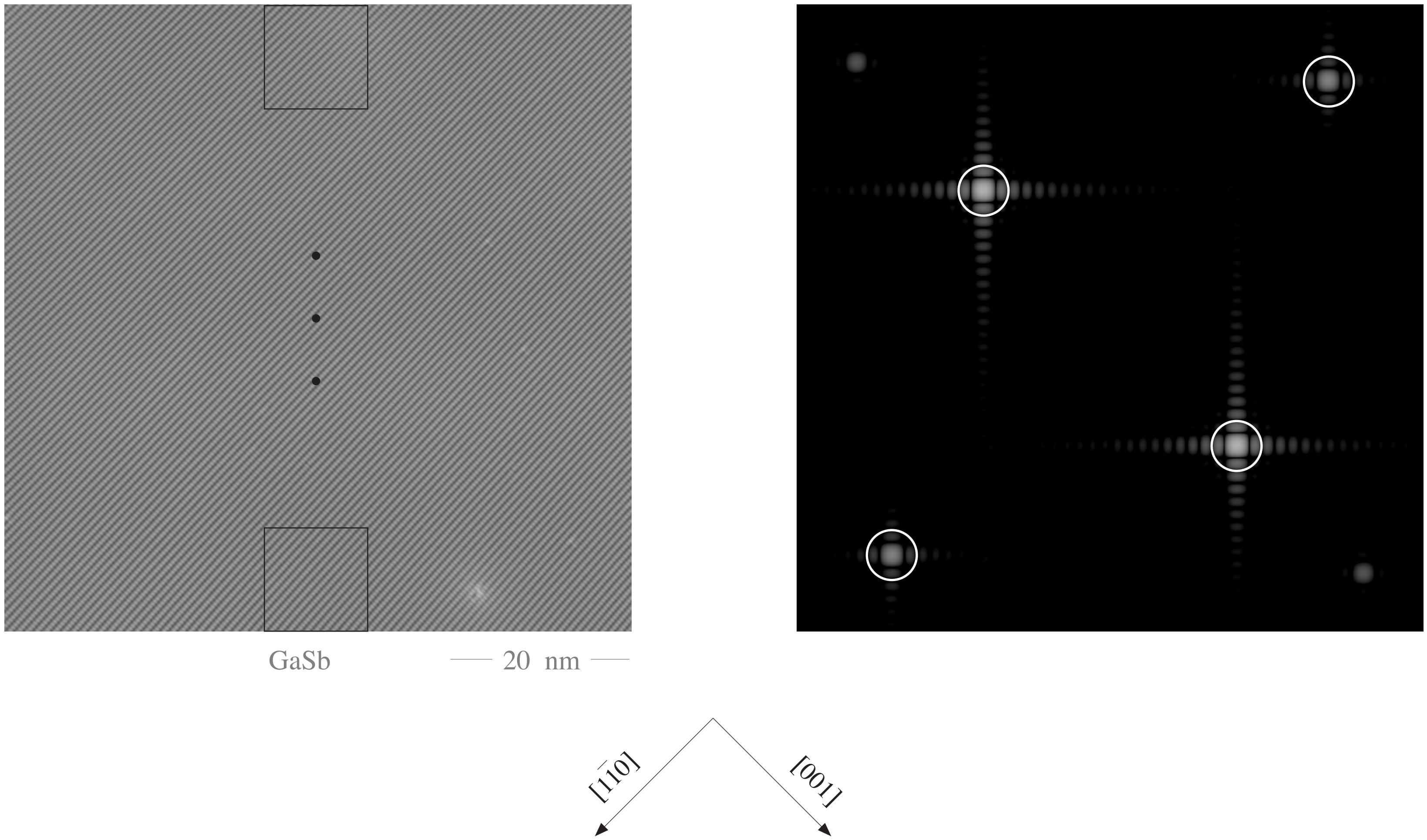


FIGURE 2.18. A series of horizontally centered windows (left), complementary to Fig. 2.17, were cropped from an image of the GaSb substrate, displacing each 100 Å window by 50 Å to cover the image from bottom edge (excluding whiplash) to top edge. The location of the reciprocal–lattice vectors (circled in white) within the resulting power spectra (right) varies with crop location. This progression of crop windows allows one to map out the nonlinearity of the y piezo.

perpendicular to the window, any change in window shape or alignment will produce in a corresponding change to the ringing pattern. The ringing can be abated by softening the edges of an image via a "windowing" function, such as a cosine function centered on the image crop, but at the cost of throwing away unacceptably large amounts of hard–won experimental data. This pattern is present in Fig. 2.15, but a combination of the broadened peak and the Sinc suppression due to the larger image size made it difficult to distinguish the ringing from background noise. It is natural to suspect the larger spot sizes in Fig. 2.17 and Fig. 2.18 simply cover up the doubled peaks observed in Fig. 2.15, but sections through the RLVs in Fig. 2.17 and Fig. 2.18 monotonically increase to respective maxima whereas those in Fig. 2.15 likely would not.

Mapping out the *x*– and *y*–scan coordinates for the [001] (Fig. 2.19) and <110> (Fig. 2.20) reciprocal lattice vectors, we observe that when the crop coordinate is held constant (the slow–scan coordinate in the left graphs or the fast–scan coordinate in the right graphs) there is no variation in the corresponding RLV component. This symmetry between *x* (fast scan) and *y* (slow scan) with respect to small–area, vertical and horizontal crops is at first surprising[11], but indicates that despite all the aforementioned shortcomings of these piezos, they reproducibly yield the same frequencies at the same point in the scan over the relatively small ranges used to acquire individual images.

The crop coordinate that is systematically varied throughout each sequence tells a different story however. As the crop window is swept from the left to the right across the

---

[11] While the two series of small–area measurements are symmetric in many ways with respect to *x* and *y* coordinates, they are not symmetric with respect to the immediate history of the fast–scan (*x*) piezo, since it is constantly being swept through a range of values for each slow–scan (y) piezo increment.

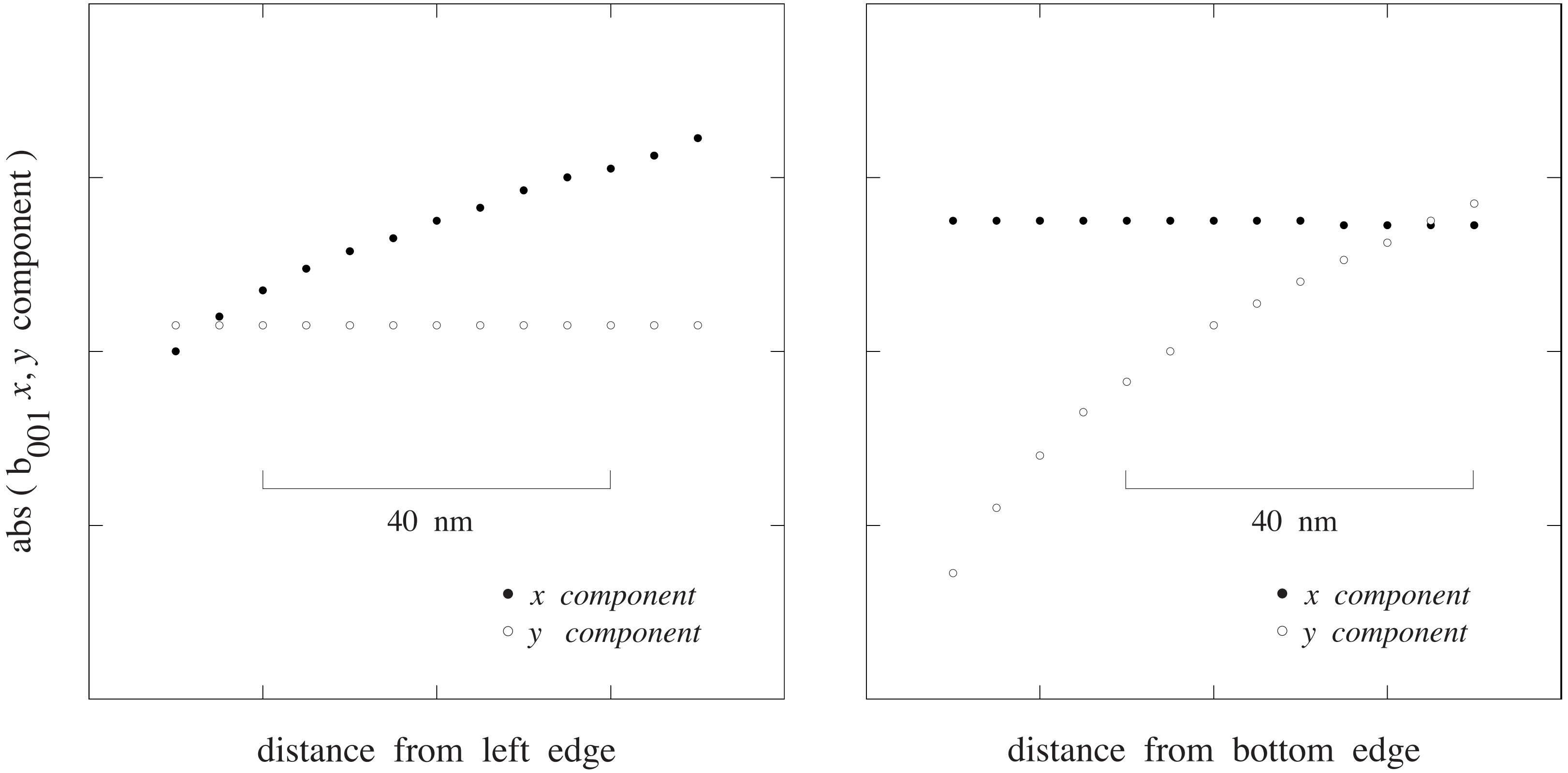


FIGURE 2.19. Coordinates for <001> reciprocal lattice vectors stemming from small crops outlined in Fig. 2.17 (left panel) and Fig. 2.18 (right panel). Full vertical scale (with suppressed zero) for both graphs is set to 20% of the midpoint, so that they may be directly compared. The *y*–coordinate of the crop windows is fixed throughout the series in the left panel, whereas the *x*–component varies from 0 to 600Å. The opposite occurs in the right panel, where the *x*–coordinate of the crop windows remains fixed, and the *y*–component varies from 0 to 600Å. Scale bar represents area encompassed by standard DFT crop window shown later in Fig. 2.21.

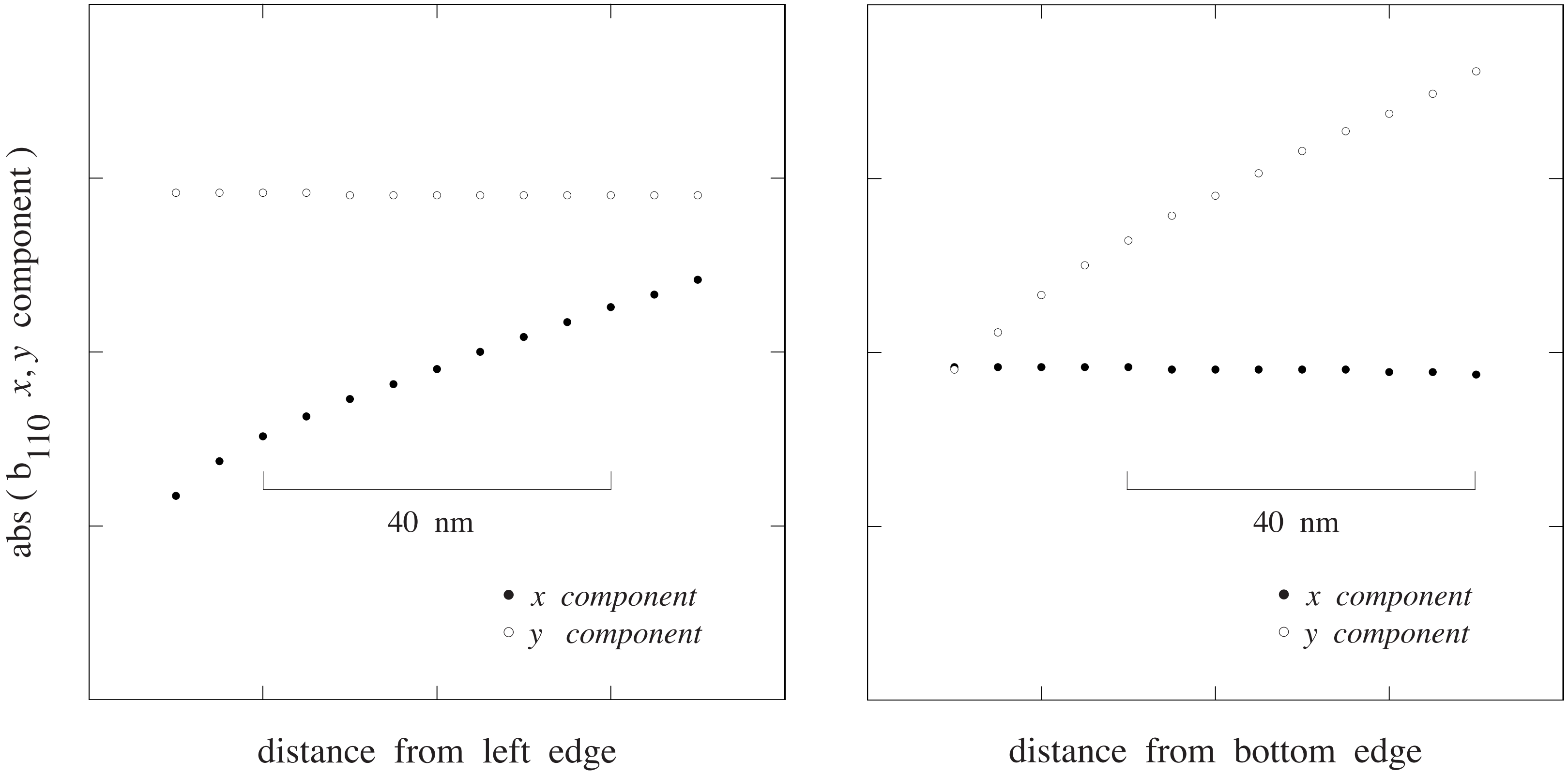


FIGURE 2.20. Coordinates for <110> reciprocal lattice vectors stemming from small crops outlined in Fig. 2.17 (left panel) and Fig. 2.18 (right panel). Full vertical scale (with suppressed zero) for both graphs is set to 20% of the midpoint, so that they may be directly compared. The *y*–coordinate of the crop windows is fixed throughout the series in the left panel, whereas the *x*–component varies from 0 to 600Å. The opposite occurs in the right panel, where the *x*–coordinate of the crop windows remains fixed, and the *y*–component varies from 0 to 600Å. Scale bar represents area encompassed with standard DFT crop window shown later in Fig. 2.21.

image (Fig. 2.17, left) the *x*–component (fast–scan) of both RLVs increases; likewise as the crop window is swept from the bottom to the top of the image (Fig. 2.18, left) the *y*–component (slow–scan) of both RLVs increases. This shift to higher spatial frequencies means that the atoms appear closer in real space, but since the atoms in the underlying substrate are, in reality, equidistant, the piezos must be physically contracting more for each voltage step as the scan progresses.

The slope of the RLV *x*–component versus distance from the left side of the reverse–scan image (fast–scan expanding) is opposite that in Fig. 2.19 (left) and Fig. 2.20 (left). Said another way, due to the hysteresis illustrated in Fig. 2.12, the *x*–component of the RLV increases as the crop window is moved from the right to the left across the reverse image, mirroring the left to right increase in the forward image. This increase from right to left in the reverse image is nevertheless, consistent with the piezo step size increasing for each additional voltage step as the scan progresses, since the piezo is expanding in going from the right side of the image to the left side.

The image size can be increased from 100 Å x 100 Å to 400 Å x 400 Å with little consequence to the quality of the FFT all the while suppressing ringing of the Sinc (and therefore the width of the RLVs) due to the increase in window size. For DFT windows larger than 400 Å x 400 Å the <110> RLV begins to bifurcate, and not long after, the [001] RLV bifurcates as well as observed in Fig. 2.15. The best crop location, empirically determined, is illustrated in Fig. 2.21, and marked by the scale bar positions in Fig. 2.19 and Fig. 2.20. We instinctually chose the location of the standard crop window in a region that happened to coincide with nearly linear spatial "chirp" in the

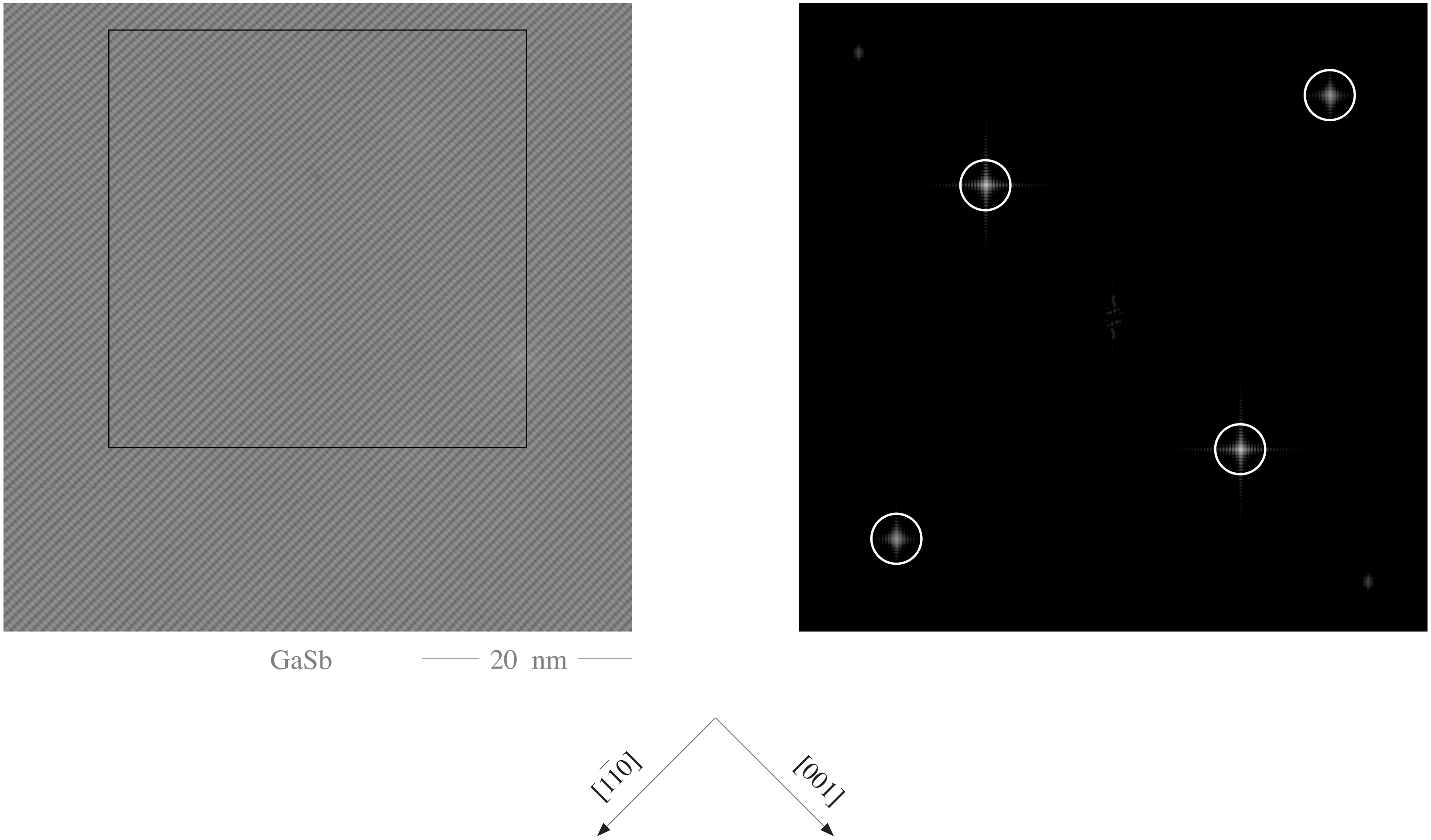


FIGURE 2.21. Central 40 nm x 40 nm crop from atomic–resolution STM image (left) and corresponding reciprocal–space map (right). Highlighted spots denote the [001] and <110> reciprocal–lattice vectors defining the two–dimensional surface mesh. As in Fig. 2.17 and Fig. 2.18, the power in each reciprocal–lattice vector is concentrated in a single spot circumventing the worst effects of piezo nonlinearity observed in Fig. 2.15. Growth direction is from top left to bottom right.

RLVs. While not mathematically addressed here, it seems plausible that the RLVs obtained this way are averages over this linear chirp within our standard crop window.

Survey Protocol

To collect the ensembles needed to drive down statistical uncertainties in our measurements, we need to construct physically appropriate image surveys. Studying interface roughness in QCL materials [39] required the development and adoption of lateral survey protocols that follow a small subset of superlattice repeats for up to a micron in the [110] or [1-10] directions, thereby sampling a fixed point in time during the growth history. Additional time slices may then be similarly reconstructed through successive vertical displacement of these lateral surveys along the [001] growth direction.

Given the need for lateral surveys, one must still choose which of two possible navigation routes (Fig. 2.22, left) – affected by corresponding offsets applied to the successive scan frames – to adopt. Simultaneous expansions of both fast– $(-x_{offset})$ and slow–scan piezos $(-y_{offset})$ illustrated in Fig. 2.11 (right) and Fig. 2.22 (left), effects a [1–10] translation of the scan–frame origin within the (–1–10) cleavage plane illustrated. Simultaneous contraction of both piezos $(+x_{offset}$ and $+y_{offset})$ on the other hand effects a [–110] translation of the scan–frame origin.

When we discussed piezo creep, which causes whiplash at the bottom of the images, it was pointed out that this creep causes additional problems whenever an image frame is offset parallel to the positive y scan direction. The problem, while subtle in real space, becomes frustratingly apparent in reciprocal space (Fig. 2.22, right). There are

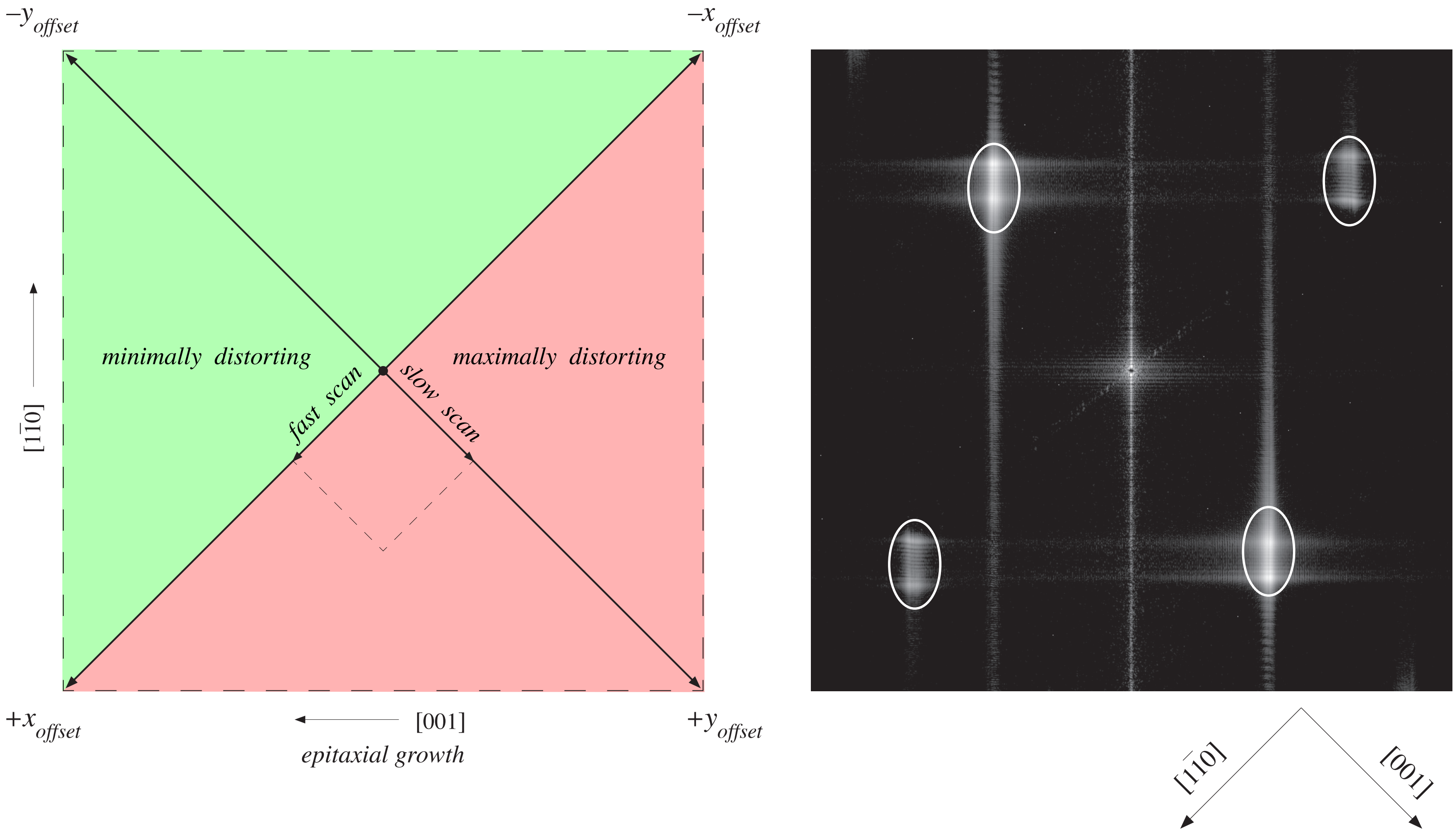


FIGURE 2.22. STM scan frame located at the origin indicates fast– and slow–scan directions oriented along x– and y–piezo axes, respectively. Minimally distorting zone corresponds to navigation routes that avoid positive increments of the y–piezo offset. An example power spectra of the distortion caused by positive y–piezo offsets is shown to the right with doubled reciprocal lattice vectors highlighted. Schematic reprinted with permission from [39].

two *very* different RLV *y* components present in the DFT when the frame offsets contract (i.e. parallel) the slow–scan piezo. This effectively restricts the lateral survey direction to a small choice[12], dictated by offsets that simultaneously expand x (fast–scan) and y (slow–scan) piezos (Fig. 2.22, left).

Enforcing the standard 40 nm x 40 nm window size for DFTs we may map the RLV components throughout an "allowed" lateral survey as shown in Fig. 2.23. Surprisingly, the two components are not symmetric; the *x* (fast–scan) component is very stable throughout the course of the survey while the *y* (slow–scan) component drifts by as much as 4% of it's value. This asymmetry can be reconciled by considering the time scales involved with each piezo; the fast–scan piezo is actuated every few seconds while the slow–scan piezo changes over the course of ~ 15 minutes; this longer time scale makes it proportionally more sensitive to $1/f$ noise such as thermal drift of the sample. For this reason we choose to rely on the fast–scan component in Chapter III.

The [001] and <110> lattice constants averaged over the 40 nm x 40 nm standard crop (Fig. 2.24) can be calculated from the RLV components in Fig. 2.23; this image–by–image conversion map in Figs. 2.23 and 2.24 is calculated for each and every survey and provides the ruler we ultimately adopt to normalize our STM measurements, either directly (correlation lengths in Chapters III and IV) or indirectly via the RLV (period measurements in Chapter III).

---

[12] The extreme doubling also restricts the vertical survey direction if minimally–distorted images and DFTs are sought. It is worth calling attention to the serendipitous alignment of the [001] direction with a minimally distorting survey route, given the modified sample mounting in Fig. 2.8, this was not true of the inherited mounting so that vertical surveys "naturally" progressing from repeat 1 through 100 were heavily distorted.

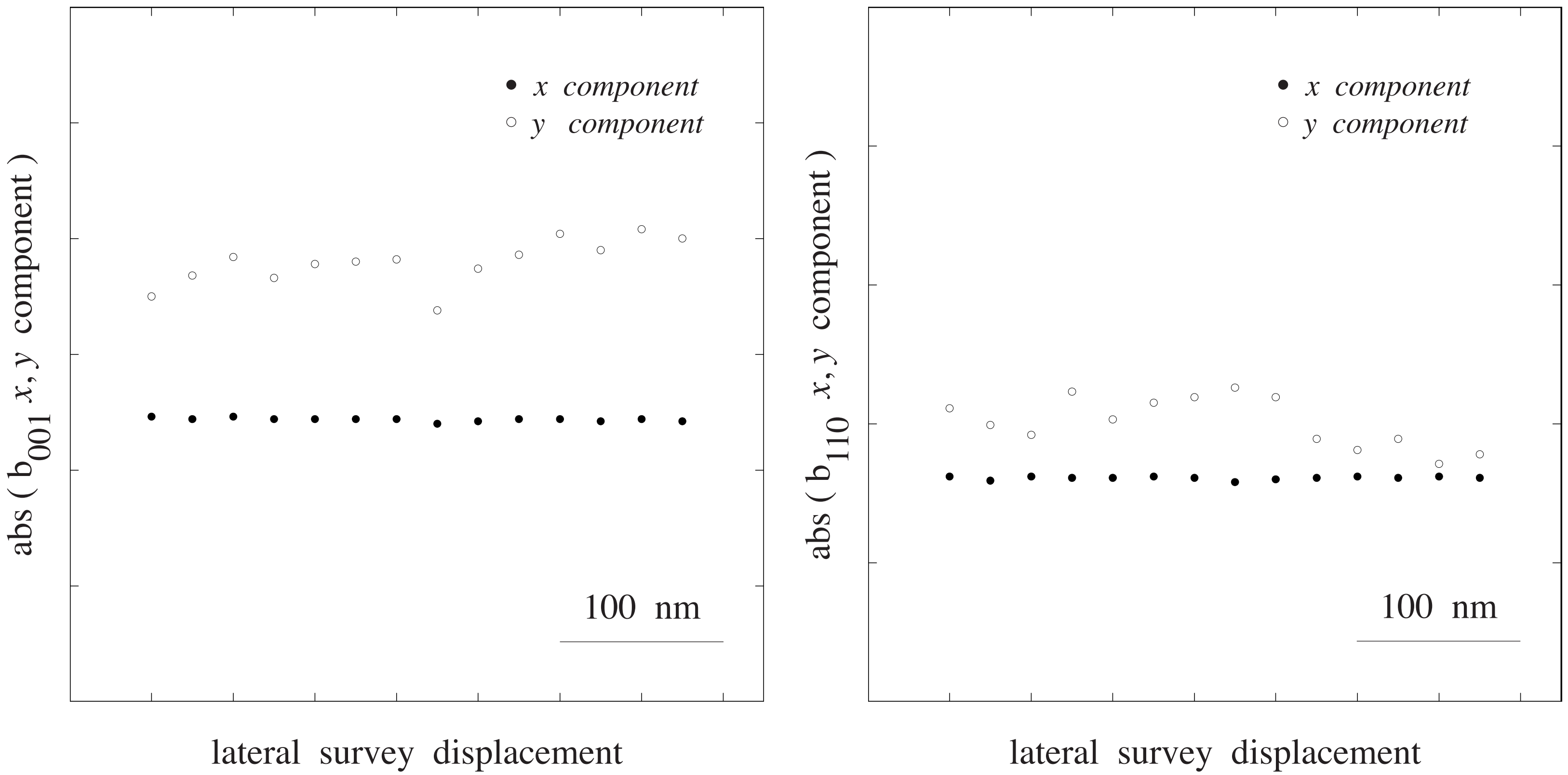


FIGURE 2.23. Reciprocal lattice vectors vary from image–to–image along a typical lateral survey. The fast scan (*x*) component is stable throughout a lateral survey, but the slow scan (*y*) component varies by as much as 4% (shown here for a survey with 2% variation). Full vertical scale (with suppressed zero) for both graphs is set to 20% of the midpoint so that they may be directly compared.

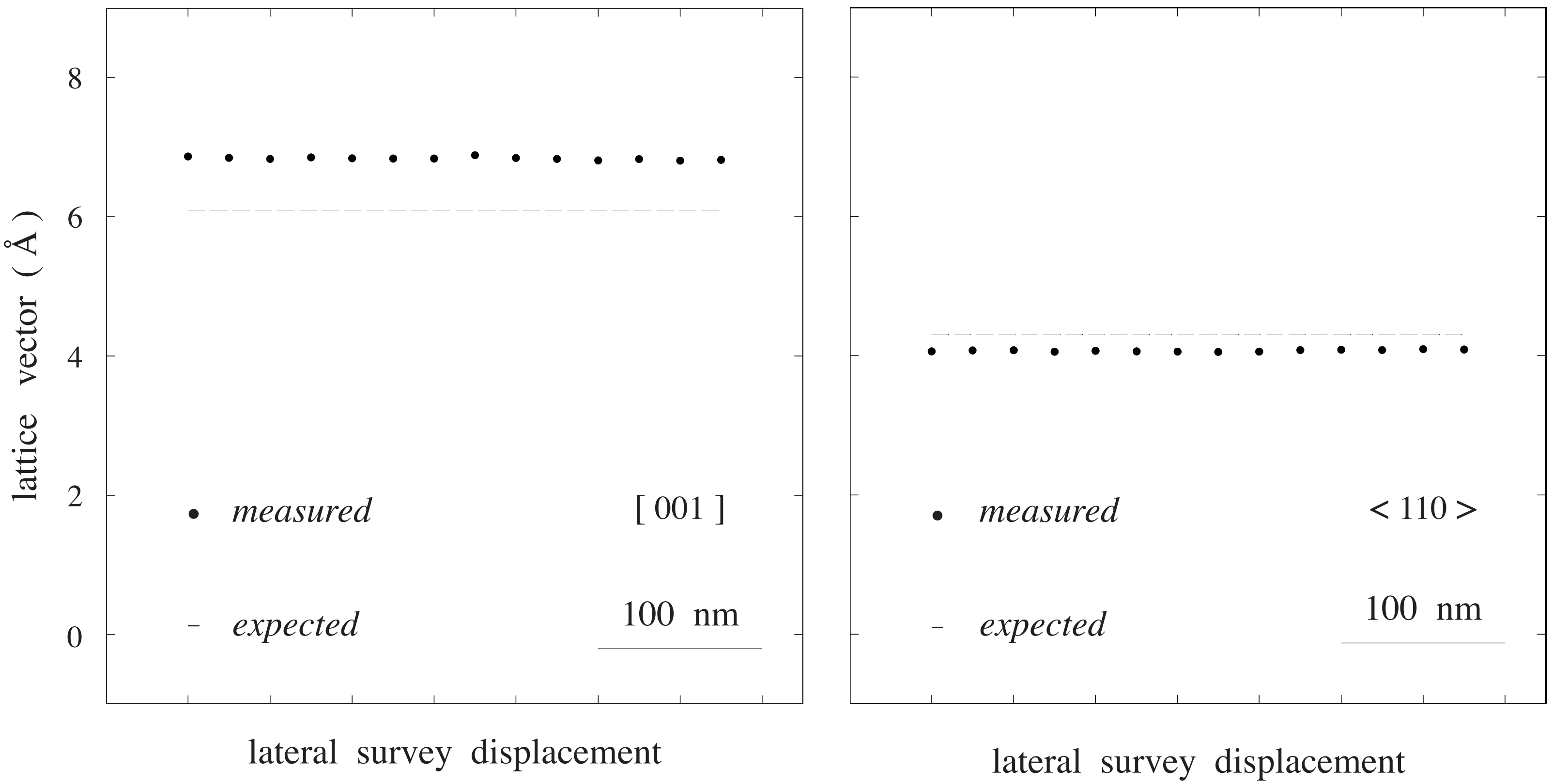


FIGURE 2.24. Real–space lattice vectors calculated from image–by–image reciprocal lattice vector components in Fig. 2.23 taken in quadrature. The [001] lattice constant is ~ 12 % larger than expected, whereas the <110> lattice constant is ~ 5 % smaller than expected based on nominal *x*– and *y*–piezo calibrations.

Every lateral survey follows a chosen subset of superlattice repeats. These repeats are then systematically varied via vertical surveys that extend the entire length of the superlattice growth in the [001] direction. Representative examples, along with arrows illustrating the allowed survey directions, are shown in Fig. 2.25. We will develop additional methods for sidestepping the distortion present in each of these survey images in the next two chapters as we seek to measure the local superlattice period in Chapter III, and the growth–plane correlations between isovalent impurity atoms in Chapter IV.

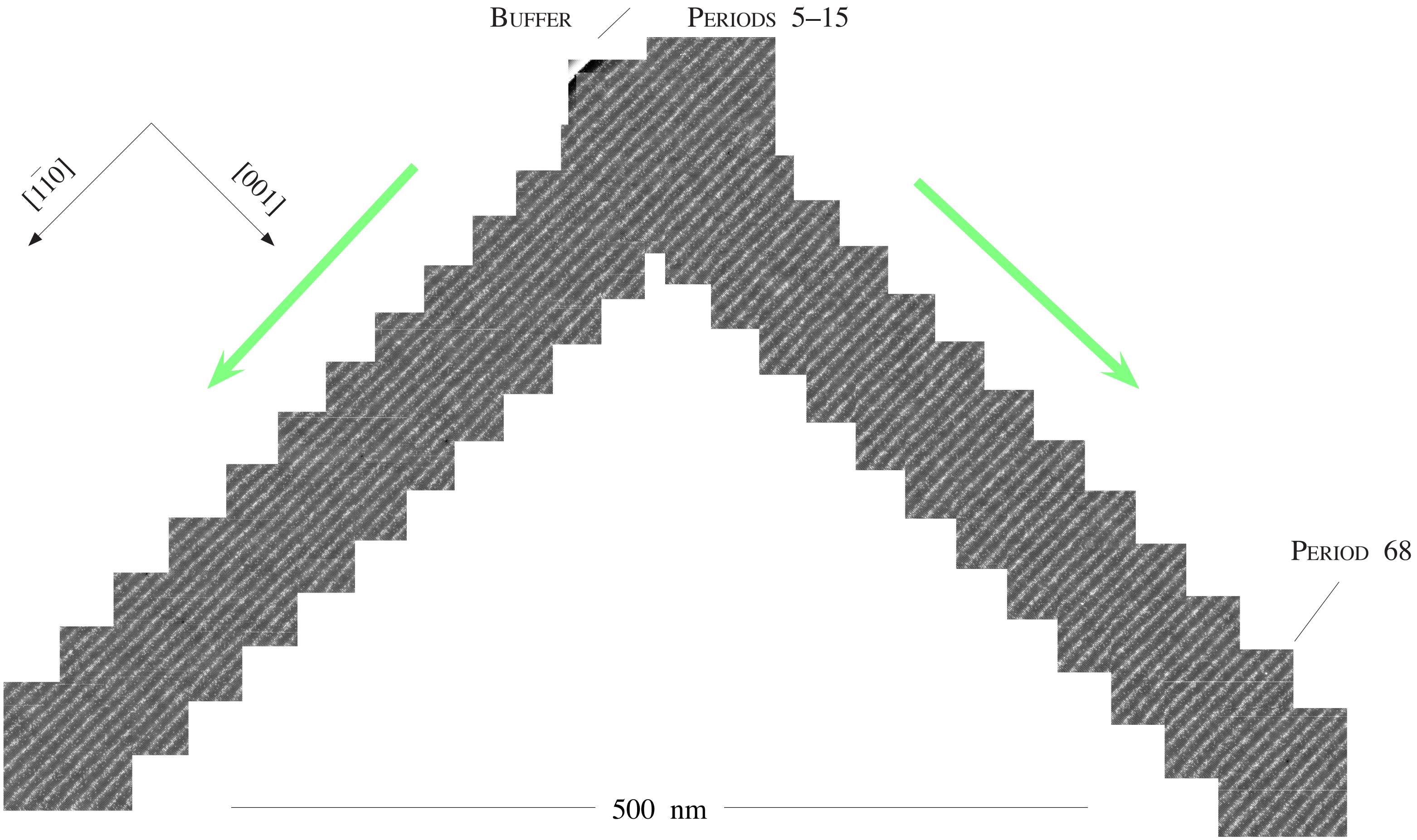


FIGURE 2.25. The precise navigational control afforded by STM facilitates micron–length lateral samplings of deliberately chosen subsets of superlattice repeats (shown here for repeats 5–15) – indexed by vertical surveys extending throughout the structure on either (–1–10) or (1-10) cross sections – providing superlattice repeat resolved statistical ensembles for subsequent analysis. Bright layers are InAsSb. Green arrows illustrate allowed survey directions from Fig. 2.22; the *y*–scan offset is decreasing in both cases, *x*–scan offset selects between lateral (decreasing *x*–offset) and vertical (increasing *x*–offset) surveys. Growth direction is from top left to bottom right.

CHAPTER III

LONG–RANGE STRUCTURAL DISORDER

Introduction

It might seem natural to begin our discussion of the differences between *as–grown* and intended structures with short–range alloy order since our data are comprised of *atomic*–resolution STM images, but long–range disorder provides a more fitting introduction to correlated statistics. We begin that discussion with a consideration of the long–range disorder manifested in superlattice period measurements.

This chapter, like the next, relies on numerical calculations to build intuition and inform analyses that might then be used with experimental measurements to extract physically relevant insights. Before embarking on those calculations, we describe how the [001]–convolved reciprocal–space satellite peaks obtained from discrete Fourier transforms (DFTs) of individual STM images can be used to measure the period variations in InAs / InAsSb superlattices. These measurements can be localized to within ~ 5 superlattice repeats in the [001] growth direction, as well as to within ~ 40 nm along <110> directions in the growth plane, and either subsequently pooled (over lateral and/or vertical survey ensembles) to compare with global measurements of the average superlattice period, or treated independently to analyze the image–by–image period fluctuations within a given set of repeats.

We show by way of semi–quantitative arguments that these localized fluctuations are inconsistent with the naïve assumption of laterally–uncorrelated interfaces for the system studied here. We then turn to numerical experiments that explore how correlated interface roughness influences the local period – even when the correlation lengths involved are small on the scale of our measurement window – and show how the magnitude of image–to–image period fluctuations depend on the parent parameters in a stochastic description of interface roughness. Finally, we use our numerical experiments to fully explore the dependence of period fluctuations on interface roughness parameters.

Further insight is offered by a MOCVD–grown InAlAs / InGaAs superlattice, whose strong layer contrast makes it amenable to direct measurement of the stochastic parameters used to describe correlated interface roughness in our numerical experiments. These directly measured roughness values confirm that the local period fluctuations in InAlAs / InGaAs originate with interface roughness. We then use this knowledge to set limits on the corresponding roughness parameters for our InAs / InAsSb material system.

The idea to use DFTs to measure superlattice periods was developed by Dr. Federico Lopez to contrast layer uniformity of InAlAs / InGaAs superlattices grown by MOCVD and by MBE. I am indebted to Dr. Lopez as well as Dr. Kara Kanedy for the experimental benchmarks used in this chapter.

Analysis of Experimental Data

Chapter II concluded with the introduction of device–scale surveys (Fig. 3.1) composed of atomic–resolution STM images (Fig. 3.2). As discussed at length in Chapter II, these images are inevitably distorted, and that distortion undermines straightforward attempts to measure a superlattice period in real space. We turn instead,

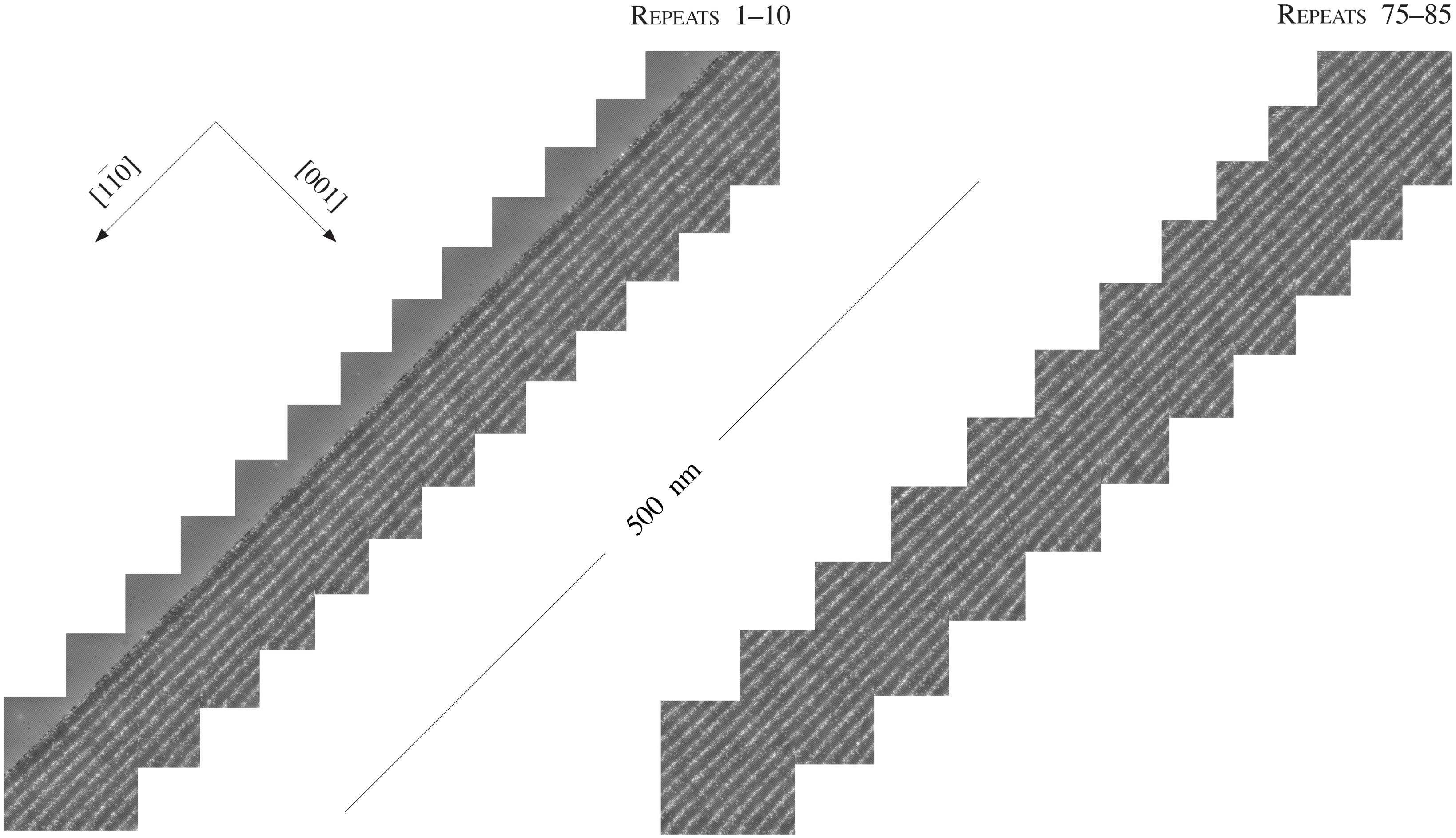


FIGURE 3.1. Device–scale surveys across a (–1–10) cleavage plane through an InAs/InAsSb superlattice showing early and late superlattice repeats. Bright layers are InAsSb, [001] growth direction is from top–left to bottom–right. Adapted from [20].

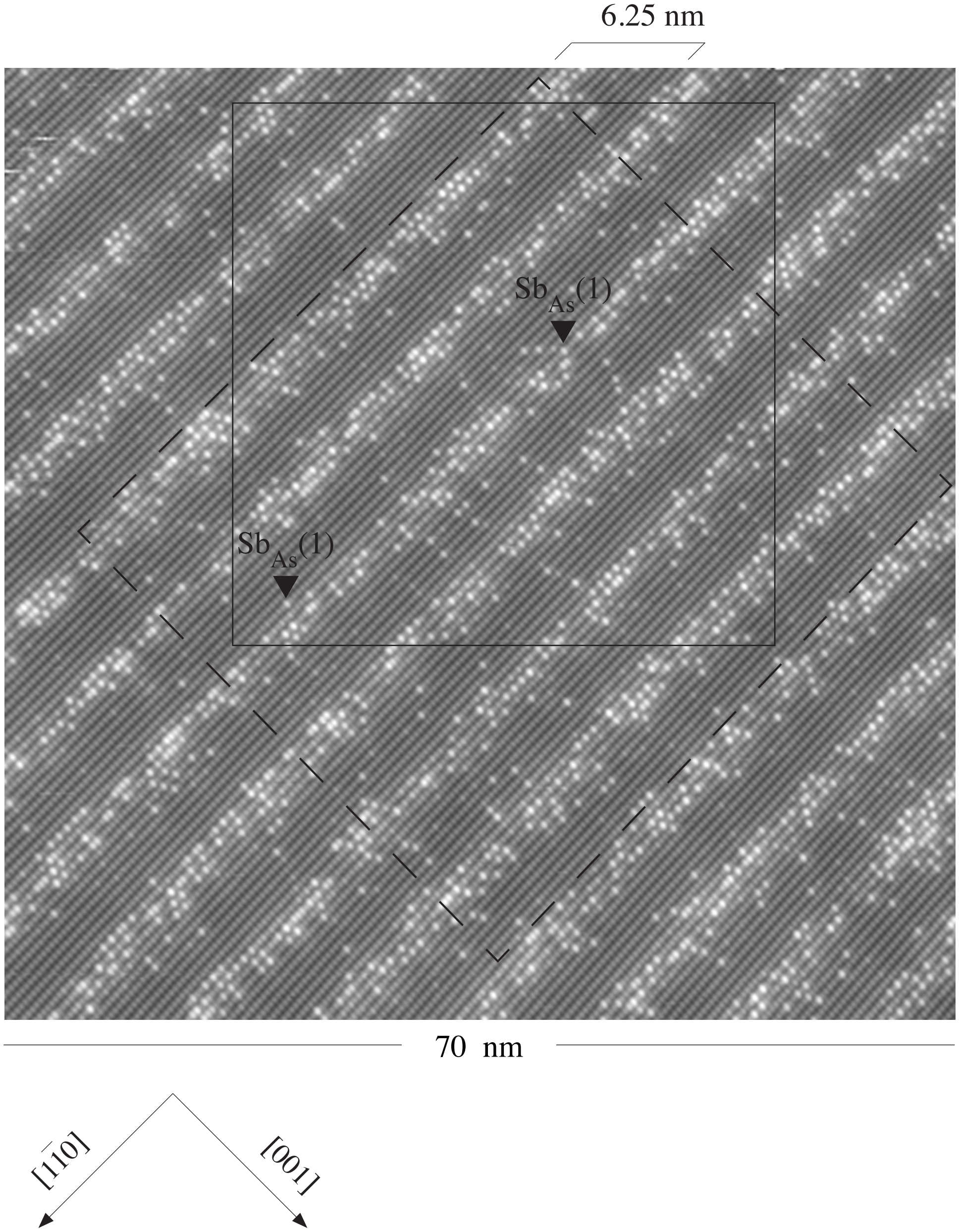


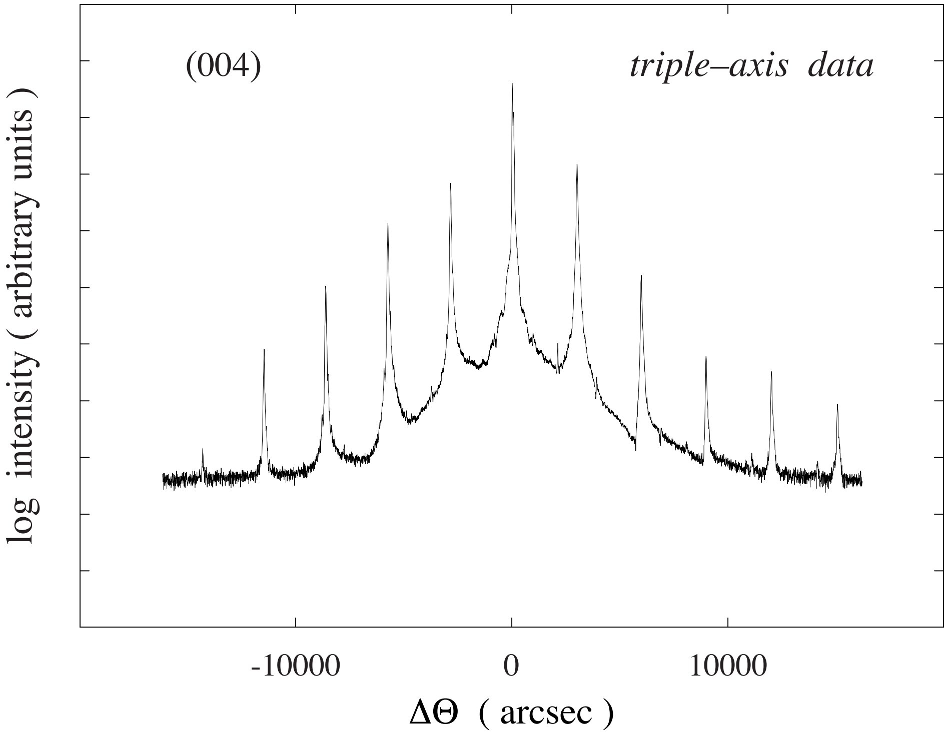
FIGURE 3.2. Individual atomic–resolution image with antimony–for–arsenic substitutions indicated by carets and approximate thickness of a single repeat annotated above. Solid box identifies standard DFT crop window illustrated in Fig. 2.22; dashed box indicates a representative counting window used to compile the antimony fraction in successive [001] monolayers. Adapted from [20].

to a measurement approach similar in spirit to high–resolution x–ray diffraction (HRXRD), where Bragg's Law yields a superlattice period of 20.62 ± 0.01 ML (Fig. 3.3). Just as the x–ray spectrum has superlattice peaks convolved with the (004) substrate reflection, the DFT (Fig. 3.4) calculated from the standard image crop corresponding to the solid box in Fig. 3.2 exhibits superlattice peaks (encircled in red) convolved with the [001] reciprocal lattice vector[1].

An [001] section through these data (Fig. 3.5, top) looks very similar to a HRXRD rocking curve (Fig. 3.3, top) and can be analyzed in like manner, save one very important difference. Recall from Chapter II that the DFT reciprocal–lattice vectors (RLVs) vary from image–to–image; since the superlattice peaks are convolved with the [001] RLV they must vary in exactly the same way. Normalizing the superlattice satellite peaks to the [001] RLV image–by–image therefore provides a measurement denominated in units of the atomic mesh that is potentially insensitive to image distortion. The x (fast–scan) component of each superlattice peak normalized to the x component of the RLV is a linear function of satellite order whose slope is inversely proportional to the superlattice period (Fig. 3.5, bottom). The slow–scan (y) axis is proportionally more susceptible to $1/f$ noise than the fast–scan (x) axis, as described in Chapter II, so while the slow–scan component yields perfectly serviceable period measurements, we rely exclusively on the fast–scan component in what follows; in so doing we've implicitly assumed our superlattice is perfectly oriented perpendicular to

---

[1] X–ray reflections from zinc blende structures require the Miller indices sum to 4*n and all three indices to be either even or odd; the (004) plane therefore is the first allowed reflection in these materials [58]. This does not apply to the two–dimensional STM image, which is a rectangular lattice.



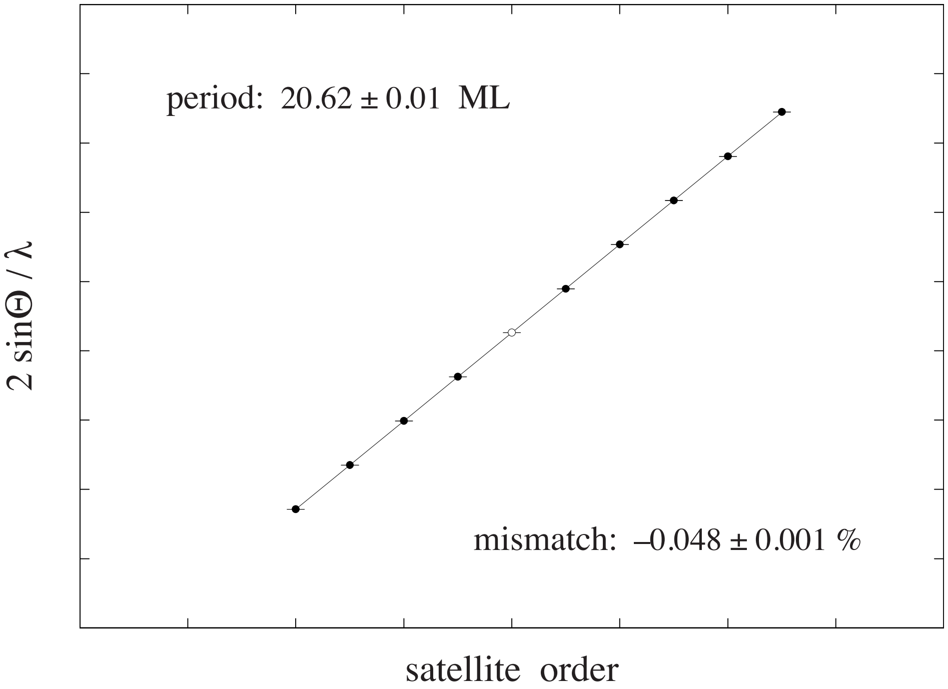


FIGURE 3.3. Triple–axis HRXRD measurement about the (004) reflection (top) and superlattice period calculated using Bragg's law (bottom). Adapted from [20].

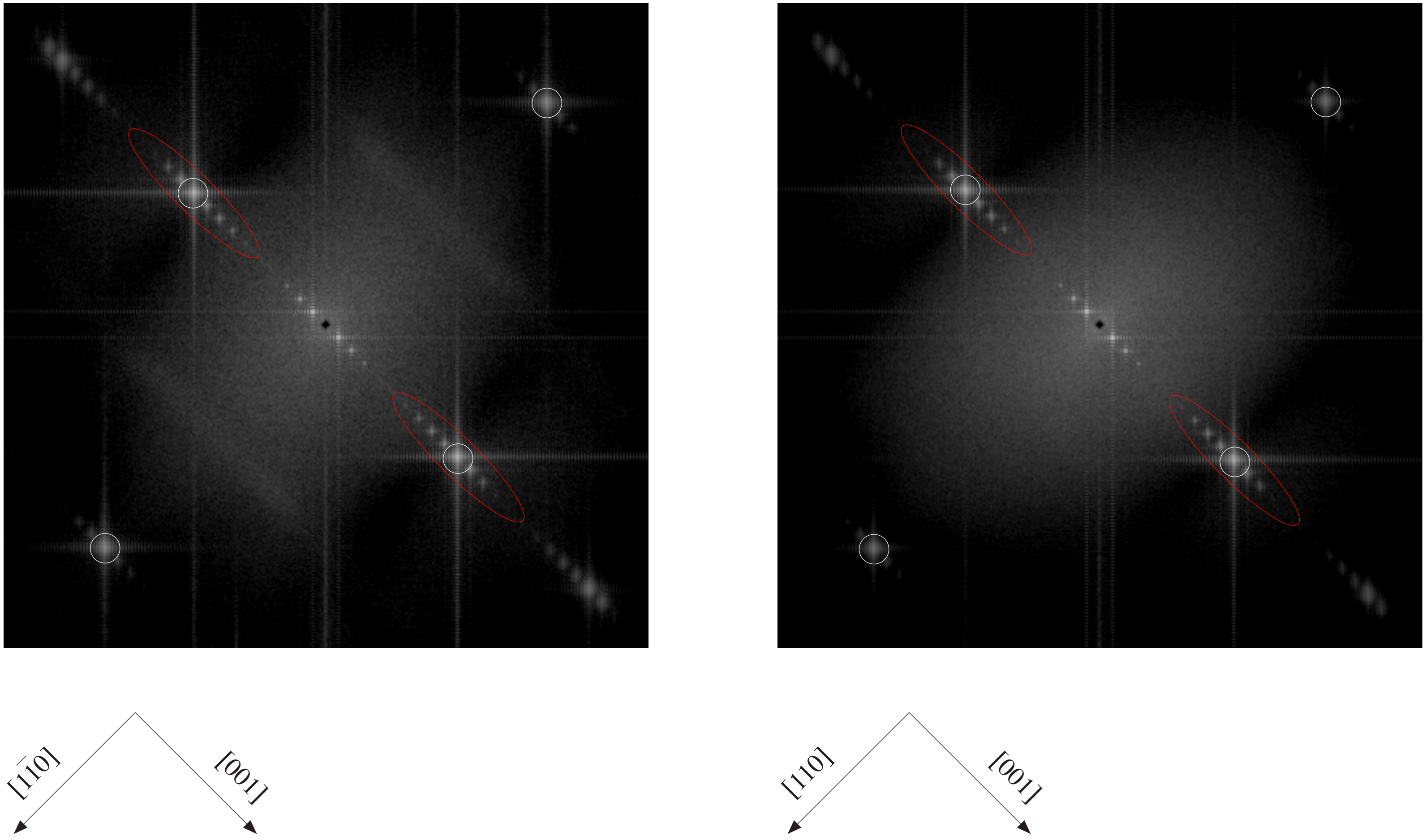


FIGURE 3.4. Two–dimensional reciprocal space maps from 40 nm x 40 nm crops in (–1–10) and (1–10) cross section (left and right respectively). Reciprocal lattice vectors are circled in white, [001]–convolved superlattice satellite peaks are circled in red. Satellite peaks are similar in both cross sections. Adapted from [20].

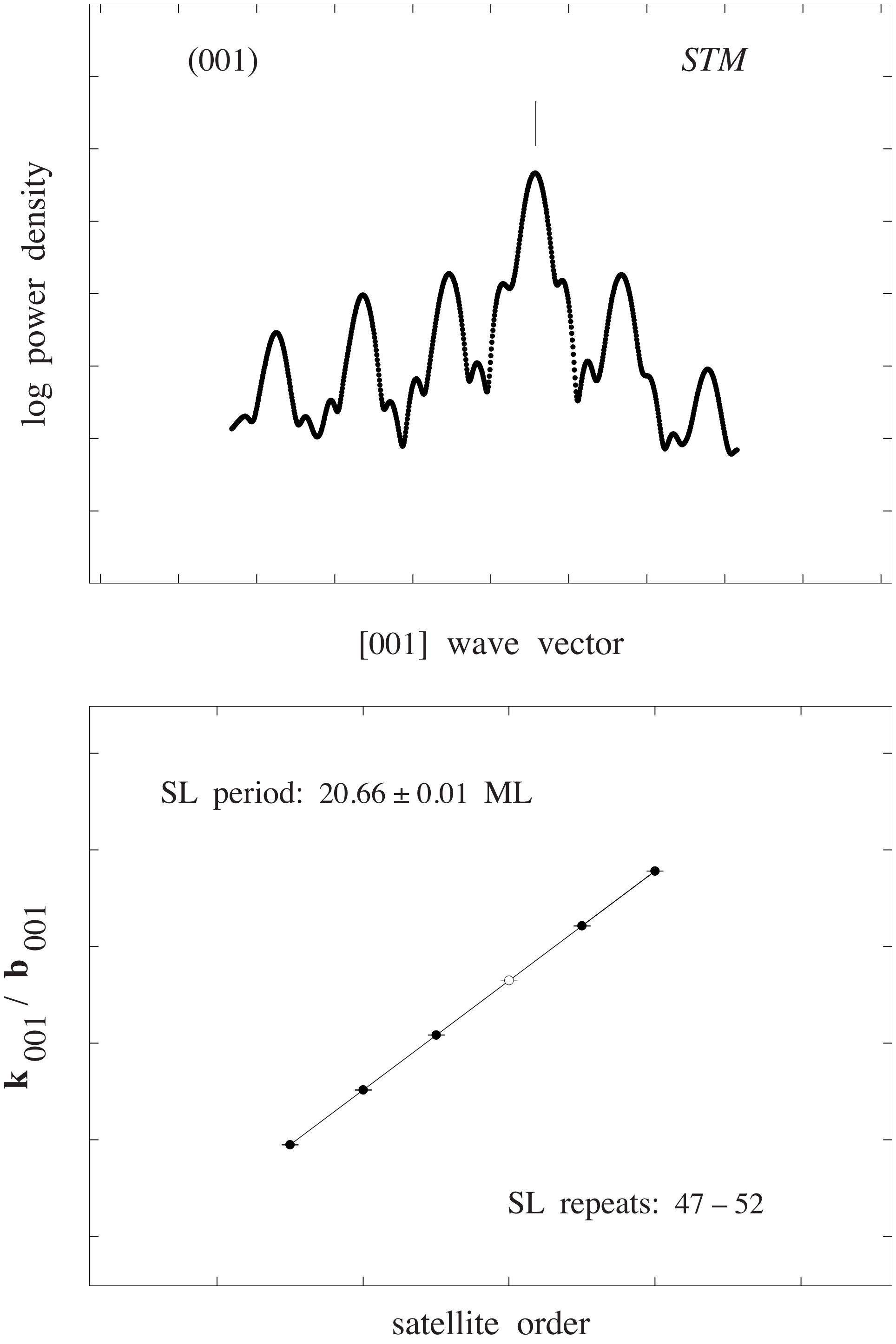


FIGURE 3.5. Section through survey–averaged reciprocal–space map (top) in Fig. 3.4 emphasizing superlattice satellites convolved with the [001] reciprocal lattice vector (tick). The x–coordinate of each satellite is normalized image–by–image to the x–coordinate of the reciprocal lattice vector and then averaged (bottom); the resulting slope versus satellite order is inversely proportional to superlattice period. Adapted from [20].

the [001] growth direction. This is not always strictly true, since even high–quality substrates may have miscut angles on the order of 0.1°. Our particular sample was grown on a substrate whose miscut was specified as ≤ 0.03°, potentially introducing a systematic error of less than 0.02%; this error is roughly eight times smaller than the single–image Bragg errors[2].

We see from Fig. 3.6 that as we survey along a <110> direction, the periods extracted from interleaved forward and reverse (contraction and expansion of the *x* piezo) images track one another. As described in Chapter II, these measurements are taken on opposite sides of a piezo–hysteresis curve and are, in that sense, independent; they may thus be averaged to reduce the experimental period uncertainty. The close agreement between forward– and reverse–scan data in Fig. 3.6 explicitly demonstrates that normalizing to the image–by–image RLVs effectively circumvents the piezo non–idealities detailed in Chapter II. Interestingly, each set of measurements fluctuates well outside experimental error and we return to consider this finding more thoughtfully in a moment. First, however, we ask if the periods measured this way are as accurate as they are reproducible.

---

[2] It's important here to clarify the hierarchy of DFT–related errors in this discussion. Each single–image satellite peak as well as its associated reciprocal–lattice vector has an uncertainty of ± 1 reciprocal–space pixels. The resulting single–image Bragg errors stem from least–squares fits to the normalized satellite peaks, given these uncertainties. Survey–pooled errors (as in Fig. 3.5) are the corresponding single–image uncertainties reduced by the square root of the number of images.

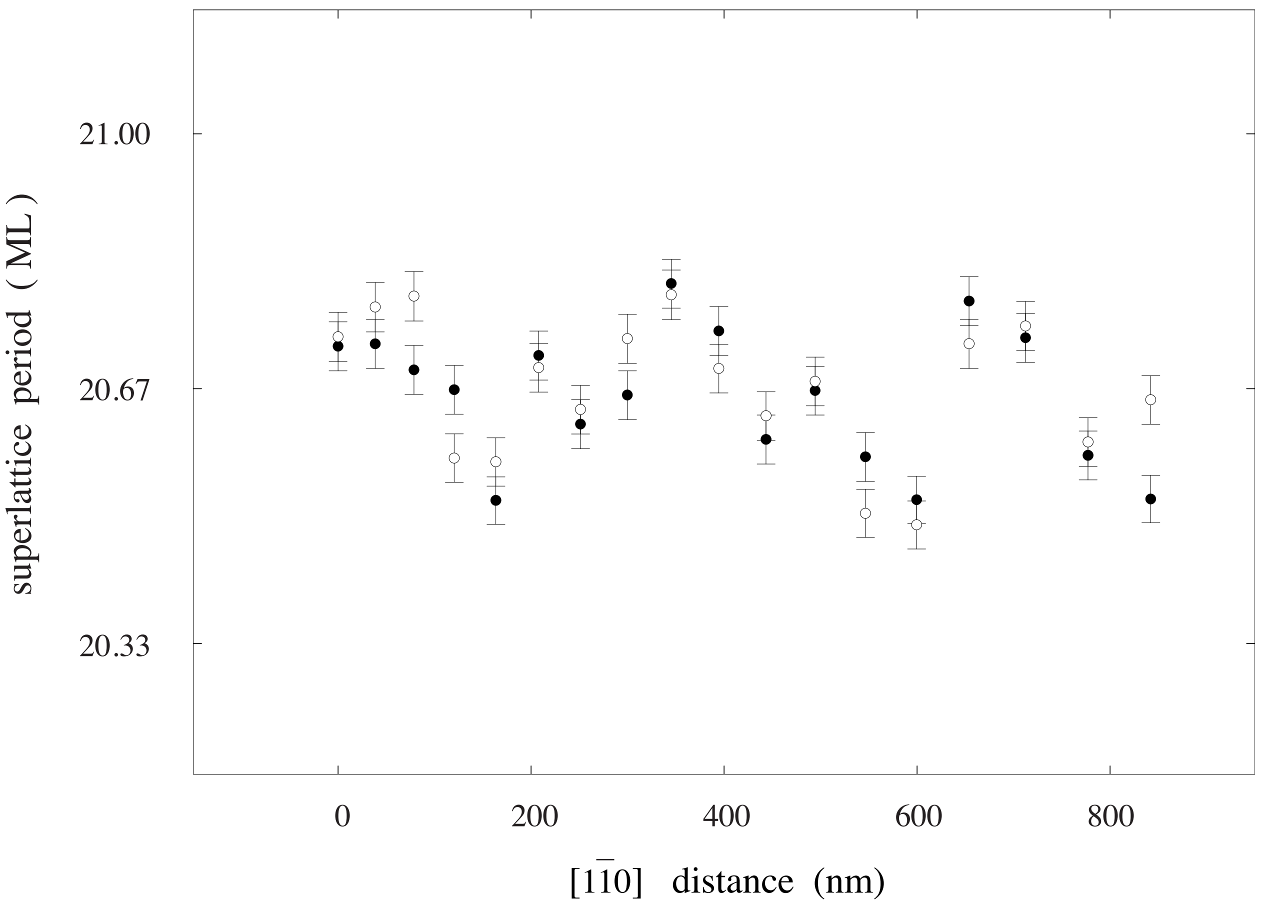


FIGURE 3.6. Image–by–image superlattice periods from fast–scan “Bragg” plots similar to the one in Fig. 3.5, illustrating the close agreement between forward– (x piezo contraction) and reverse–scan (x piezo expansion) images. Laterally the surveys evidence fluctuations measurably exceeding those reflected in the experimental “Bragg” errors. Reprinted with permission from [20].

To answer this question, we compare the laterally–averaged periods[3] from six[4] disjoint, vertical subsets of superlattice repeats occurring early on, midway through, and late in the growth. These periods, summarized together with their attending Bragg errors in Fig. 3.7, are clustered near 20.67 ML and 20.5 ML. We see a vertical drift in (–1–10) periods that is quantitatively mirrored in a similar set of (1–10) surveyed repeats taken over a different location on the wafer (~ 10 mm apart). Careful re–examination of the x–ray data guided by these STM results reveals each superlattice satellite can be modeled as a sum of three overlapping Gaussian components whose minority side–bands are consistent with the periods 20.53 and 20.69 ± 0.01 ML respectively (Fig. 3.7). The comparison with these x–ray side bands is made even more convincing by the following observation: since there is a 1:1 correspondence between the coherently–strained lattice constant and the survey–averaged alloy composition [20], our STM period measurements can be converted from monolayers to absolute Ångstroms. So doing (Fig. 3.8), we find that upper and lower clusters agree with their matching side band to better than 0.2% in absolute terms.

It is also noteworthy that the dominant 20.62 ML component in the x–ray spectrum (Fig. 3.3) is nowhere to be found across six STM surveys. Dynamical x–ray simulations suggest this “period” could easily arise from a convolution of the two STM side bands, identified here, with a previously described vertical evolution in substrate

---

[3] Averaging over all image–by–image measurements (Fig. 3.6) to obtain a representative period and Bragg error for each surveyed set of repeats.

[4] A seventh survey containing reduced statistics was not included in these averages, it nevertheless agrees well with the trends observed in the other six surveys.

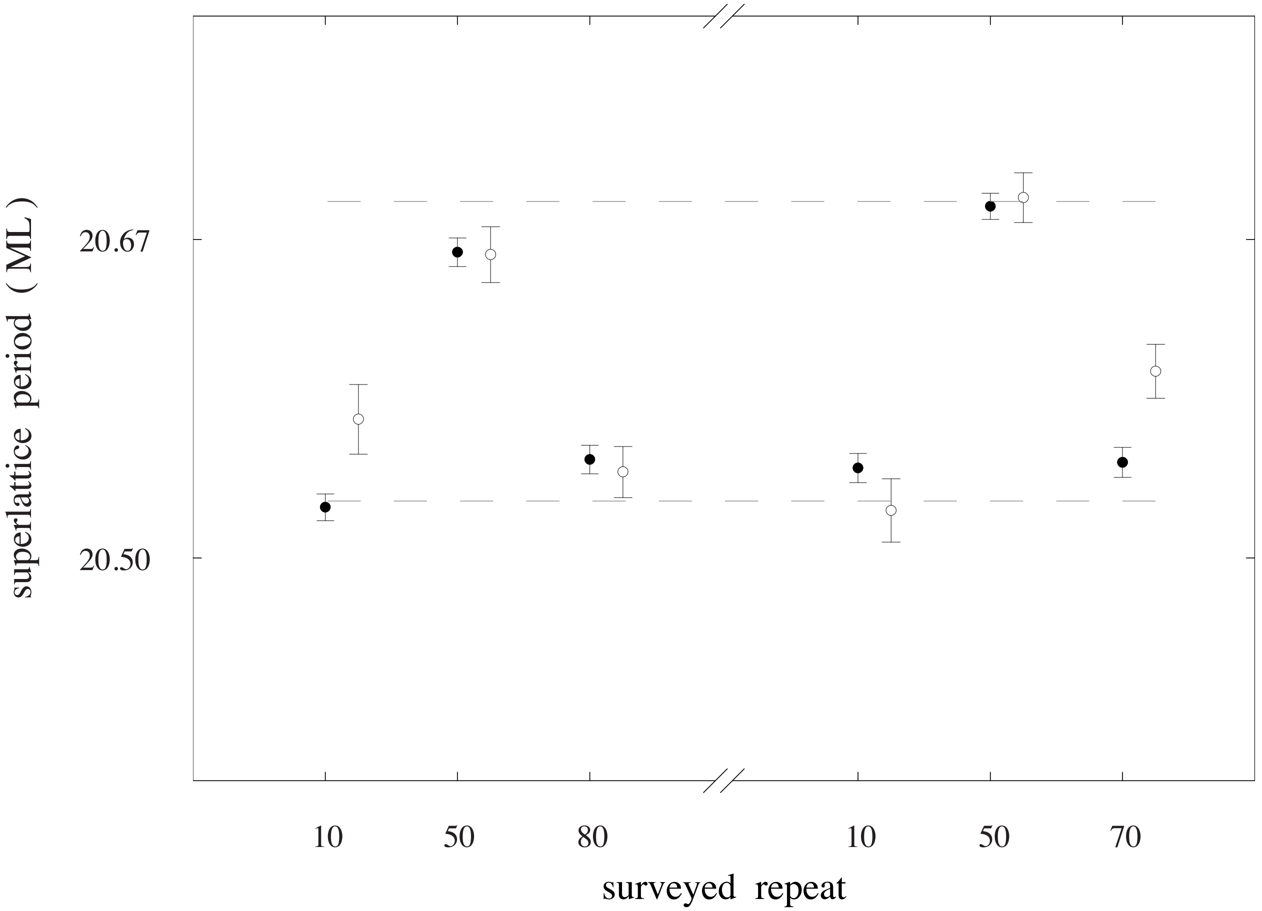


FIGURE 3.7. DFT periods (closed circles, forward/reverse averaged over Fig. 3.6 for each survey) over like repeats in both cross sections show similar vertical period drift; this drift is independently confirmed by real–space fits to the corresponding antimony segregation profiles (open circles) reconstructed from atomic counts. Uncertainties in DFT periods are propagated from Bragg plots (Fig. 3.5, bottom); uncertainties in atomic counting (Fig. 3.2) reflect binomial errors. Careful re–examination of HRXRD spectrum (Fig. 3.3) guided by STM shows each satellite peak may be described as a sum of three Guassian components, the minority components (dashed lines) of which agree well with STM measurements. Adapted from [20].

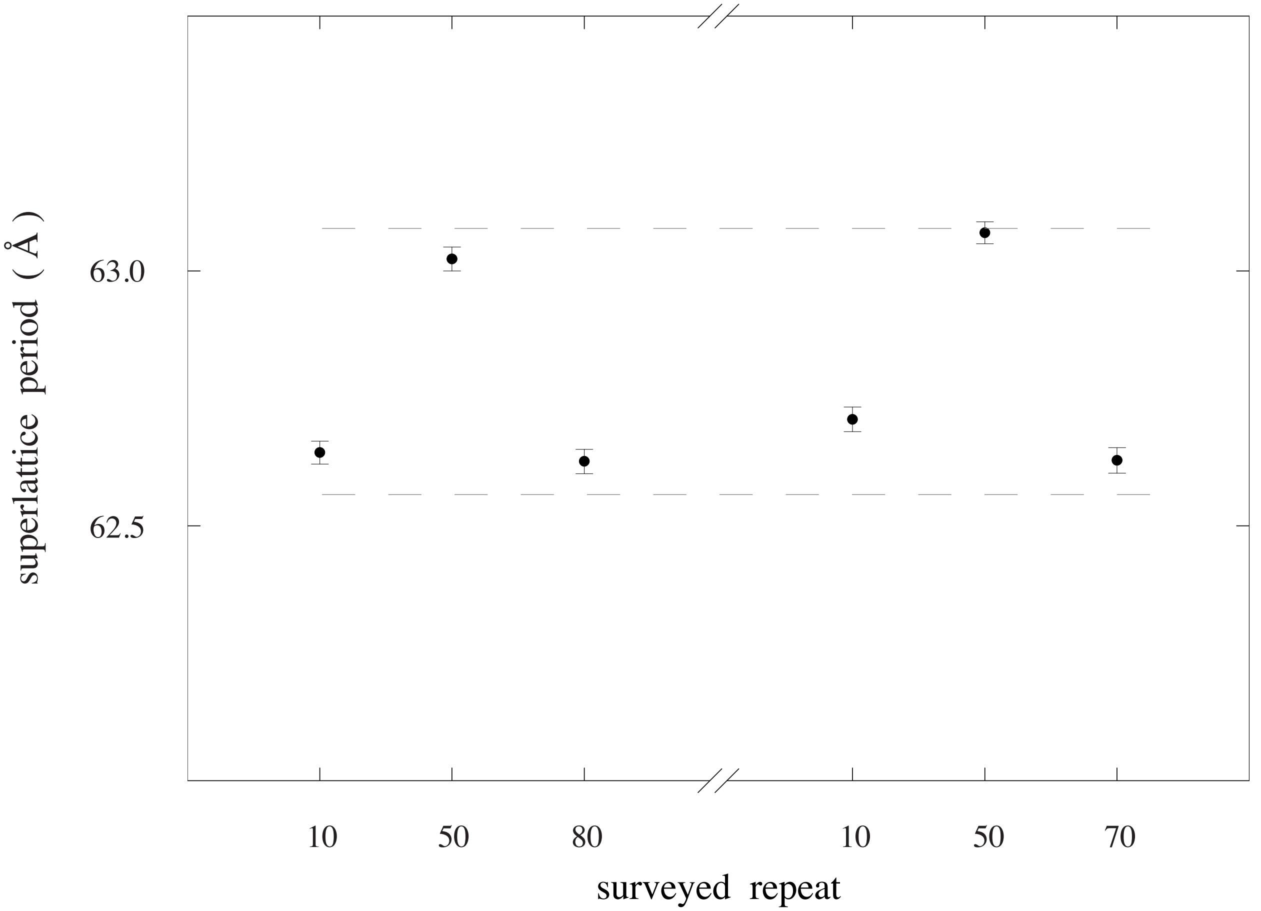


FIGURE 3.8. Laterally–averaged DFT periods (circles) from Fig. 3.7 rescaled by the corresponding survey–averaged lattice constant, from integration of the bulk antimony profile over indicated repeats. Absolute agreement (in Å) between Bragg plots assembled from minority components of satellite peaks (dashed lines) in the HRXRD spectrum (Fig. 3.3) and DFT periods, suggest that the STM composition is accurate to within 0.2%. Reprinted with permission from [20].

mismatch [23]. The beauty of such painstakingly acquired HRXRD data notwithstanding, they may in fact be subtly misleading in the sense just described.

Independent support for the period drift documented in Fig. 3.7 comes from altogether separate reasoning. Fits to the laterally–averaged [001]–monolayer–indexed antimony profiles whose reconstruction via atomic counting is facilitated by the robust, atomic–resolution, impurity discrimination in these materials (showcased in Fig. 3.2), provide detailed, quantitative insights into the compositional grading associated with antimony segregation and its effect on the resulting HRXRD spectrum [23]. Introducing the superlattice period as an additional fit parameter to this well–established segregation model [59] offers satisfying agreement between direct–and reciprocal–space period measurements assembled over essentially identical spatial domains (Fig. 3.2).

The close agreement between HRXRD, STM DFT, and real–space, counting–based period measurements again affirms that normalizing STM–satellite–peak spacings to the image–by–image DFT RLVs successfully circumvents all scan–related distortions inherent to STM. We may thus return to Fig. 3.6 confident that the lateral fluctuations exhibited there are not a measurement artifact, but the accurate sampling (with small measurement uncertainty, namely the single–image Bragg error) of a comparatively–larger–variance stochastic phenomenon "frozen into" the structure.

In light of the size of the image–to–image fluctuations documented in Fig. 3.6, we might ask, again, what uncertainty best characterizes the pooled period measurements assembled in Fig. 3.7, since every lateral distribution of periods will also be characterized by its own variance (Fig. 3.9). Within error[5], this survey variance appears to be homogeneous (i.e. independent of superlattice repeat or location on the wafer) as well as isotropic (independent of cleavage face), so that a better estimate may be obtained by averaging over all surveys. We will see the utility in this later.

It turns out the squared uncertainties in the survey–averaged periods (Fig. 3.7) are best estimated by the variance of each survey (Fig. 3.9) normalized to the number of images in that survey, a quantity known as the standard error of the mean [61]. Even though this standard error is approximately 4 times larger than our survey–averaged Bragg uncertainty (Fig. 3.10), the period differences identified in Fig. 3.7 remain statistically significant, with less than 1% probability they originate by chance [62].

Uncorrelated Interfaces and Comparison with Experiment

A superlattice is defined by the regularly spaced interfaces between two materials and the period likewise defined by successive interfaces of like type (e.g. InAsSb–on–InAs or InAs–on–InAsSb). Thus, at its core, the period we measure via the DFT is simply the difference between two like interfaces averaged over the image pixels within our window. Viewed this way, it is reasonable to examine how interface fluctuations might translate into layer thickness fluctuations and thus period fluctuations.

To begin with, we consider whether a randomly fluctuating interface can cause period fluctuations of the magnitude actually observed. To simplify matters, we first address the case where the STM–scan axes, and hence the bounding edges of the image,

---

[5] The variance errors in Fig. 3.9 are calculated from the variance of the variance, and are proportional to the variance itself [60].

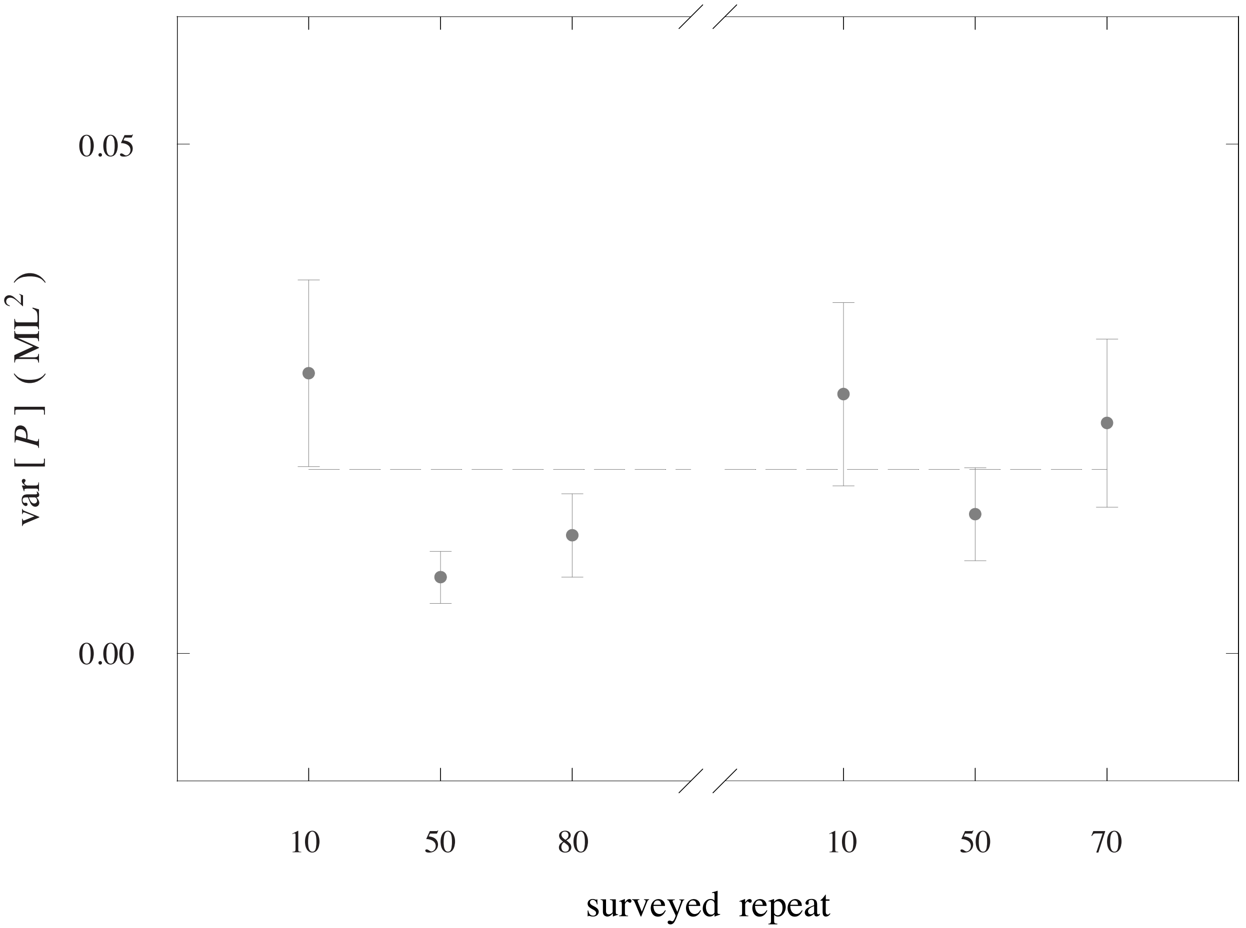


FIGURE 3.9. Image–to–image period variance due to fluctuations in Fig. 3.6 is calculated for each survey whose average is shown in Fig. 3.7. Average value (dashed line) is 0.018 ± 0.003 ML^2.

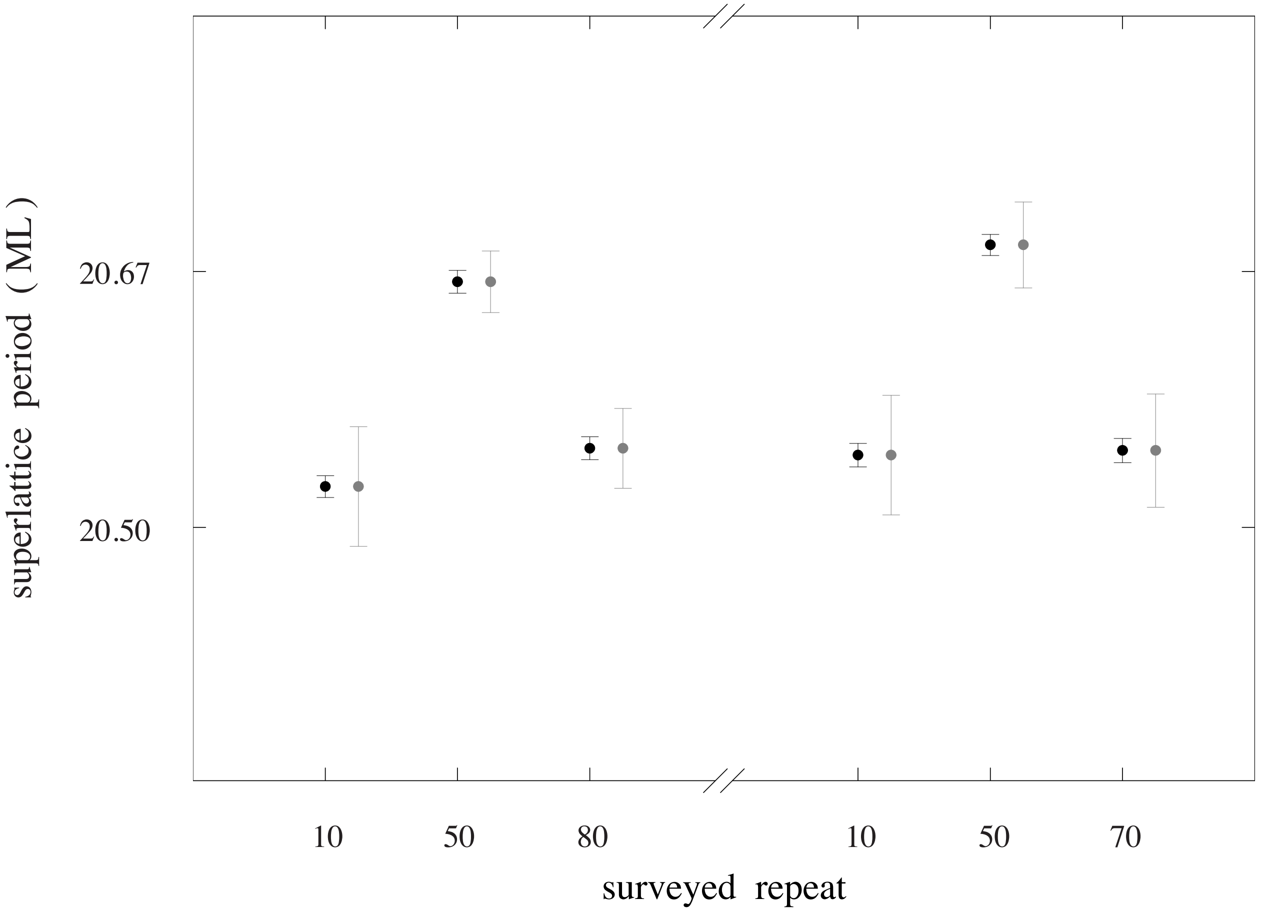


FIGURE 3.10. Laterally–averaged DFT periods (circles) from Fig. 3.7 with attending Bragg–like experimental uncertainties (black) and (more representative) standard errors of the mean (grey), which reflect the fluctuations observed in Fig. 3.6. These standard errors are approximately four times larger than the Bragg errors. Larger uncertainties notwithstanding, there is still less than a 1% probability that the observed drift occurs by chance.

are aligned with the crystal axes, setting aside for now the actual situation in which our image crop is rotated 45° relative to these axes[6]. We also adopt the convention that all explicit sums refer to a single interface in a single image and that all variances refer to image–by–image calculations over lateral surveys.

Then a single interface sampled at *N* sequential points along a <110> direction in a given image will have mean location

$$\langle z\rangle = \frac{1}{N}\sum_{i=1}^{N} z_i\,, \tag{3.1}$$

where the z_i are statistically–independent, point–to–point [001] coordinates distributed with parent mean μ and parent variance σ^2. The image–to–image distribution of means will then have expected variance

$$\mathrm{var}[\langle z\rangle] = \mathrm{var}\left[\frac{1}{N}\sum_{i=1}^{N} z_i\right] = \frac{\sigma^2}{N}\,. \tag{3.2}$$

In dealing with single–image experimental data, μ would be estimated from (3.1) and σ^2 from

---

[6] We hereafter distinguish between the simplified case and experimental reality by referring to the former as a vertical window and the later as a rotated window.

$$s^2 = \frac{1}{N-1}\sum_{i=1}^{N}(z_i - < z >)^2. \quad (3.3)$$

In dealing with lateral surveys, as in (3.2), μ would instead be estimated from E$[< z >]$ and σ^2 from E$[s^2]$, where these expectation values are survey averages. Here we assume these parent parameters are already known.

Let us ignore, for the time being, the interface located in the middle of the superlattice repeat and focus on the two bounding interfaces, thereby assuming that any asymmetry between interface templates is weak (so that $(\sigma_A - \sigma_B)/(\sigma_A + \sigma_B) \ll 1)$. The single–image, average period is then given by the difference in mean locations for every second interface[7], averaged over the integer number of repeats *M* in the [001] direction, where *M* is calculated as the greatest integer of the ratio between our DFT window size and the expected period (Fig. 3.7). Since each interface (labeled *A* through *A*+*M*) in a vertical window has the same length, this sum of differences reduces to

$$P = \frac{\langle z \rangle_A - \langle z \rangle_{A+M}}{M}, \quad (3.4)$$

with all intermediate terms cancelling in pairs. This same logic justifies our neglect of the interfaces separating constituents within a period.

---

[7] Since both averaging and subtracting are linear operations subtracting mean values is equivalent to averaging point–by–point differences; the former is easier to describe, however the latter is the operation actually performed in our simulations.

The image–to–image variance of (3.4), taken with respect to E[P] across a lateral survey, is then computed from

$$\mathrm{var}[P] = \frac{\mathrm{var}[\langle z\rangle_A - \langle z\rangle_{A+M}]}{M^2} = \frac{2\mathrm{var}[\langle z\rangle]}{M^2}, \tag{3.5}$$

where we've explicitly assumed the bounding interfaces (A and $A+M$) are statistically independent of one another, *i.e*. vertically uncorrelated[8]. Substituting (3.2) into (3.5) we finally obtain

$$\mathrm{var}[P] = \frac{2\sigma^2}{M^2 N}, \tag{3.6}$$

for the predicted period variance with interfaces whose point–to–point fluctuations are statistically independent, or laterally uncorrelated.

Fits to the antimony segregation profile in our InAs/InAsSb superlattice [20,23] provide physically reasonable estimates of the order of 1 ML for the σ in (3.6). The predicted period variance for a 40 nm x 40 nm vertical window (N=400, M=6) is then 0.0013 Å^2. As we establish later in this chapter, that prediction must be multiplied by a geometric correction (~ 1.5) before comparing to actual data, which is acquired in a rotated window. The period variance in this system (Fig. 3.9) is then ~ 50 times larger than expected on the basis of (3.6).

---

[8] An additional covariance term must be included in (3.5) whenever this is not the case.

To lowest order, lateral correlations will reduce the number of statistically independent degrees of freedom along each interface from N to $N_{eff} = N/\Lambda$, and this observation offers a natural explanation for the factor of 50 discrepancy between (3.6) and the data[9]. A more sophisticated treatment, below, indicates this estimate of the lateral correlation length is too large by about 40%, but the rough agreement between our simple approach and well–established lower bounds on interface island sizes (~ 300 Å) [63] is nevertheless noteworthy. Experiments linking (intrinsic) lateral [63] as well as (deliberately introduced) vertical [64] period fluctuations to carrier localization [65] and inhomogeneous broadening [29] in related systems underscore the potential importance of these types of STM measurements.

Numerical Calculations: Simulated Periods

To gain a better appreciation of interface roughness we will simulate profiles with specified correlation lengths using a Monte Carlo method. The real–space interface profiles are generated by inverting Fourier coefficients that pair a (assumed) Gaussian amplitude with random phases. Choosing the unnormalized Fourier amplitude, Fig. 3.11, to be

---

[9] The results in Fig. 3.9 assume the period measurement in each image is statistically independent of that in any other. The lateral correlation length arrived at here is small on the scale of image separations (50 nm), physically justifying this assumption, but explicit calculation of the image–to–image period autocorrelation from the data in Fig. 3.6, for example, confirms the result.

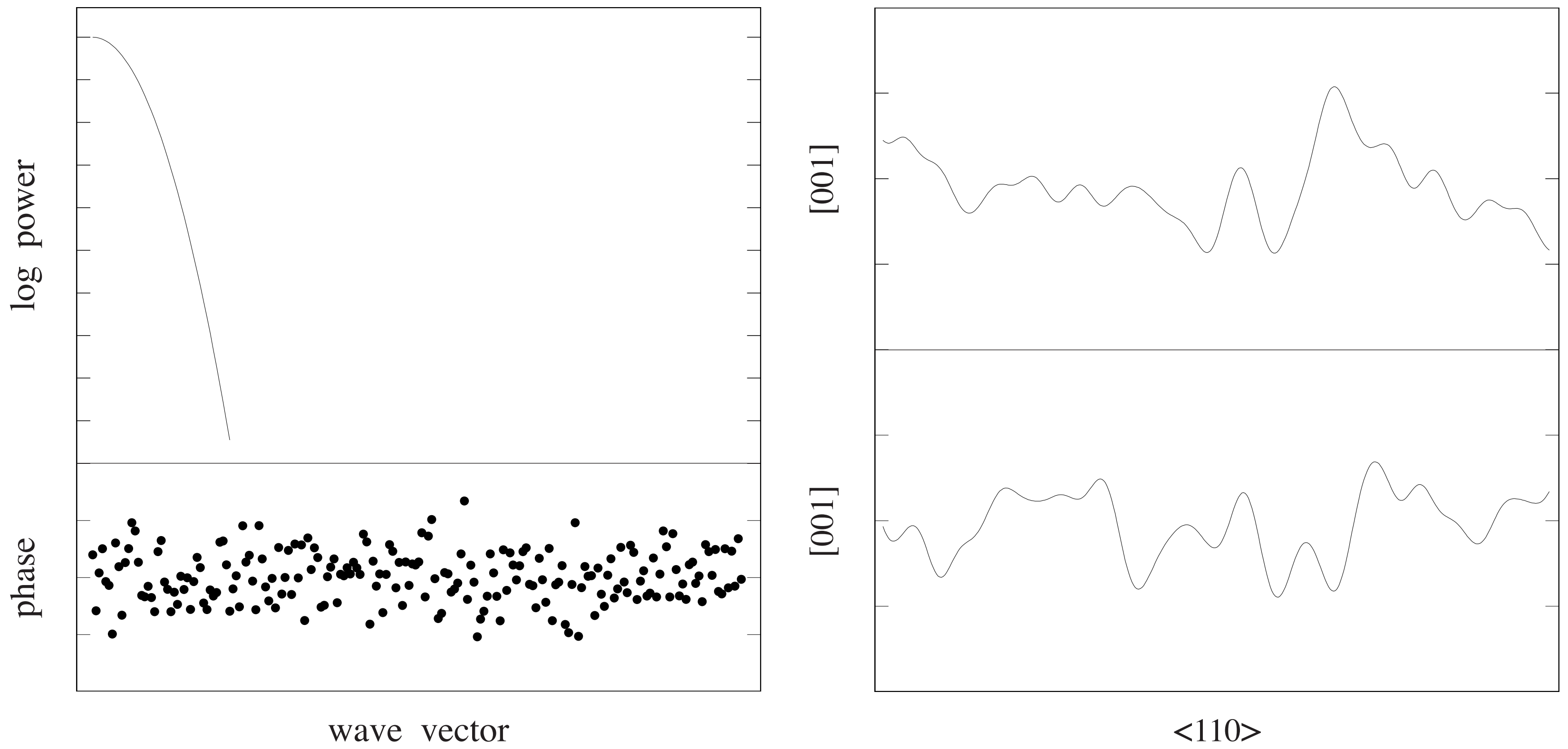


FIGURE 3.11. Constructed Gaussian power spectra corresponding to 10Å correlation length (upper left). Associating a random phase (lower left) with the square root of the power and inverse transforming produces a stochastic profile with the desired correlations, two examples of which are shown to the right.

$$|F_k| = \exp\left[-\frac{k^2}{4\Lambda_k^2}\right] \tag{3.7}$$

guarantees a Gaussian functional form[10] for both the resulting power spectral density and autocorrelation function where the real–space correlation length is given by

$$\Lambda_\mathrm{x} = \frac{L}{2\pi\Lambda_\mathrm{k}} \tag{3.8}$$

if L is the total number of points that the transform is calculated over. The k = 0 value is zeroed to remove the DC component of the spectrum and ensure each profile has a mean of zero when it is first created, then the amplitudes are scaled so that the squared area is equal to the real–space roughness variance σ^2. Each wave vector is assigned a random phase under the constraint that $F_k^* = F_{-k}$ so that the resulting profile will be real valued.

Nine profiles, which will enclose eight complete repeats, each with a length of 10.5 μm are generated for each of eight different correlation lengths ranging from 12.5Å to 100Å in steps of 12.5Å. There are two reasons that we choose to simulate a very long profile instead of several short ones despite being computationally unfavorable. Any potential endpoint problems are eliminated when the sampling window and correlation lengths are much smaller than the total length of the profile, but just as important shorter

---

[10] We use a Gaussian functional form in our examples, however this can just as easily be computed for Exponential power spectra. We will see later in this chapter that for all practical purposes our results are insensitive to this choice of functional form.

interfaces do not accurately reflect the sampling of the experimental interfaces that we wish to represent with these simulations.

Sampling the correlated interfaces (Fig. 3.12) has some interesting consequences for the mean and variance calculated over any finite window size, where the profile wanders from the parent mean for larger stretches. These values normalized to parent variance, sampled over N = 40 nm, and plotted as a function of the dimensionless quantity $\Lambda/_N$ show that as the correlation length is increased, a given sample mean is likely to be found further from the parent mean (Fig. 3.13, left) and as a consequence the sampled variance is reduced (Fig. 3.13, right). The variance of sampled means, (3.2), modified to include correlations, is given by [66]

$$\mathrm{var}[\langle z\rangle] = \frac{\sigma^2}{N}\left[1 + 2\sum_{i=1}^{N-1}\left(1 - \frac{i}{N}\right)\rho_i\right], \tag{3.9}$$

which when the correlation function, ρ_i, is replaced with a delta function correctly reproduces the uncorrelated limit of (3.2).

When applying this to the simulations we replace the autocorrelation function with it's assumed Gaussian form

$$\rho_i = \exp\left[-\frac{i^2}{2\Lambda_x^2}\right], \tag{3.10}$$

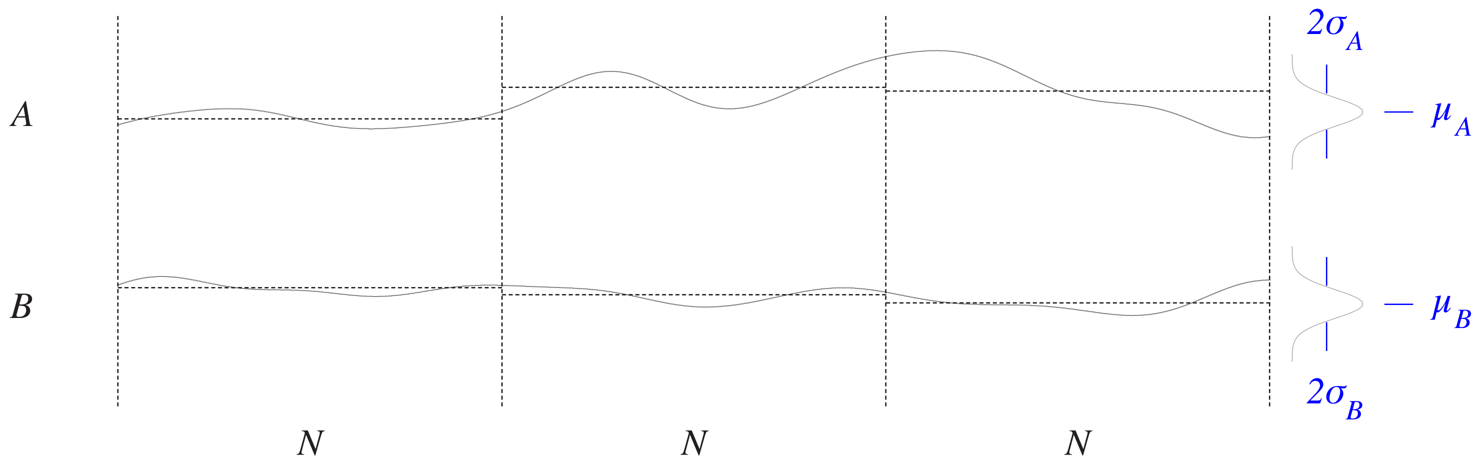


FIGURE 3.12. Correlated interface profiles like those in Fig. 3.11 have an increased probability of long stretches being located far from the parent mean (μ_A or μ_B). The N–point sampling (vertical dashed lines) of the mean interface location (horizontal dashed lines) will vary by much more than predicted for statistically independent interfaces, and so, therefore, will the period sandwiched by two of these correlated profiles.

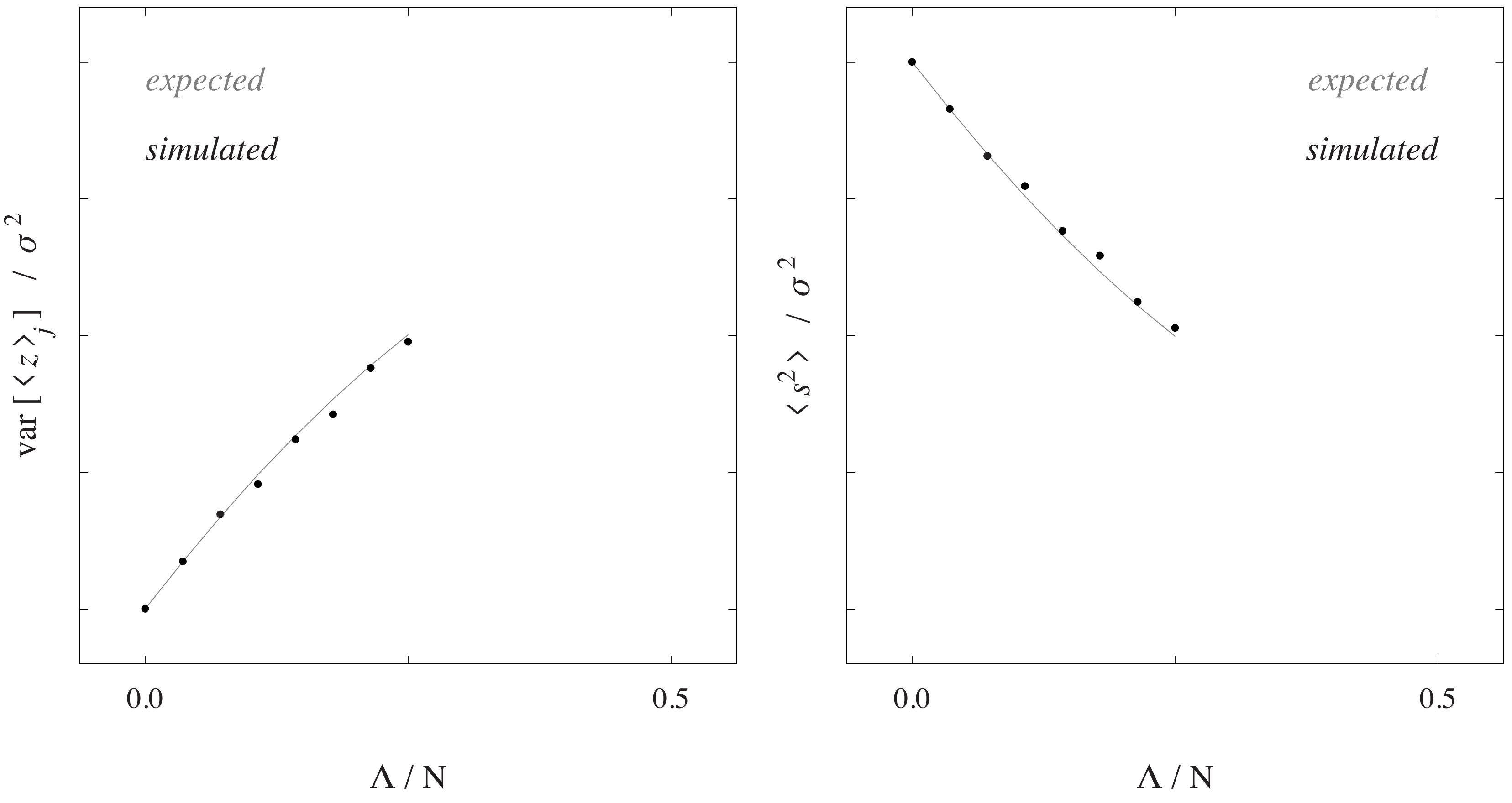


FIGURE 3.13. The variance of the sampled mean (left) and sampled mean variance (right) in units of parent roughness amplitude range from 0 to 1 vertically. The expected sampled mean calculated from (3.11) increases as the correlation length increases. As a consequence of this, the expected sampled mean variance (deviations from the sampled mean) calculated from (3.13) decreases as the correlation length increases. Adapted from [39].

and so long as the correlation length is not too large or too small ($4 \lesssim \Lambda \lesssim {}^{N}/_{4}$) the sum in (3.9) can be approximated by an integral to yield the expression

$$\mathrm{var}[\langle z \rangle] = \sigma^2 \left[\sqrt{2\pi} \left(\frac{\Lambda}{N} \right) - 2 \left(\frac{\Lambda}{N} \right)^2 \right], \tag{3.11}$$

which agrees well with the simulated points in Fig. 3.13 (left), however since we assumed that the correlation length was not too large or too small (3.11) is not strictly valid in either the large or small correlation length limit; as an example of this (3.11) implies that the uncorrelated value should be zero when we know it should actually be given by (3.2).

The expected value of s^2 (3.3) is modified similarly to (3.9) to get [66]

$$\mathrm{E}[s^2] = \sigma^2 \left[1 - \frac{2}{N-1} \sum_{i=1}^{N-1} \left(1 - \frac{i}{N} \right) \rho_i \right], \tag{3.12}$$

and plugging (3.10) into (3.12) yields

$$\mathrm{E}[s^2] = \sigma^2 \left[1 - \sqrt{2\pi} \left(\frac{\Lambda}{N} \right) + 2 \left(\frac{\Lambda}{N} \right)^2 \right], \tag{3.13}$$

again in good agreement with the simulated points in Fig. 3.13 (right) and unlike (3.11) actually does reproduce the uncorrelated limit of σ^2.

Numerical Calculations: Vertical Windows

We now move on to offset the profiles so that their means are separated by a chosen period, Fig. 3.14. Two different periods were simulated, one representative of the InAs / InAsSb material system whose period is approximately 62.5 Å; the other, 119.7 Å corresponds to the InAlAs / InGaAs material system whose interface roughness parameters have previously been directly characterized [39] and will be described in a later section. The newly created repeats are sampled by averaging over *N* points laterally along the interfaces and across *M* repeats, and the entire process replicated continuously sliding *N*+1 points laterally in between each sampling until reaching the end of the 10.5 μm stretch. For each combination of correlation length and period there are: 262 samplings of a 400 Å window, 350 samplings of a 300 Å window, and 525 samplings of a 200 Å window.

Plugging (3.11) into (3.5) and rearranging terms yields the dimensionless variance

$$\frac{\mathrm{var}[P]\,M^2}{\sigma^2} = 2\left[\sqrt{2\pi}\left(\frac{\Lambda}{N}\right) - 2\left(\frac{\Lambda}{N}\right)^2\right], \tag{3.14}$$

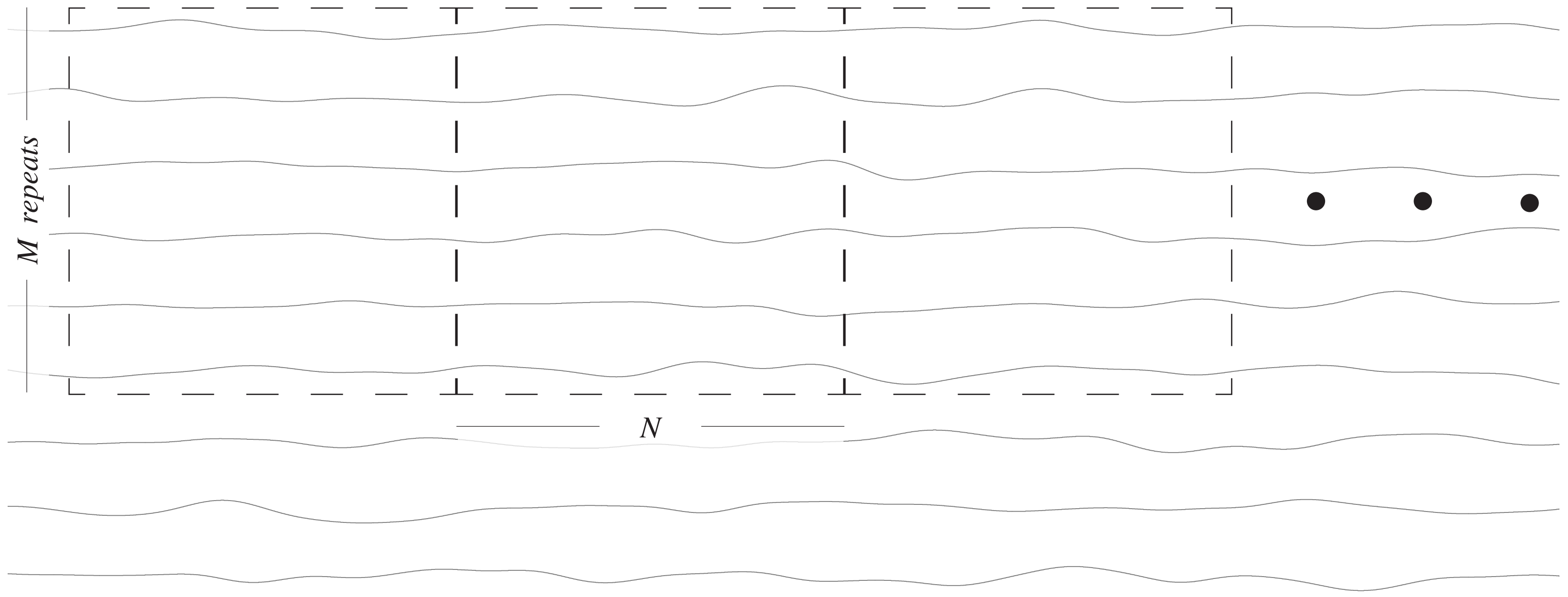


FIGURE 3.14. Correlated profiles, such as those simulated in Fig. 3.11, offset to mimic the interfaces that bound 8 superlattice repeats, overlaid with 3 vertical (crystal–aligned) sampling windows encompassing N points laterally and M repeats vertically.

which when plotted versus $\Lambda/_N$ is independent of period, roughness amplitude, and window size. This allows us to compare both period thicknesses and all three window sizes in a single universal graph (Fig. 3.15) and despite the rule of thumb that $\Lambda/_N \lesssim 0.25$ [67] we see that (3.14) is a good match to the corresponding values calculated over the simulated profiles out to $\Lambda/_N \approx 0.5$. While it appears that the expected curve is turning over and will eventually return to zero, this ignores the fact that we assumed the correlation length was not too large in going from (3.9) to (3.11). Instead, the large correlation length limit must be obtained by inserting a constant value for the autocorrelation function in equation (3.9) and simplifying to find the scaled variance approaches 2 as the correlation length gets very large; this agrees well with a 1000 Å correlation length test profile where $var[P]\, M^2/_{\sigma^2} = 1.97 \pm 0.06$ averaged over six samplings (3 window sizes in each of 2 periods). If the period variance had in fact been measured in a vertical window, this asymptotic behavior would set a minimum roughness amplitude of 1.4 Å (0.46 ML) in the InAs / InAsSb superlattice.

Since the correlation length is independent of sampling window size and the right hand side of (3.14) depends only on $\Lambda/_N$ the ratio of scaled variances for two window sizes should in theory be solvable for Λ. In practice the dependence of the ratio on correlation length is much weaker than the statistical uncertainty based on a finite number of samplings (Fig. 3.16) ruling out the possibility of simultaneously determining both σ and Λ from our period measurements. Instead what is needed is a distinct measurement relating our two unknown quantities. As we will see in a later section the

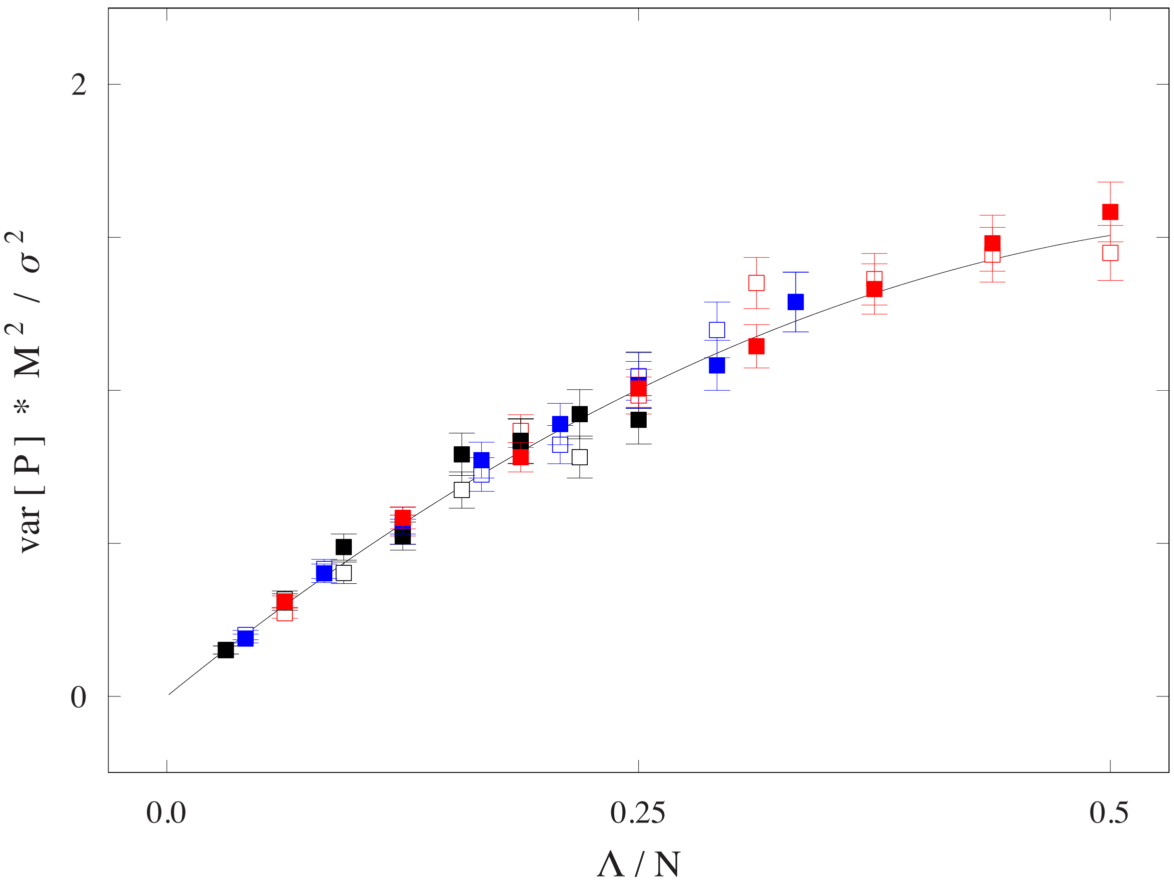


FIGURE 3.15. Dimensionless variance of simulated periods calculated in vertical (crystal–aligned) windows and scaled by the interface roughness amplitude. Further scaling by the number of interfaces (*M*) squared collapses the disparity between periods (open squares 119.7Å period, closed squares 62.5Å) and interface lengths (red squares are 200Å, blue 300Å, and black 400Å) creating a universal function of the similarly dimensionless ratio of correlation length (*Λ*) to window size (*N*). Gaussian theory (line) calculated from Eq. (3.14) agrees well with simulated profiles (squares).

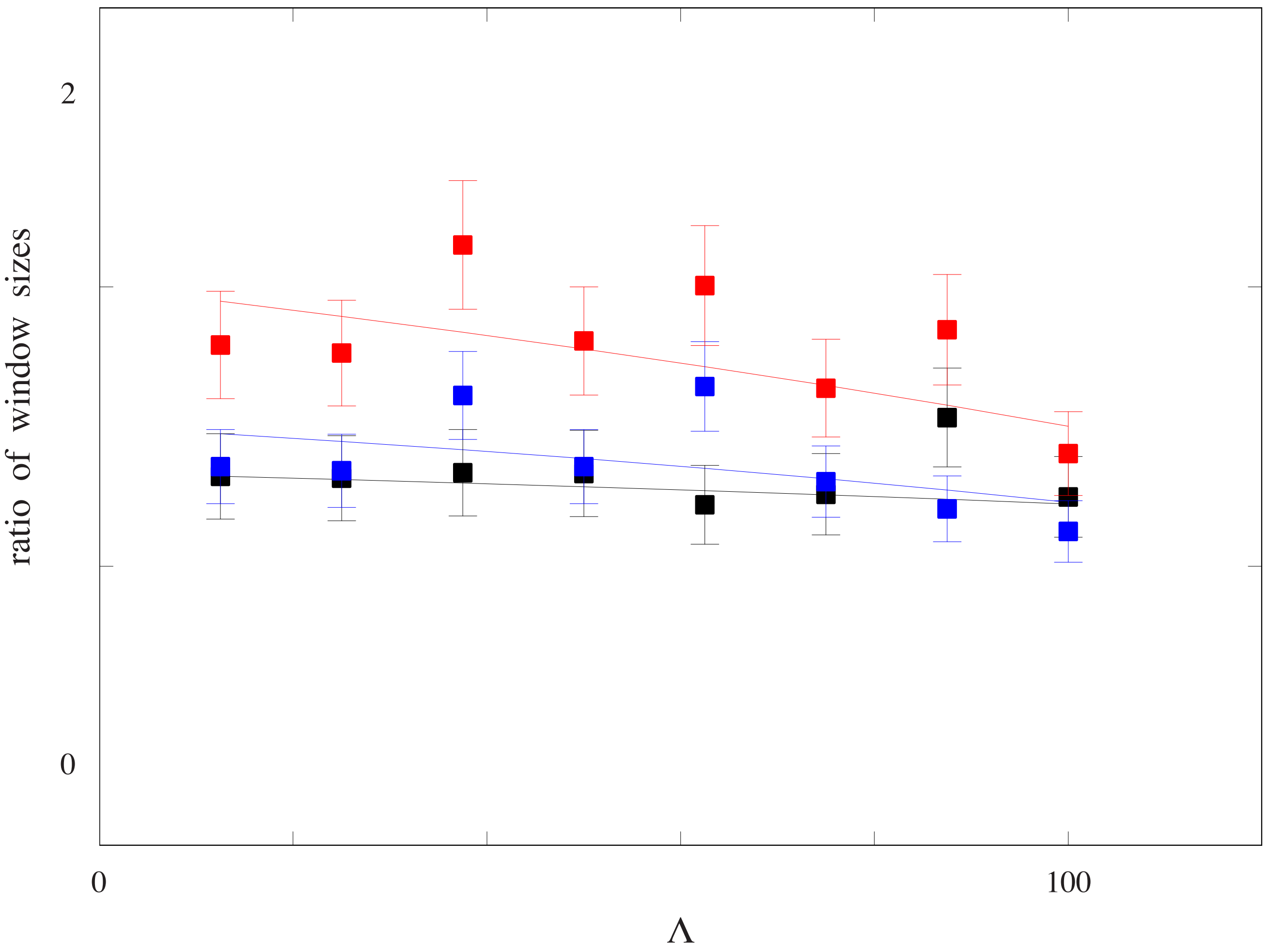


FIGURE 3.16. Ratio of simulated period variances with given correlation length calculated in different vertical window sizes (black squares are 300 Å : 400 Å ratio, blue are 200 Å : 300 Å, and red are 200 Å : 400 Å) agrees well with Gaussian theory (line). Statistical uncertainty dwarfs dependence on correlation length.

measurement of the correlation length itself is fraught with difficulties, however it is rather straightforward to measure the sampled variance so long as an interface can be defined in the STM image. It makes sense then to solve (3.13) for σ^2 and place it into (3.14). Rearranging terms the variance of periods is rescaled by the sample interface variance to get

$$\frac{\mathrm{var}[P]\,M^2}{s^2} = \frac{2\left[\sqrt{2\pi}\left(\frac{\Lambda}{N}\right) - 2\left(\frac{\Lambda}{N}\right)^2\right]}{1 - \sqrt{2\pi}\left(\frac{\Lambda}{N}\right) + 2\left(\frac{\Lambda}{N}\right)^2}. \tag{3.15}$$

Fig. 3.15 can be replotted incorporating the change from parent variance to sampled variance via (3.13) in Fig. 3.17, along with the expected curve (3.15). If s^2 is known, such as for the simulated profiles, we can directly determine the uncertainty in Λ by inferring a value from Fig. 3.17 and compare it to the value simulated (Fig. 3.18, left). A histogram of these relative uncertainties (Fig. 3.18, right) indicates that even with several hundred images any inferred correlation length will only be accurate to ~ ± 7.5%.

Numerical Calculations: Rotated Windows

Having established a solid understanding of the sampled periods with vertical windows aligned to the crystal axes, we now explore the effects of sampling the interfaces at 45° relative to the crystal axes. In the experimental data our scan axes are aligned with the image *x*,*y* axes and the crystal axes are rotated by 45°, however for the simulations it is much easier to rotate our windows in Fig. 3.14 by 45° over a fixed set of

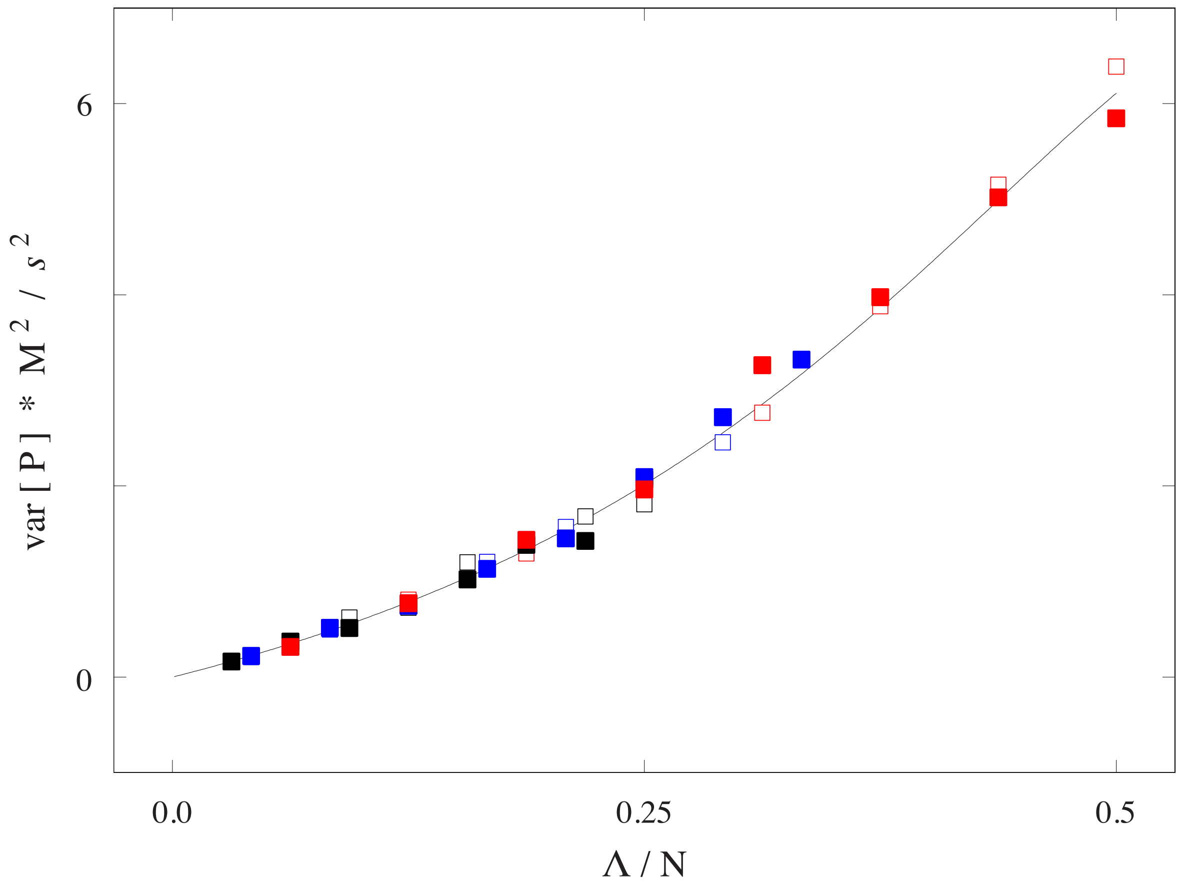


FIGURE 3.17. Variance of simulated periods from Fig. 3.15 is now scaled by the sampled roughness amplitude. Further scaling by the number of interfaces (*M*) collapses the disparity between periods (open squares 119.7Å period, closed squares 62.5Å) and window sizes (red squares are 200Å, blue 300Å, and black 400Å) creating a universal function of the similarly dimensionless ratio of correlation length (Λ) to window size (*N*). Gaussian theory (line) calculated from Eq. (3.15) agrees well with simulated profiles (squares).

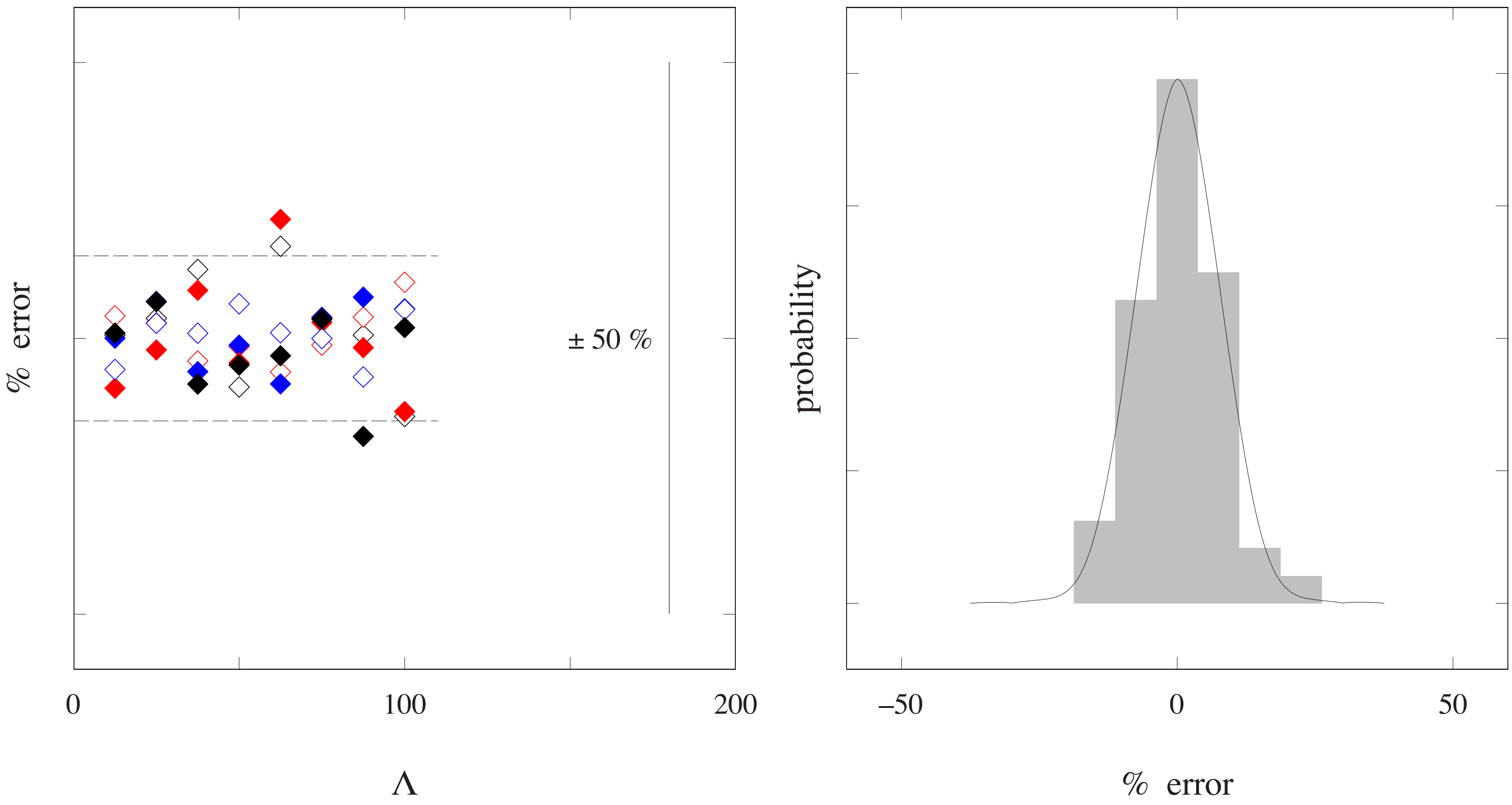


FIGURE 3.18. Relative error between known simulated correlation lengths and those predicted by theory given the simulated vertical–window period fluctuations associated with said lengths (left). A histogram of these values (right) is an approximate stand–in for the uncertainty in any given measurement. Inferred 95% confidence interval given by ± 15.0% is illustrated as dashed lines in left panel. Colors correspond to those in Fig. 3.17.

interfaces, shown in Fig. 3.19, rather than rotating the interfaces under a fixed set of sampling windows. When analyzing the rotated windows we must be careful that the window size is not made too large, thereby artificially excluding interfaces that should be present inside the sampling window. Instead of using a spreadsheet program to calculate differences like was done in the vertical window, we write our own routine that calculates differences between every consecutive interface for all 10.5 μm, then averages the differences if and only if both interfaces that were subtracted are within a border defining the rotated window.

Plotting the variances calculated over the new rotated (*x*–*y* image–aligned) window (Fig. 3.20) scaled the same as the vertical window, we see that the curves are no longer universal. An immediate concern is that the window could be in some way aligned so that longer interfaces are presented for certain window sizes, to ease this concern we can directly count the number of differences in both the rotated and vertical windows.

The number of differences in the vertical window is simply the length of each interface multiplied by the number of repeats contained within the window. The number of rotated differences is smaller, as expected since the number will be set by the shorter of two consecutive interfaces, and dividing the number of rotated differences by the number of vertical differences shows a trend within a family of windows for a given period and a split between the families for two different periods (Fig. 3.21, left). However these disparities collapse when each value is multiplied by $\sqrt{P/_N}$, leaving all of the measurements clustered around 0.33 (Fig. 3.21, right). Therefore, due to the

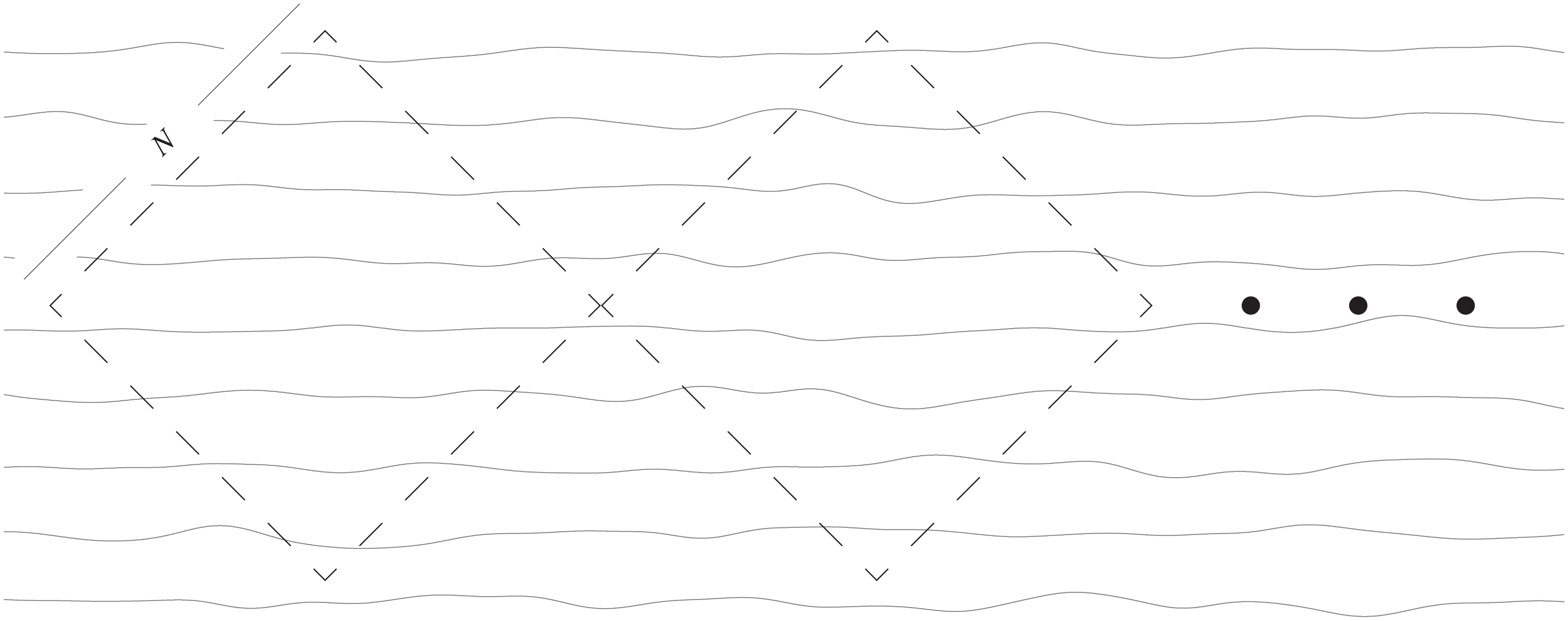


FIGURE 3.19. Same correlated profiles as Fig. 3.14 now overlaid with rotated (*x*–*y* image–aligned) sampling windows of size *N* x *N*. Superlattice repeats contained in the sampling window are no longer equally sampled.

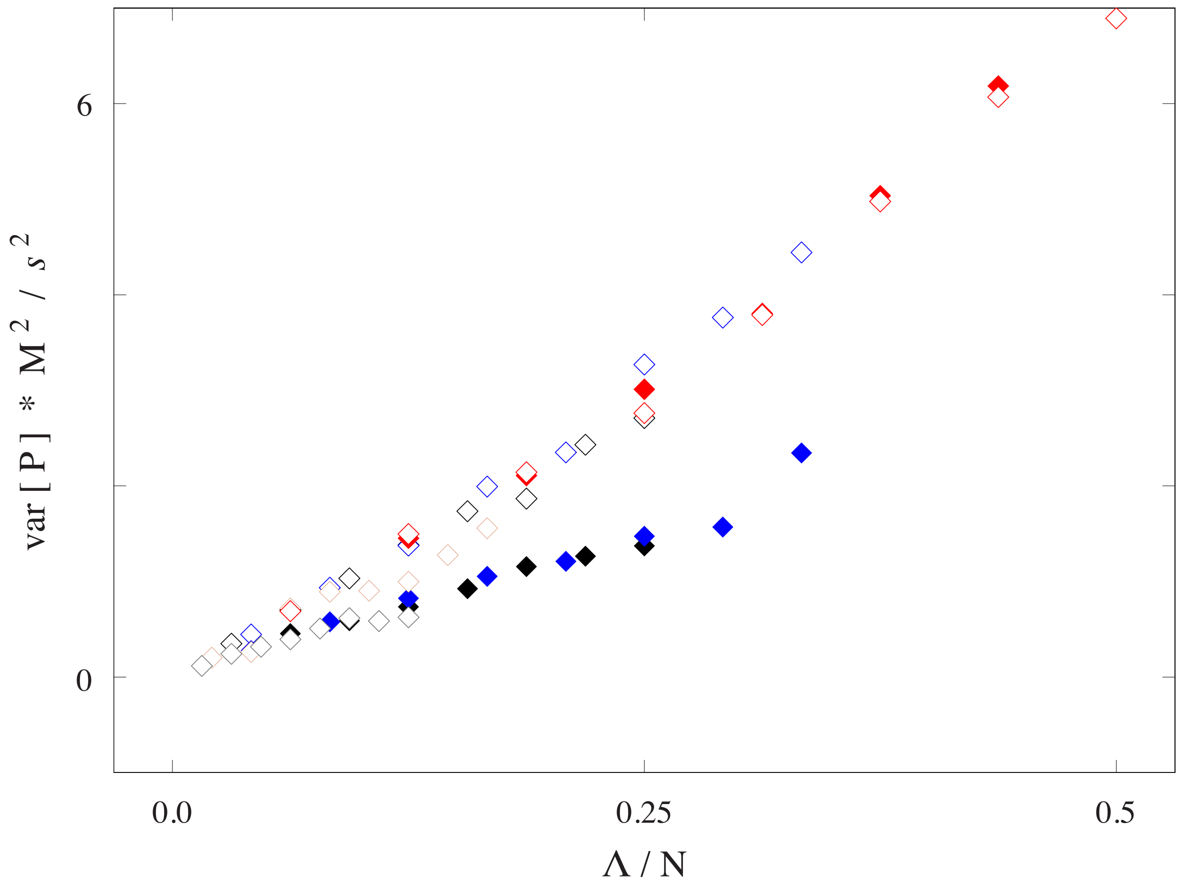


FIGURE 3.20. Dimensionless variance of simulated periods calculated in rotated windows and scaled by the sampled roughness amplitude. Further scaling by the number of interfaces (*M*) no longer collapses the disparity between periods (open diamonds 119.7Å period, closed diamonds 62.5Å) or window sizes (red diamonds are 200Å, blue 300Å, black 400Å, tan 600Å, and grey 800Å).

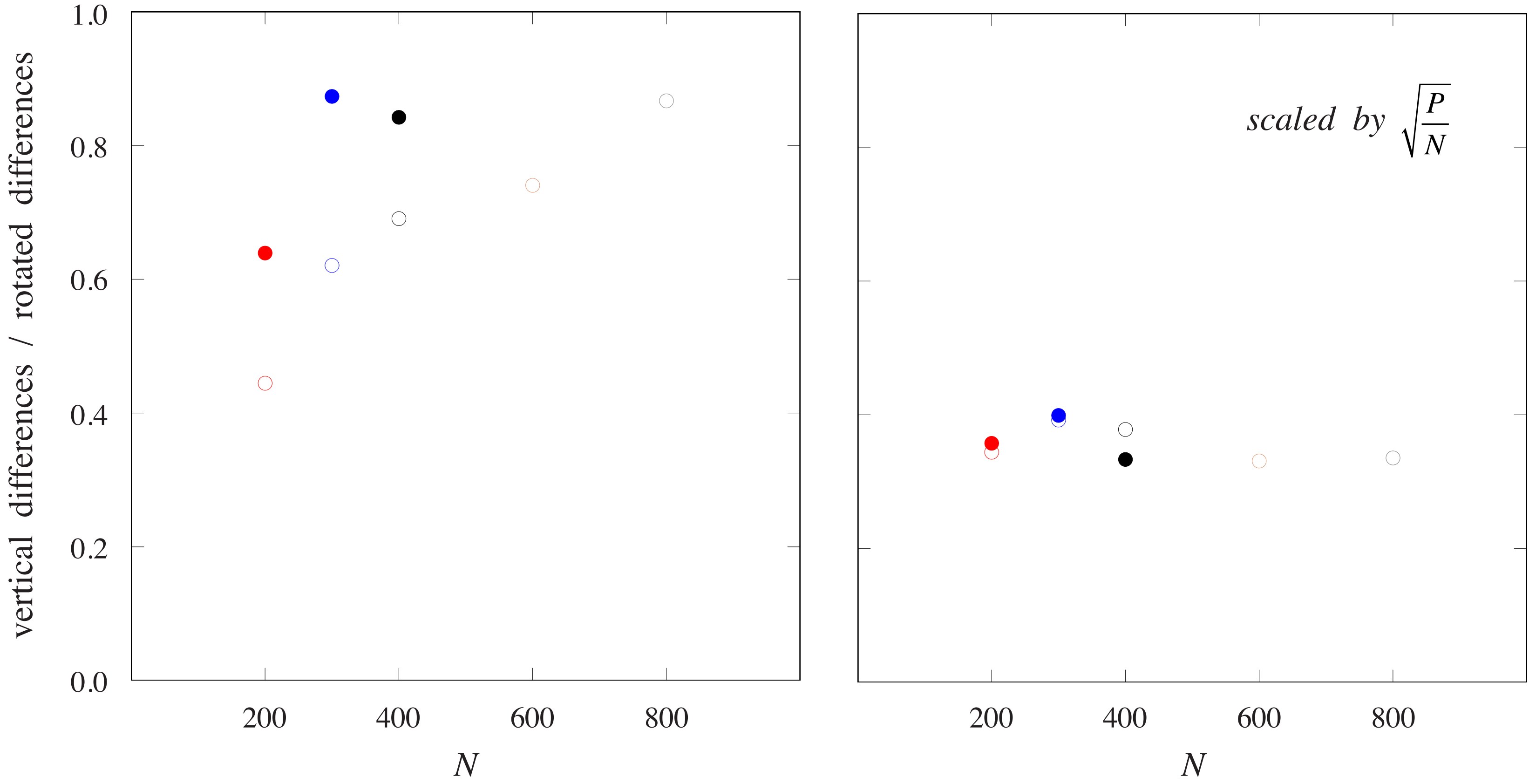


FIGURE 3.21. Fewer differences between adjacent interfaces are completely enclosed within a rotated window than within a vertical window. Their ratio (left) depends on period (open circles 119.7Å period, closed circles 62.5Å) and window size (red circles are 200Å, blue 300Å, black 400Å, tan 600Å, and grey 800Å). Scaling the ratio by the square root of period divided by window size corrects for this discrepancy (right).

specific geometry of a window at 45°, there are $3\sqrt{N/_P}$ times fewer differences in the rotated window than in the vertical window.

To see how our vertical axis in Fig. 3.20 needs adjusted we rewrite the left hand side of (3.15) as

$$\frac{\mathrm{var}[P]\, M^2}{s^2} = \frac{\mathrm{var}[P]\, N^2 M^2}{s^2 N^2}, \tag{3.16a}$$

where the multiplication by N^2M^2 counteracts the dependence of the variance on the total number of differences. Then since the total number of differences goes down when we rotate the window by 45°, (3.16a) needs to be multiplied by $N/_P$ to once again counteract the corresponding change in variance[11], resulting in

$$\frac{\mathrm{var}[P]\, N^2 M^2}{s^2 N^2}\left(\frac{N}{P}\right) = \frac{\mathrm{var}[P]\, M^2}{s^2}\left(\frac{N}{P}\right). \tag{3.16b}$$

Adjusting our vertical axis to correspond to (3.16b), closes the gap between the different sampling windows considerably (Fig. 3.22) but does not result a universal curve as observed in the vertical window variance in Fig. 3.17. To understand why we consider the schematic in Fig. 3.19. The interfaces in the diamond are all different lengths suggesting that an integer number of repeats, while well motivated for the

---

[11] We chose not to include the constant 3 in our rescaling, letting it be wrapped up in the dependence on $\Lambda/_N$ and instead focused solely on scaling out the $N/_P$ dependence.

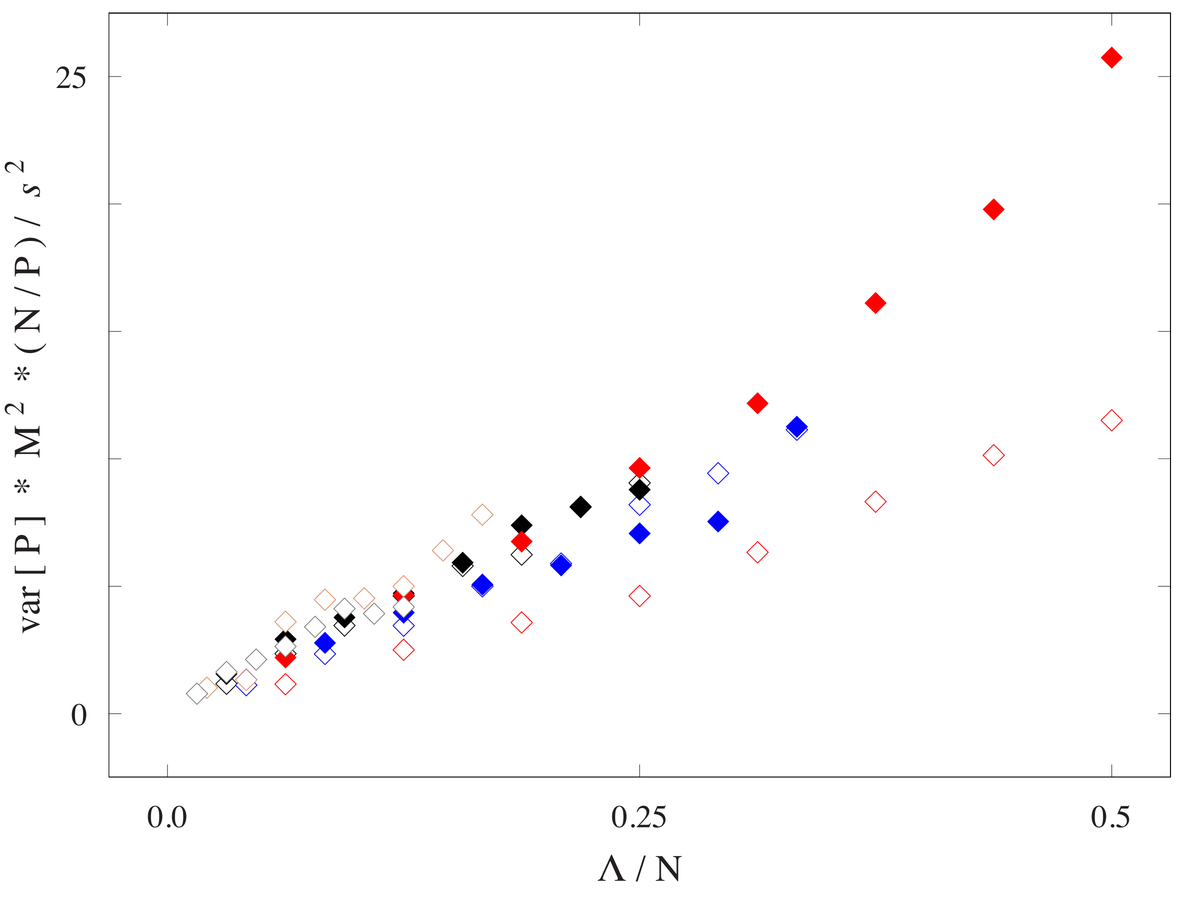


FIGURE 3.22. Scaling Fig. 3.20 to account for the smaller number of differences in rotated windows relative to vertical windows (illustrated in Fig. 3.21) still does not collapse disparity between periods (open diamonds 119.7Å period, closed diamonds 62.5Å) or window sizes (red diamonds are 200Å, blue 300Å, black 400Å, tan 600Å, and grey 800Å).

vertical window, may not sufficiently describe the measured periods when the interfaces are unequally weighted. To finish collapsing the different curves we allow the integer M, which was obtained by taking the greatest integer of ${}^{N}\!/_{P}$, to turn back into the floating point ${}^{N}\!/_{P}$. With this change (Fig. 3.23) we see that the various curves have once again collapsed to a universal representation, but unsurprisingly the vertical expectation (3.15) no longer agrees with simulated values in the rotated window (Fig. 3.24).

The expected variance (3.15) can be modified if we know the transfer function from the universal curve in the vertical window (Fig. 3.17) to the universal curve in the rotated window (Fig. 3.22). This transfer function, H shown in Fig. 3.25, can be obtained by dividing each point in Fig. 3.23 by the expectation (3.15). The expression missing from (3.15) is

$$H \approx 4\sqrt{{}^{\Lambda}\!/_{N}} \tag{3.17}$$

as demonstrated by the fit in Fig. 3.25. It's not immediately obvious where this expression comes from, however it is reminiscent of a chain rule, which crops up when changing variables of integration. The expected scaled variance in the rotated window

$$\frac{\mathrm{var}[P]}{s^2}\left(\frac{N}{P}\right)^3 = \frac{2\left[\sqrt{2\pi}\left(\frac{\Lambda}{N}\right) - 2\left(\frac{\Lambda}{N}\right)^2\right]}{1 - \sqrt{2\pi}\left(\frac{\Lambda}{N}\right) + 2\left(\frac{\Lambda}{N}\right)^2}\, 4\left(\sqrt{\frac{\Lambda}{N}}\right). \tag{3.18}$$

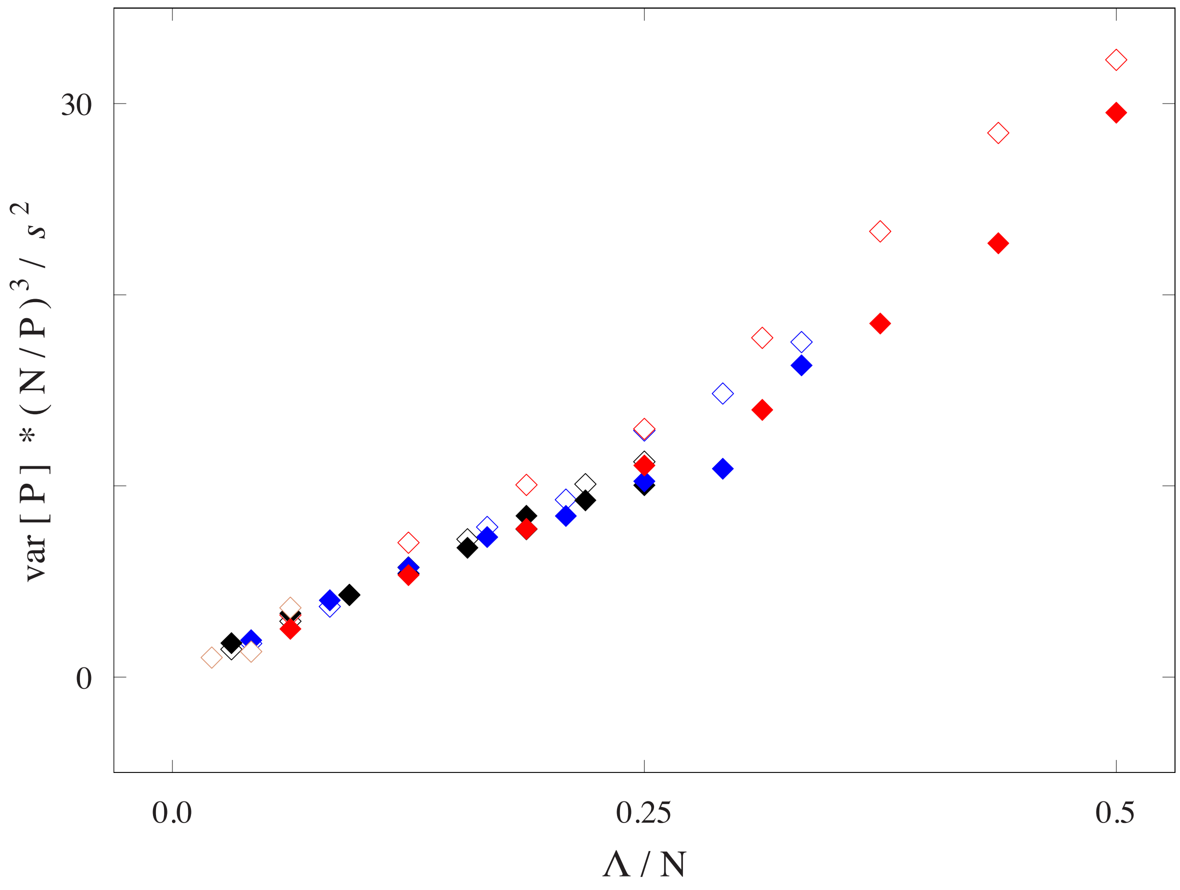


FIGURE 3.23. Scaling Fig. 3.22 by floating point values instead of integer values collapses the disparity between periods (open diamonds 119.7Å period, closed diamonds 62.5Å) and window sizes (red diamonds are 200Å, blue 300Å, black 400Å, tan 600Å, and grey 800Å).

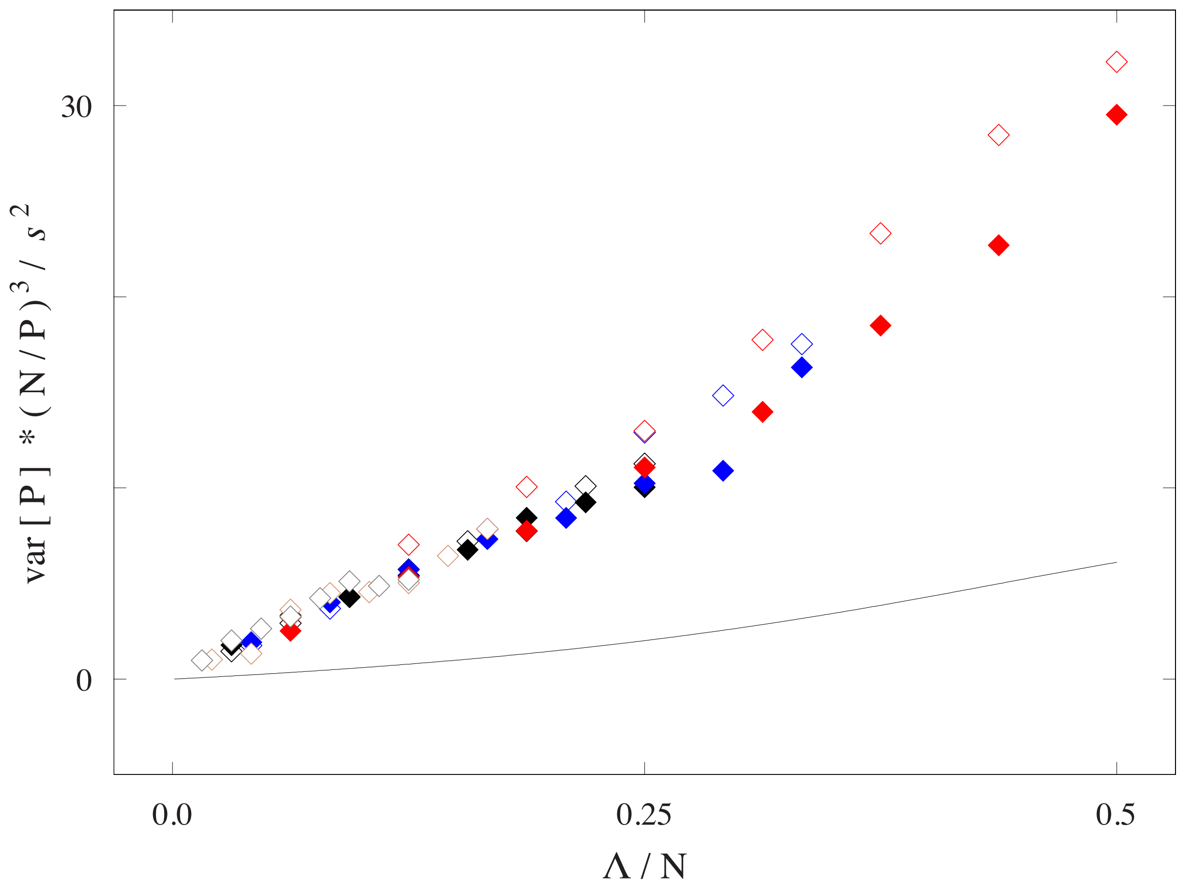


FIGURE 3.24. Scaled rotated–window period variances reproduced from Fig. 3.23 with Gaussian theory calculated via (3.15) for vertical windows (black line) overlaid.

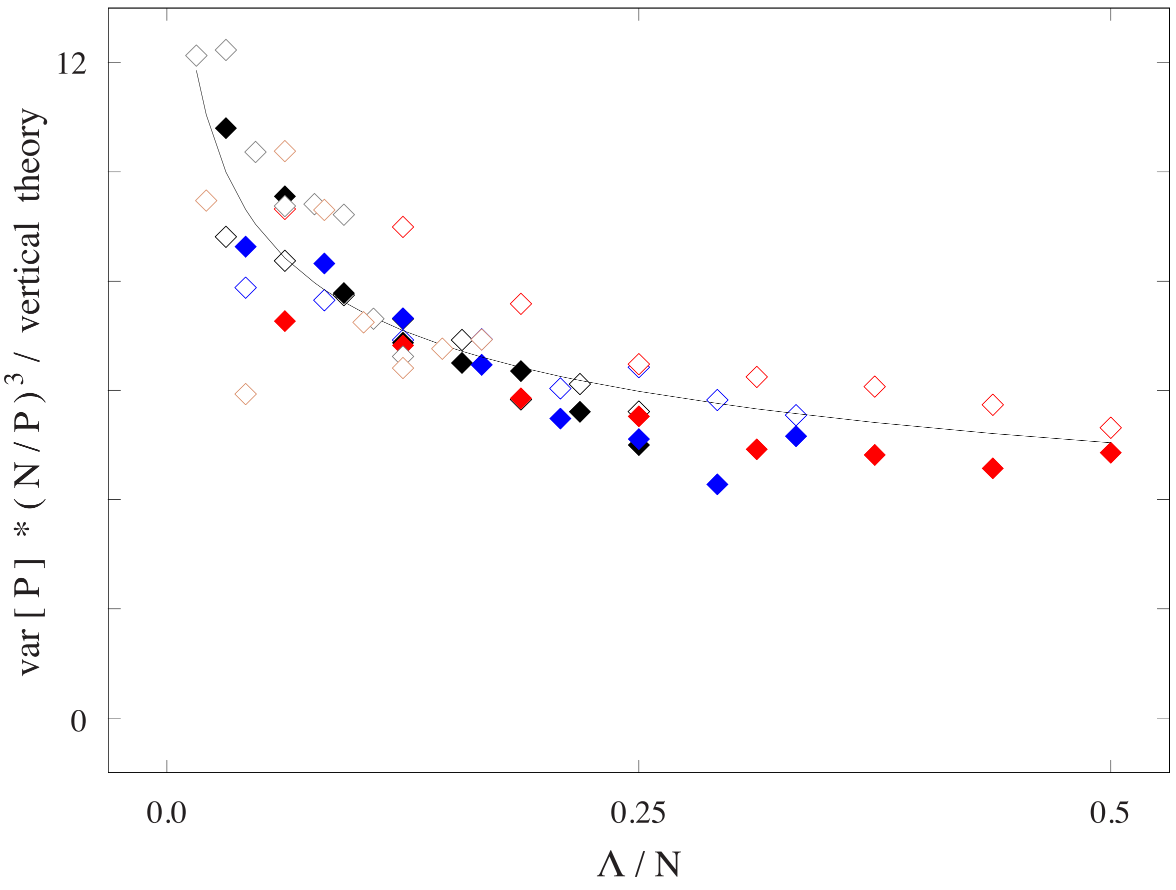


FIGURE 3.25. Dividing rotated–window period variances in Fig. 3.24 by vertical–window Gaussian theory also in Fig. 3.24 captures transformation caused by window rotation. Resulting fit (black line) produces an empirical transfer function between the two sampling windows given by (3.17).

agrees well with the values calculated for the simulated profiles up to $\Lambda/_N$ ~ 0.33 (Fig. 3.26). We once again ask if s^2 is measured separately (in a vertical window) how well do we know the correlation length. Inferring the correlation length from the simulated profiles (Fig. 3.26) just as we did in the vertical window, we notice two things, one is that the dispersion of relative uncertainties in Fig. 3.27 (left) is slightly larger reflected in the ± 11.6% width of their histogram (Fig. 3.27, right), additionally we find that our empirical theory systematically underestimates the simulated correlation length by 3%. We are now in a position to compare these theoretical underpinnings to an experimental benchmark of directly measured interface roughness[12].

Direct Measurement of Interface Roughness in Quantum Cascade Materials

We saw in the last section that the period variances in multiple windows don't provide enough information to determine both the roughness amplitude and the correlation length without unreasonable amounts of images to reduce uncertainties, so to compare with experiment we need to be able to directly identify and extract the heterojunctions that define the superlattice. From these interface profiles we may then measure the correlation length and sampled roughness amplitude. As we will see shortly the measurement of the sample roughness variance is straightforward provided that an interface profile can be identified, however measuring correlation lengths is much

---

[12] The results of which were obtained in conjunction with Dr. Federico Lopez. He obtained the data and analyzed the interface roughness amplitude. I took over and performed the reciprocal–space analysis after he graduated.

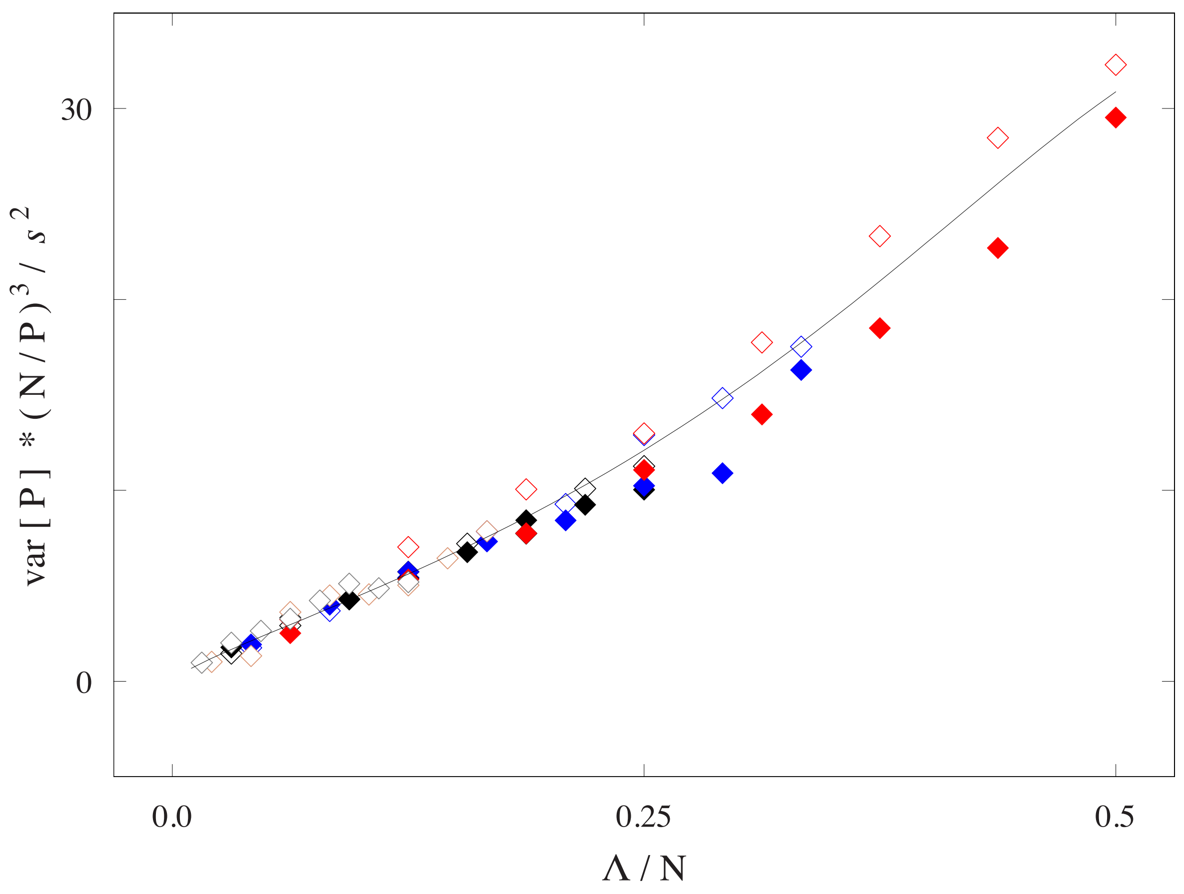


FIGURE 3.26. Rotated–window period variances in Fig. 3.24 with empirical Gaussian theory (black line) obtained from transfer function in Fig. 3.25.

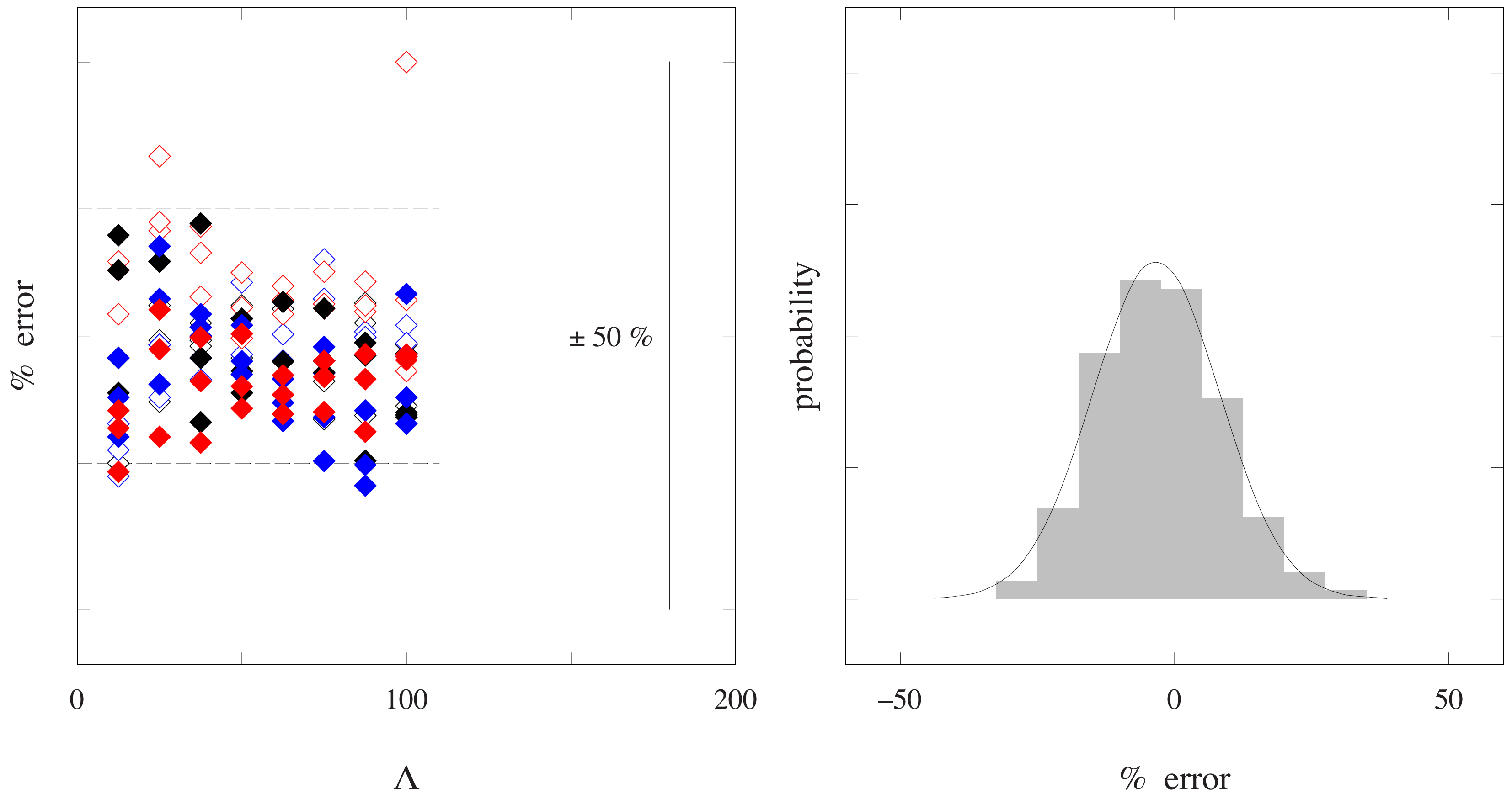


FIGURE 3.27. Relative error between known simulated correlation lengths and those predicted by theory given the simulated rotated–window period fluctuations associated with said lengths (left). A histogram of these values (right) is an approximate stand–in for the uncertainty in any given measurement. Inferred 95% confidence interval given by ± 23.2% is illustrated as dashed lines in left panel. Simulations underestimate the actual correlation length by ~ 3%. Colors correspond to those in Fig. 3.17.

trickier and prone to uncertainties. The individual InAs and InAsSb layers in the superlattice as imaged don't display an outright break between them because the electronic contribution to the contrast is weak, and the exponential rise associated with the segregation profile at the beginning of the layer produces a very diffuse border that can only be identified once the antimony fraction is averaged laterally over 100 nm.

We instead turn to a material system that *does* display a large greyscale contrast between subcomponents of the superlattice, the InGaAs / InAlAs material system, which is used extensively in quantum cascade lasers and detectors. This large difference in greyscale values has two sources that add constructively, the first is a valence band offset that is comparable to the InAs / InAsSb material system and the second is the large difference in strain which relaxes into (tensile InAlAs) and out of (compressive InGaAs) the surface post cleavage. Once atomic corrugation is filtered from the images the histogram of greyscale intensities is clearly bimodal with each individual lobe associated with one and only one of the two materials (Fig. 3.28, right). Thresholding on the value that maximally separates the two peaks creates a mask of either the InAlAs or InGaAs layer the meeting of which defines the interface profile.

The interfaces highlighted in Fig. 3.28 (left) appear to fluctuate on the scale of a monolayer in the growth direction and smoothly transition from one fluctuation to the next on the order of 5–10 lattice constants in the growth plane. It is no surprise then that a lateral mapping of period measurements in this system, similar to those tabulated in the InAs / InAsSb system, also shows fluctuations well outside experimental uncertainties (Fig. 3.29). Two surveys were taken over the same exact set of repeats in (–1–10) cross

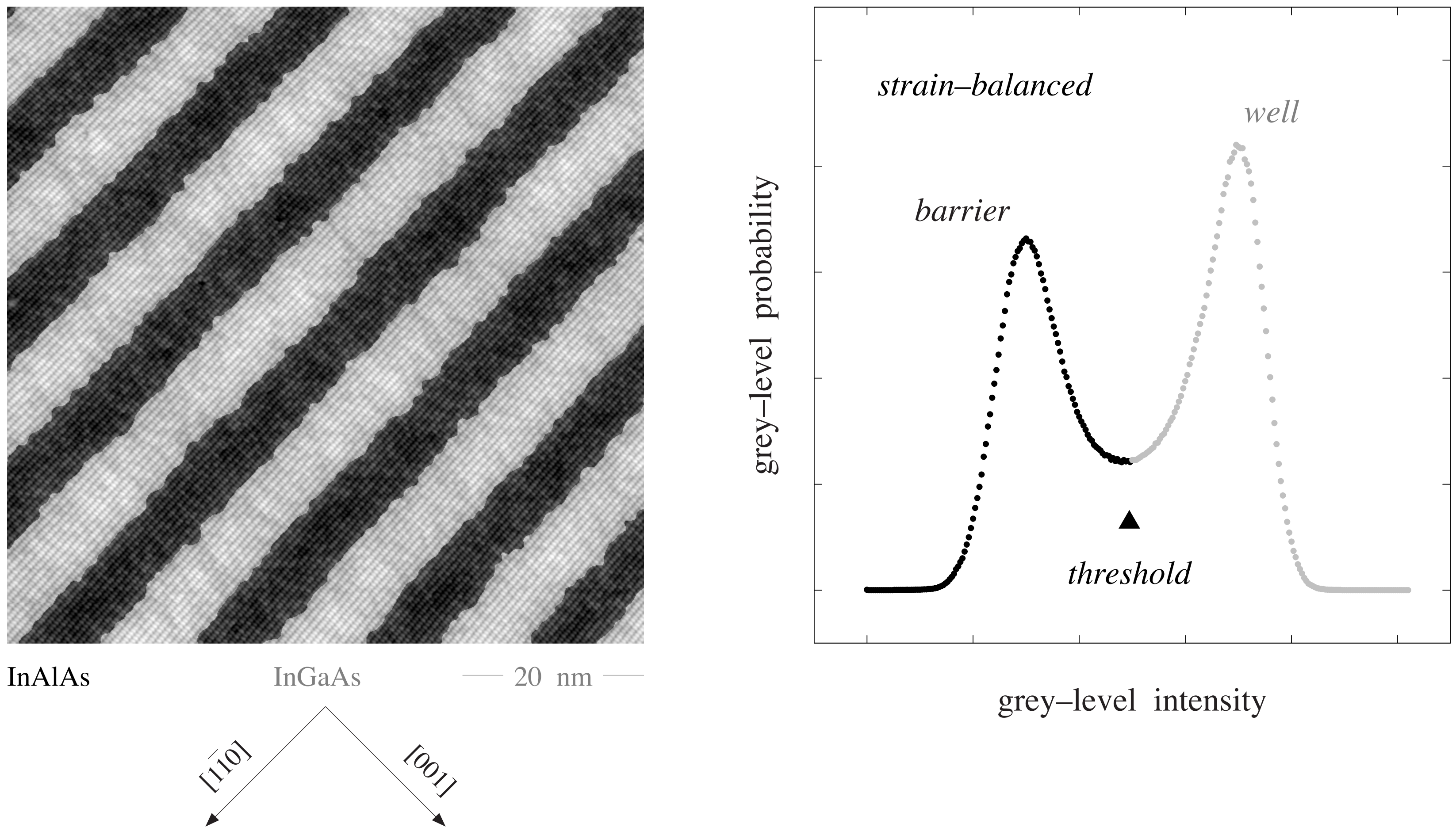


FIGURE 3.28. Atomic–resolution InAlAs / InGaAs superlattice image (left) accentuated via binary mask to emphasize delineation of the heterojunctions. The two–level mask was created by identifying regions in the image, which correspond to intensities below (black) and above (grey) the threshold value separating the two modes in a bandwidth–limited contrast histogram (right). Growth direction is from top left to bottom right. Reprinted with permission from [39].

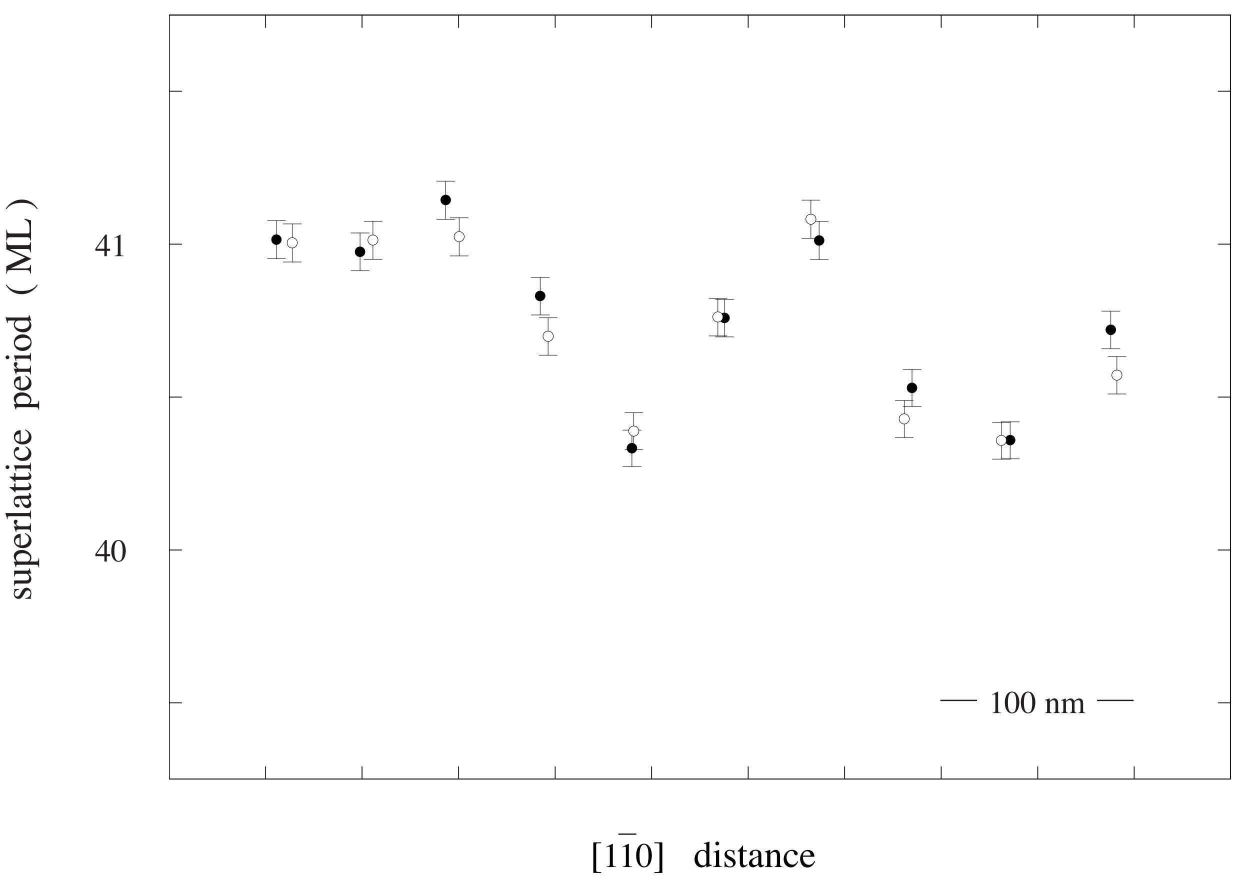


FIGURE 3.29. Superlattice periods computed from the high–frequency scan display close agreement between images taken 24 hours apart (open and closed circles). The measurements mapped out laterally throughout the survey evidence fluctuations measurably exceeding those expected on the basis of experimental (“Bragg”) errors. Reprinted with permission from [39].

section on two consecutive days. The second half of these surveys overlap strongly in space and as it turns out the period measurements are almost indistinguishable even sharing the fluctuations from one image to the next. This is another testament that these fluctuations aren't due to measurement errors, but are sampling differences that were "frozen" into the structure during growth. The variance for these two surveys in their full form is summarized in Fig. 3.30 along with a slightly shorter version from (1–10) cross section, and similar to the InAsSb material system the variance increases when sampling the period with a smaller window. Once again the variances appear to be isotropic within the confines of large uncertainties; as we will see shortly the interface roughness in this material system is nearly isotropic rendering this symmetry in the period variances unsurprising.

Before extracting the sought after interfaces, each image is filtered to remove atomic information then interpolated to four pixels per angstrom to reduce the ambiguity in the interface location. These images are then cropped to reduce distortion to acceptable levels and rotated to an angle specific to each interface in said image; these two steps were done deliberately so that all interfaces were of the same length and each successive measurement of the interface profile was equally spaced, conditions that do not affect a roughness amplitude measurement, but are quite necessary for the interface power spectra to be interpretable. The profiles are identified via masking the image on a set threshold and extracted via edge detection in each mask. The resulting profiles, are 350Å in length with four pixels for every nominal angstrom in both the lateral, <110> direction, and vertical, [001] direction.

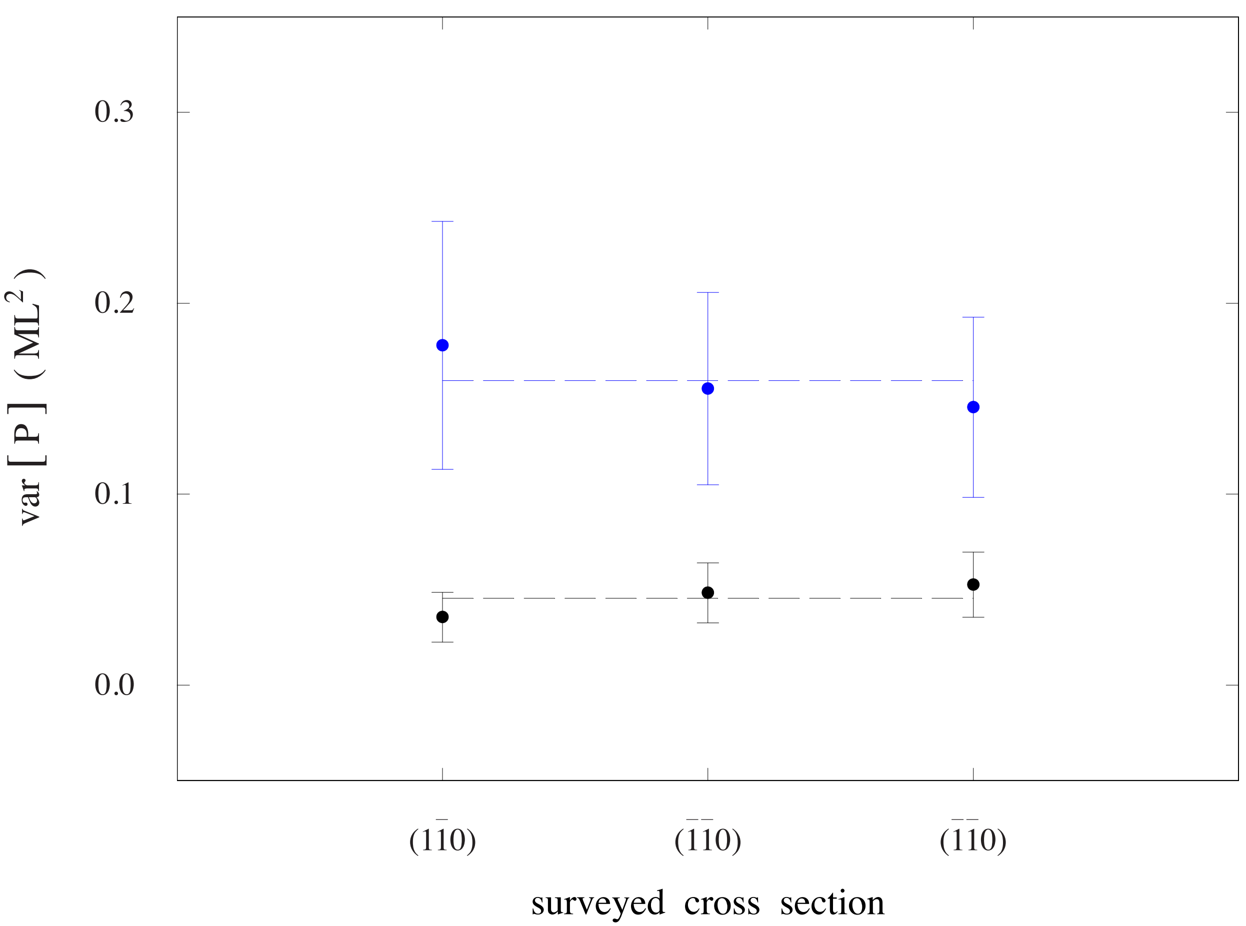


FIGURE 3.30. The variances of image–to–image fluctuations in Fig. 3.29 are tabulated for three device–scale surveys. The two (–1–10) surveys were taken 24 hours apart. Two crop sizes: 400 Å x 400 Å (black), 300 Å x 300 Å (blue) show an increase in variation as the image size is reduced.

To characterize the stochastic process underlying the interface roughness we start by computing the sampled roughness amplitude (3.3), which is easily visualized via a histogram of the [001] interface fluctuations about a locally sampled mean value (Fig. 3.31). Modeled as a normal distribution there is a small asymmetry between interface types with the aluminum template (0.83 ML) ~ 10% rougher than the gallium template (0.75 ML).

Next we compute the one–dimensional power spectra of our interface profiles. We choose not to zeropad the interfaces, in contrast to the image FFTs seen in Chapters II and III, which is especially important when characterizing low–frequency behavior. If zeros were added to the end of the profile in real space to increase the length by a factor of 40 such as we did with the images, then very low frequencies would be distorted in reciprocal space; in fact the first 40 points smoothly transition from zero power at k = 0 to the real spectrum, however the nature of that transition is not clearly defined and makes determining the functional form difficult. We instead give up the benefit of superior reciprocal space resolution to gain a power spectrum unbiased by the effects of zeropadding. As a further contrast to the image FFTs the total length in reciprocal space has actually been increased due to interpolating the interface in real–space (see Eq. 2.1b) thus mirroring the increase in real space length that is required to decrease the size of frequency bins in reciprocal space. The extra frequency bins are all located at high frequency where our atomic filter has removed any physical signal.

The interface power spectra, averaged over the survey ensemble and any combination of voltage, cross section, or interface type produces a nearly universal

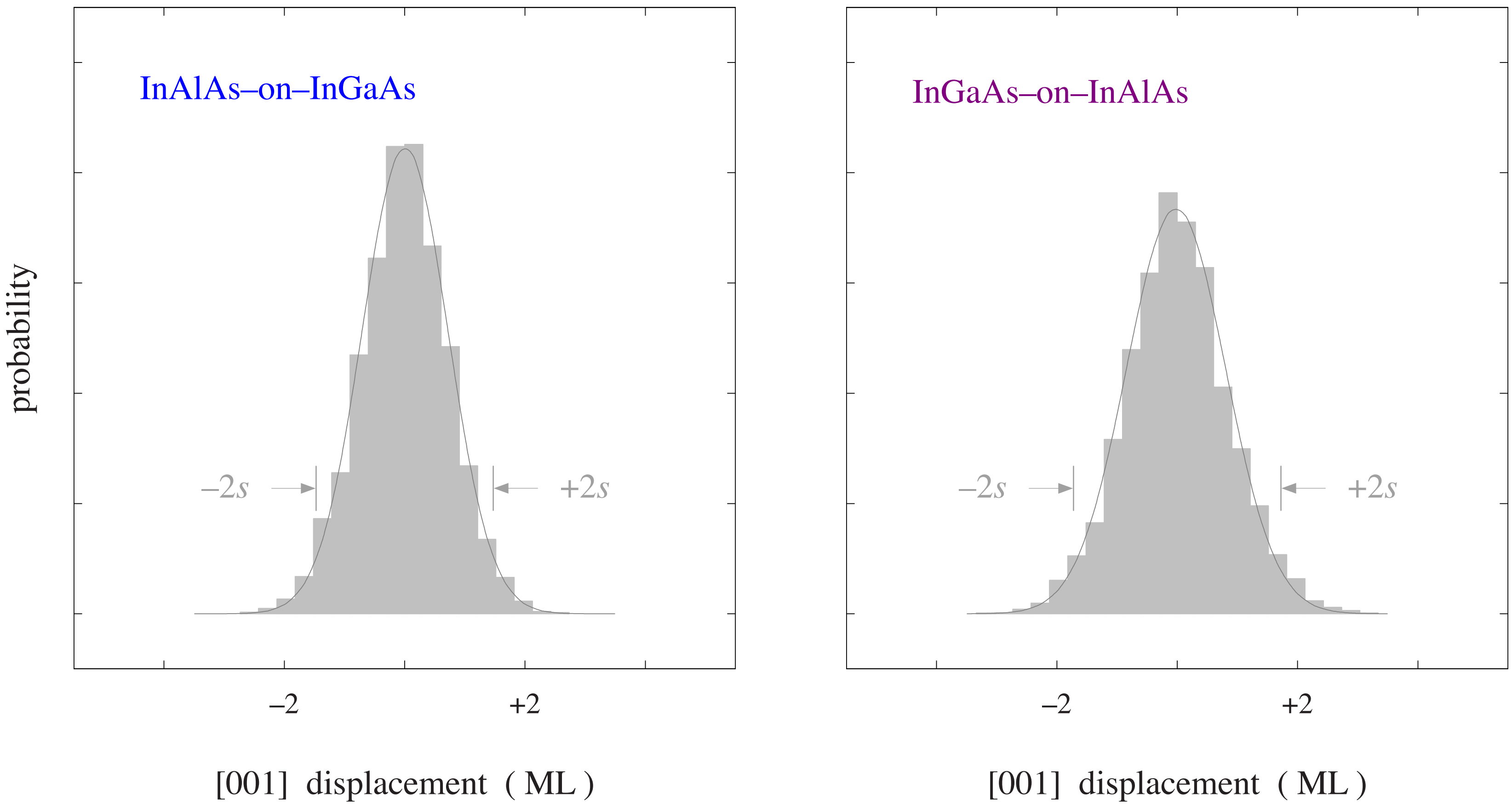


FIGURE 3.31. Distribution of vertical (z_i) fluctuations normalized to local [001] lattice constant and referenced to respective segment means from the (–1–10) cross section of the InAlAs–on–InGaAs, left, and InGaAs–on–InAlAs, right, interfaces. Sample variance (s^2) is calculated from the expectation value of (3.3). Distribution of these fluctuations are well described by a Gaussian model (grey). Three consecutive periods were pooled in each instance. Reprinted with permission from [39].

curve. This curve appears to show three distinct length scales, a low–frequency exponential–like and intermediate–frequency Gaussian–like component that show up as a straight line and a parabola respectively on a log–linear plot (Fig. 3.32, left), and a high–frequency inverse squared component that is linear on a log–log plot (Fig. 3.32, right). As we'll see shortly this third component is not real, it is in fact a byproduct of finite–length sampling.

To measure the correlation lengths we'll first develop and test a fitting strategy using simulated interfaces similar to those previously employed to define superlattice repeats, except the spacing between real–space datapoints will be decreased to 0.25 Å to mimic the extracted interfaces. When the profiles are repeatedly sampled over nonoverlapping lengths of 350 Å, then Fourier transformed and averaged together, they too contain an inverse square "Brownian Noise" component (Fig. 3.33) even though the simulated spectrum does not. Since the inverse square diverges at k = 0, we instead employ a Lorentzian function, which is an inverse square for large k, to approximate the Brownian Noise, thus the spectrum in Fig. 3.33 will be modeled as the sum of Gaussian and Lorentzian functions

$$PS = A \exp\left[-\frac{k^2}{2\Lambda_k^2}\right] + \frac{B}{1 + \left(\frac{k}{k_0}\right)^2}. \tag{3.19a}$$

The fitting function is left with too many degrees of freedom to pin down the Lorentzian cutoff frequency (k_0) or amplitude (B), so to constrain the parameter space that must be searched by the least squares fitting routine we constrain the amplitude of the Lorentzian

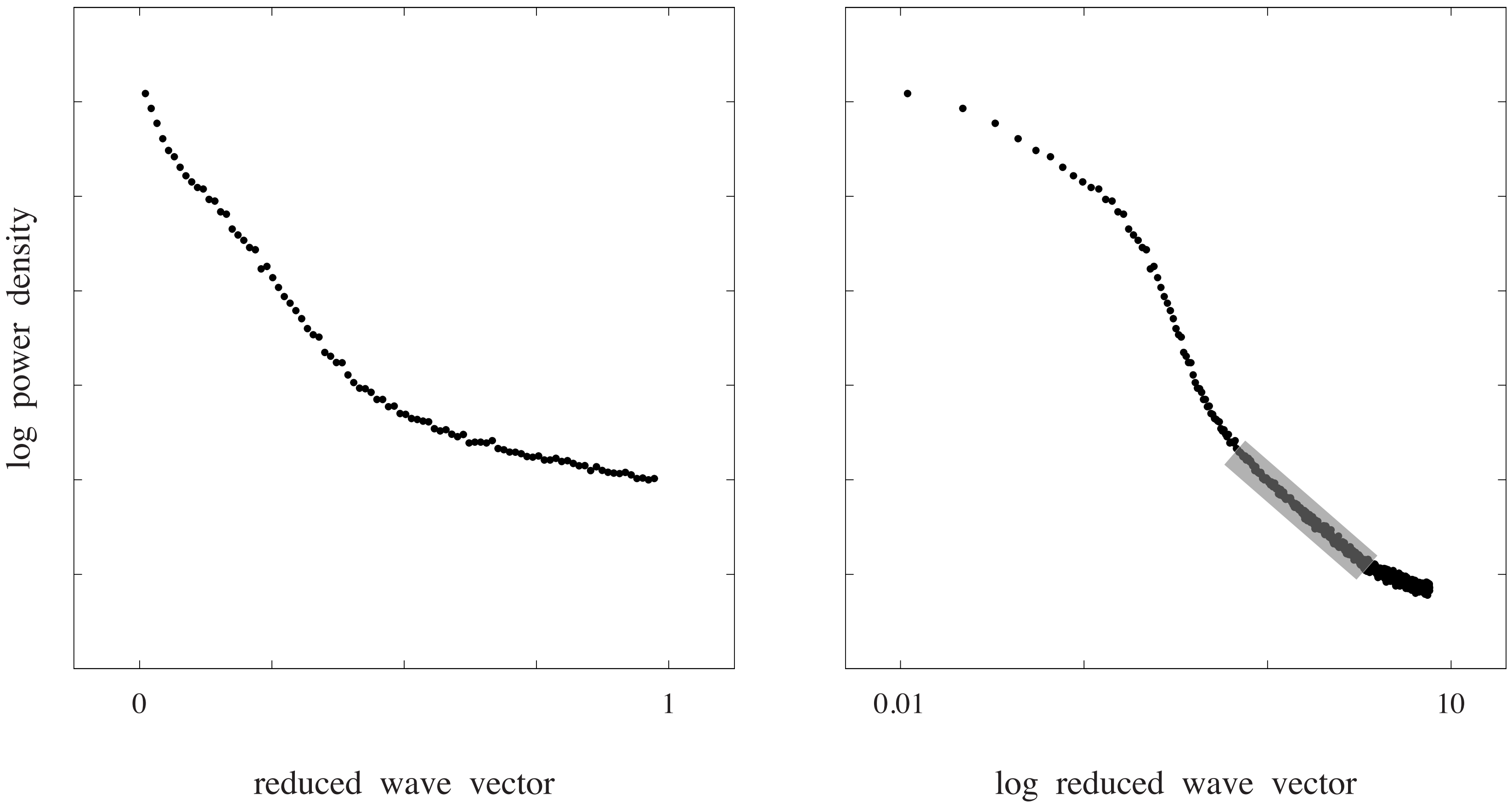


FIGURE 3.32. Cross section–, voltage–, and interface type–pooled roughness power spectrum. Three distinct length scales are distinguishable; the high–frequency component (highlighted with grey box) is identified as inverse square, “Brownian” noise.

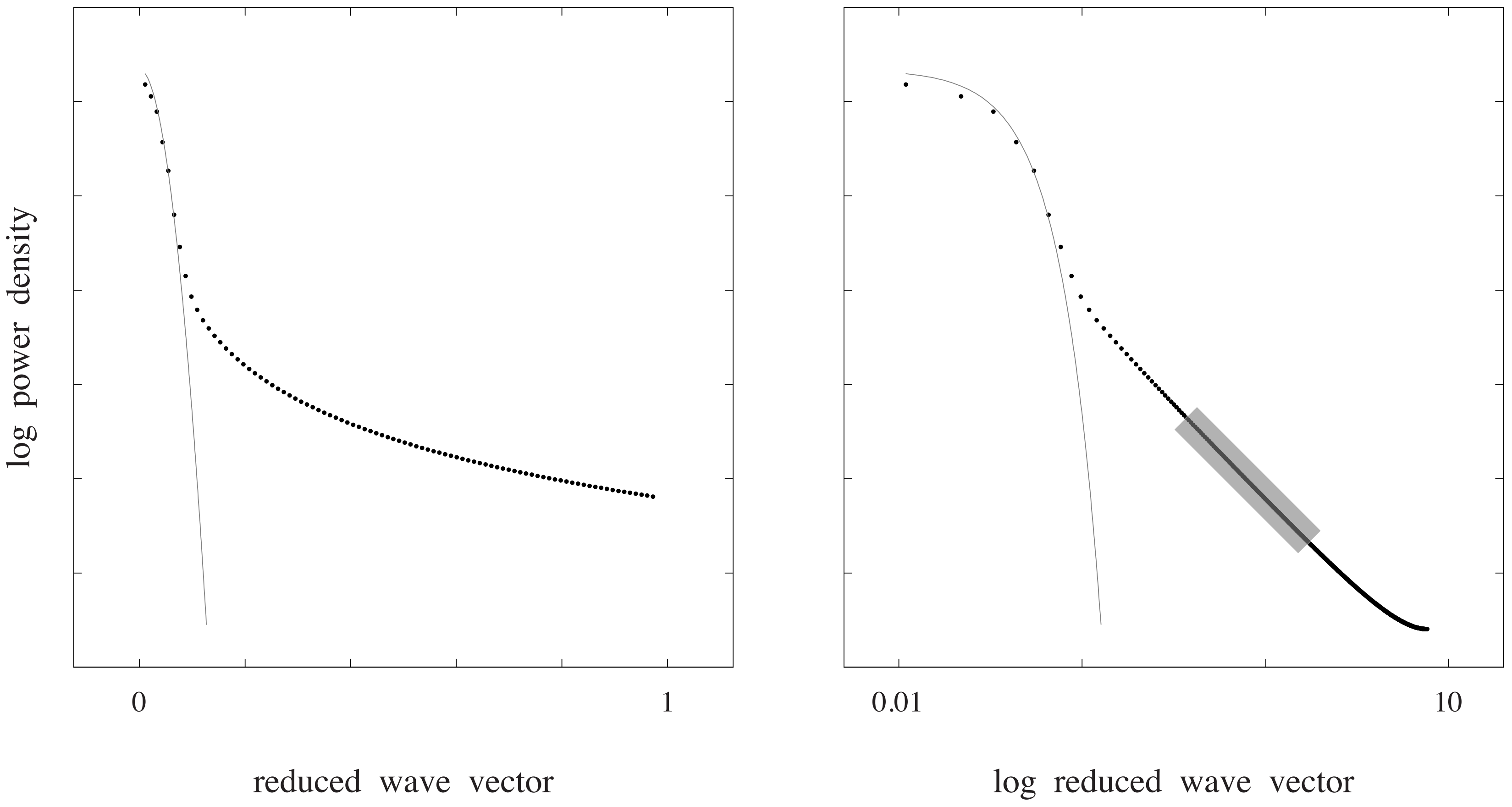


FIGURE 3.33. Power spectrum for profiles simulated with single Gaussian component. When calculated over entire profile (line) the power has the expected Gaussian form, whereas when the full profile is sampled at *N* point intervals, the power displays the same excess Brownian noise component observed in Fig. 3.32 (again highlighted by grey box).

to give the observed accumulated power between 1 and 2 inverse lattice constants, and the cutoff frequency equal to Λ_k.

This fitting function does a good job of describing the sampled power spectra of simulated profiles, shown in Fig. 3.34 (left) for a correlation length of 25 Å. The real–space correlation length is calculated from the fit Gaussian width via (3.8), and the variance is calculated from the sum of areas under the Gaussian and Lorentzian components. The resulting relative uncertainties are ~ 5 % in the real–space correlation length and ~ 0.5 % in the variance.

Generalizing the fitting function to two Gaussian components

$$PS = A \exp\left[\frac{-k^2}{2\Lambda_{k\_1}^2}\right] + C \exp\left[\frac{-k^2}{2\Lambda_{k\_2}^2}\right] + \frac{B_1}{1+\left(\frac{k}{\Lambda_{k\_1}}\right)^2} + \frac{B_2}{1+\left(\frac{k}{\Lambda_{k\_2}}\right)^2} + w \qquad (3.19b)$$

again shows good agreement between the fit and the simulated data (Fig. 3.34, right). The Brown noise is now represented by two Lorentzian components (each associated with a single Gaussian component) whose sum is constrained to give the accumulated power between 1 and 2 inverse lattice constants and whose weights are given by the proportion of the respective Gaussian components. Similar to the error analysis performed for correlation lengths extracted from period variances, we can compare the multi–component fit parameters and simulated counterparts (Fig. 3.35, left) and

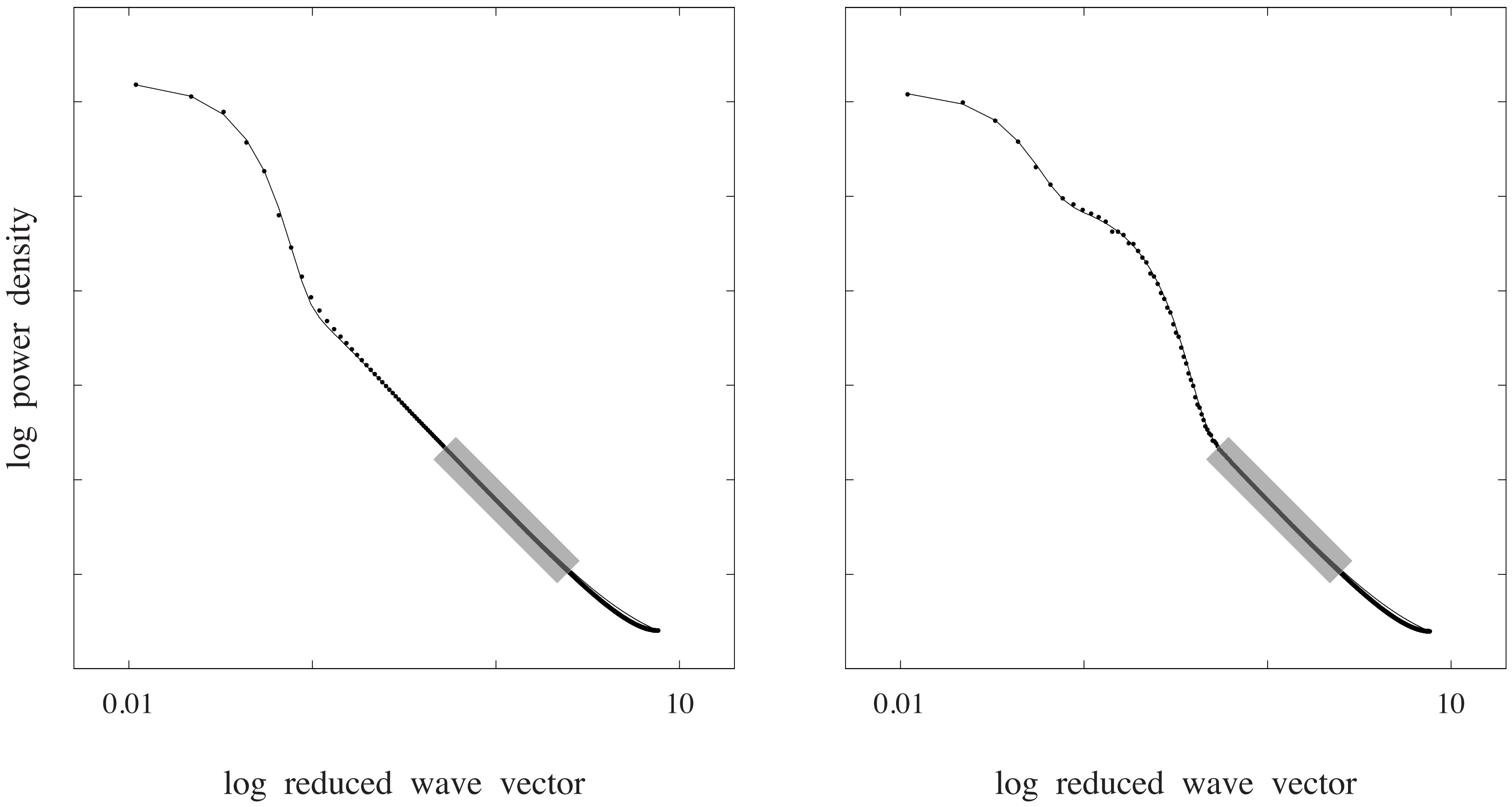


FIGURE 3.34. Single–component power spectrum reproduced from Fig. 3.33 (right) together with power spectrum from profile containing intentionally introduced second correlation length. Modeling the Brownian noise as a Lorentzian enables fits (lines) to the power spectra using (3.19a) or (3.19b) for single (left) and multiple length–scales (right) respectively.

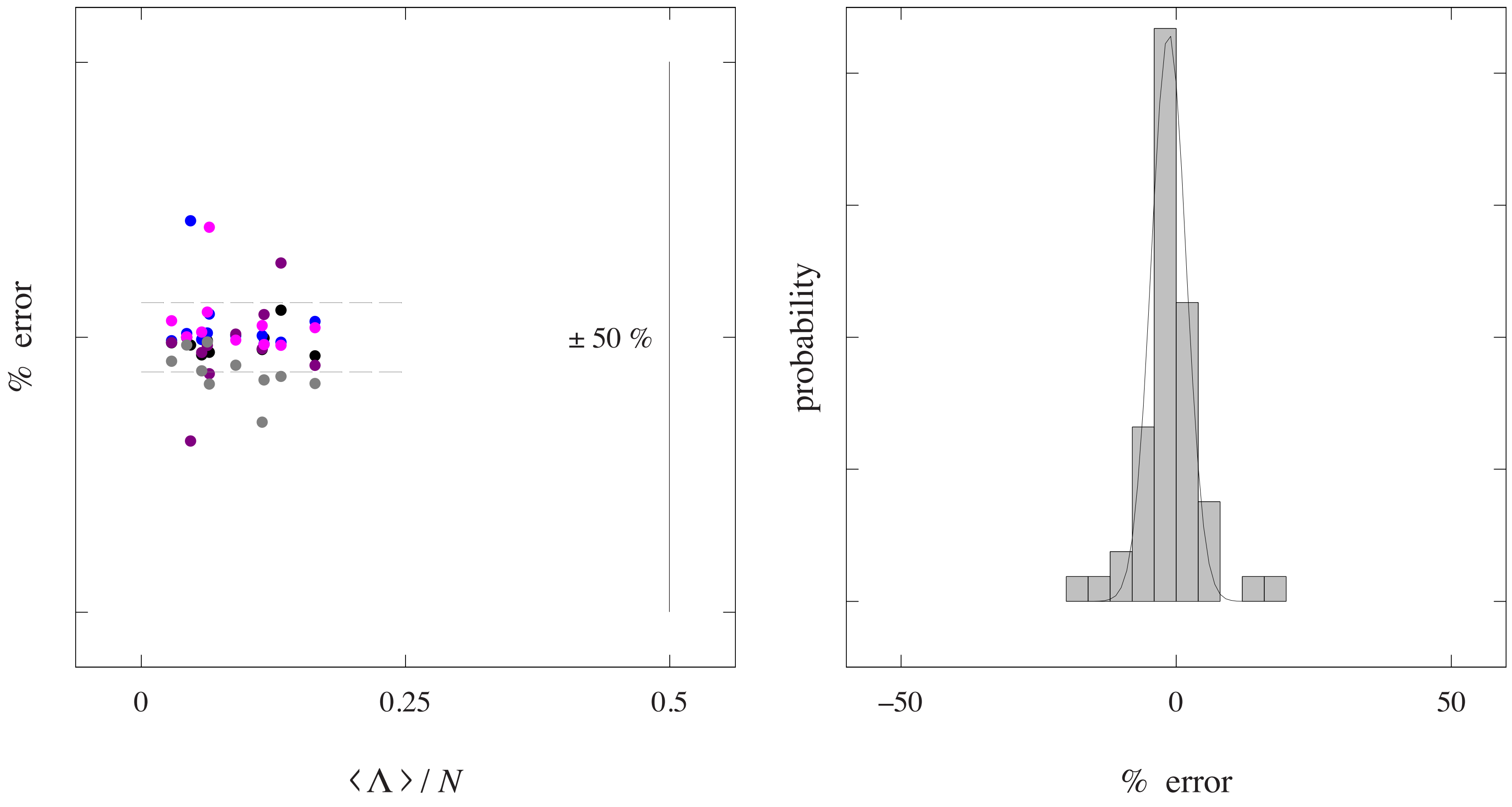


FIGURE 3.35. Relative error between fit parameters and known simulated values (left) for multi–scale Gaussian power spectra (black and grey points are length scales one and two, light and dark pink are relative weighting of two length scales, and blue is variance). A histogram of these values (right) is an approximate stand–in for the uncertainty in any given measurement. Inferred 95% confidence interval given by ± 6.3% is illustrated as dashed lines in left panel. Simulations underestimate the actual correlation length by ~ 1%.

histogram their relative errors[13] (Fig. 3.35, right) to determine each measurement has an uncertainty of 6.3% with an underestimation of ~ 1%.

Modeling the ensemble averaged power spectra similarly yields a breakdown of 40% Gaussian component with a correlation length of 5.96 Å and 60% Exponential power with a corresponding Lorentzian correlation length of 27.9 Å as shown in Fig. 3.36.

Comparison between Theory and Experimental Period Variance

The revelation that the interface power spectrum contains multiple components, and that one of said components is exponential requires rethinking our expectations for the variations of the measured periods. To consider the implications for exponential power we use a Lorenztian correlation function (the real space complement to an exponential in reciprocal space)

$$\rho_i = \frac{1}{1 + \frac{i^2}{\Lambda_x^2}} \tag{3.20}$$

in place of (3.10) when calculating (3.18). The empirical theory, after ignoring the difference between N and N–1, is given by

---

[13] All parameters were considered jointly to increase the statistical pool over which the histogram in Fig. 3.35 was calculated. This inherently assumes that all parameters have the same relative uncertainty. To test this assumption many more trials would be necessary.

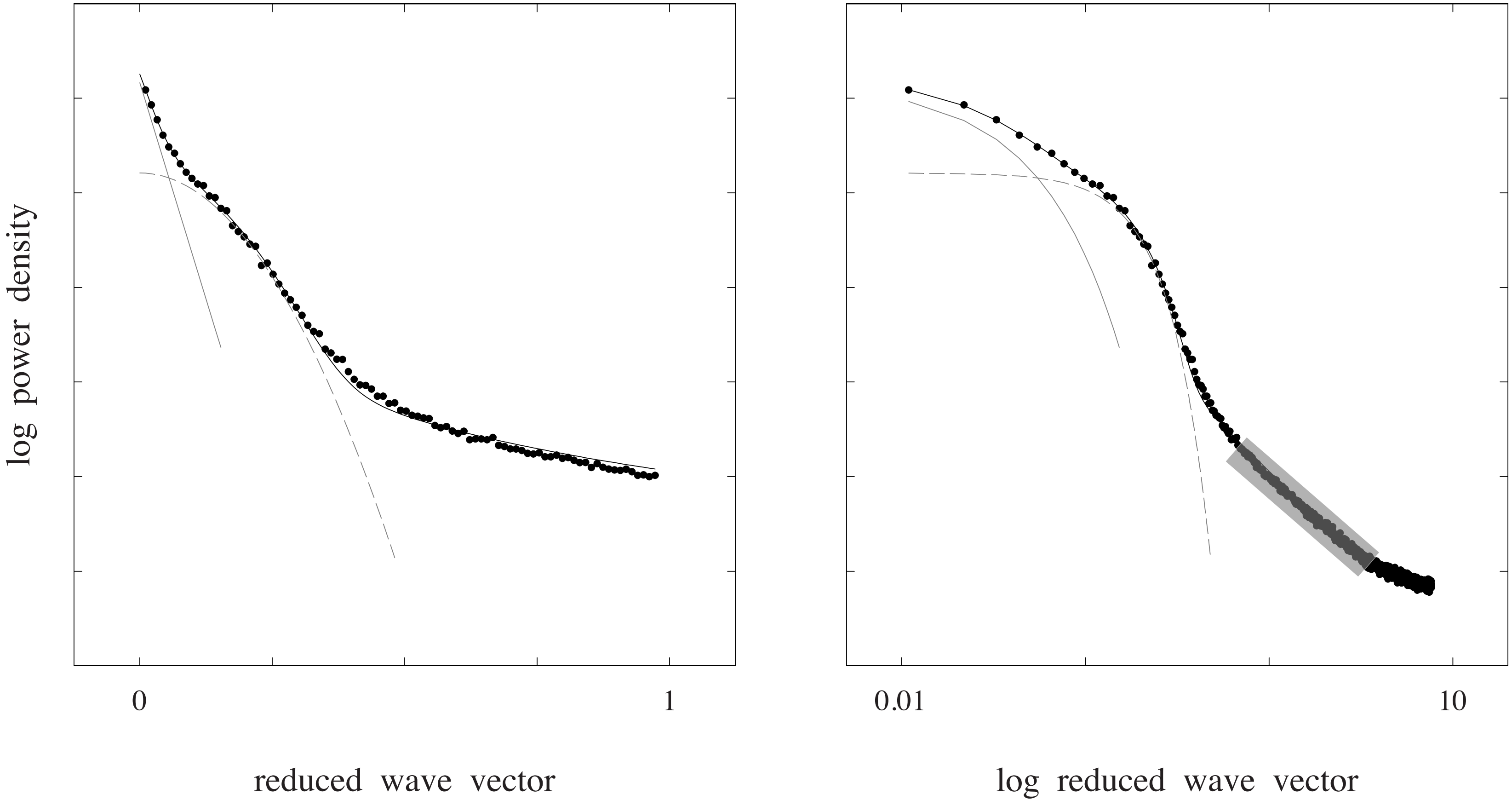


FIGURE 3.36. Fits to the power spectrum shown in Fig. 3.32 reveal a low–frequency exponential component (solid grey line), a high–frequency Gaussian component (dashed grey line) and very high frequency Brownian sampling noise (grey box).

$$\frac{\mathrm{var}[P]}{s^2}\left(\frac{N}{P}\right)^3 = \frac{4\frac{\Lambda}{N}\mathrm{ArcTan}\left[\frac{N}{\Lambda}\right] - 2\left(\frac{\Lambda}{N}\right)^2 \ln\left[1+\left(\frac{N}{\Lambda}\right)^2\right]}{1 - 2\frac{\Lambda}{N}\mathrm{ArcTan}\left[\frac{N}{\Lambda}\right] + \left(\frac{\Lambda}{N}\right)^2 \ln\left[1+\left(\frac{N}{\Lambda}\right)^2\right]} 4\left(\sqrt{\frac{\Lambda}{N}}\right) \tag{3.21}$$

for the window aligned at 45°. When the Lorentzian empirical theory is added to Fig. 3.26 we see that there is very little difference between theories obtained using Gaussian or Lorentzian correlation functions (Fig. 3.37), and the simulations themselves could easily be mistakenly attributed to Lorentzian correlations instead of to the Gaussian correlations from which they came. This justifies our use of Gaussian correlated profiles as a stand–in for Lorentzian correlated profiles. Also of note is that any linear combination of the two necessarily lies in between the two curves, therefore the logical Λ to plot on the x–axis is the probability weighted mean[14]

$$\langle\Lambda\rangle = P_1\Lambda_1 + P_2\Lambda_2\,. \tag{3.22}$$

Before the values directly measured in the last section can be added to the graph in Fig. 3.37, we must make some simplifications. First, even though our model assumes that there is no interface roughness asymmetry, we average the sampled roughness amplitude, s^2, across the the Al– and Ga–templates to obtain an "effective" roughness amplitude; additionally, we adjust the sampled roughness amplitude for the change in

---

[14] In a multi–component correlation function with mixed functional forms this quantity won't show up in the sampled variance of the mean or the sampled s^2, however it is still helpful to use as a representative correlation length for the combined processes.

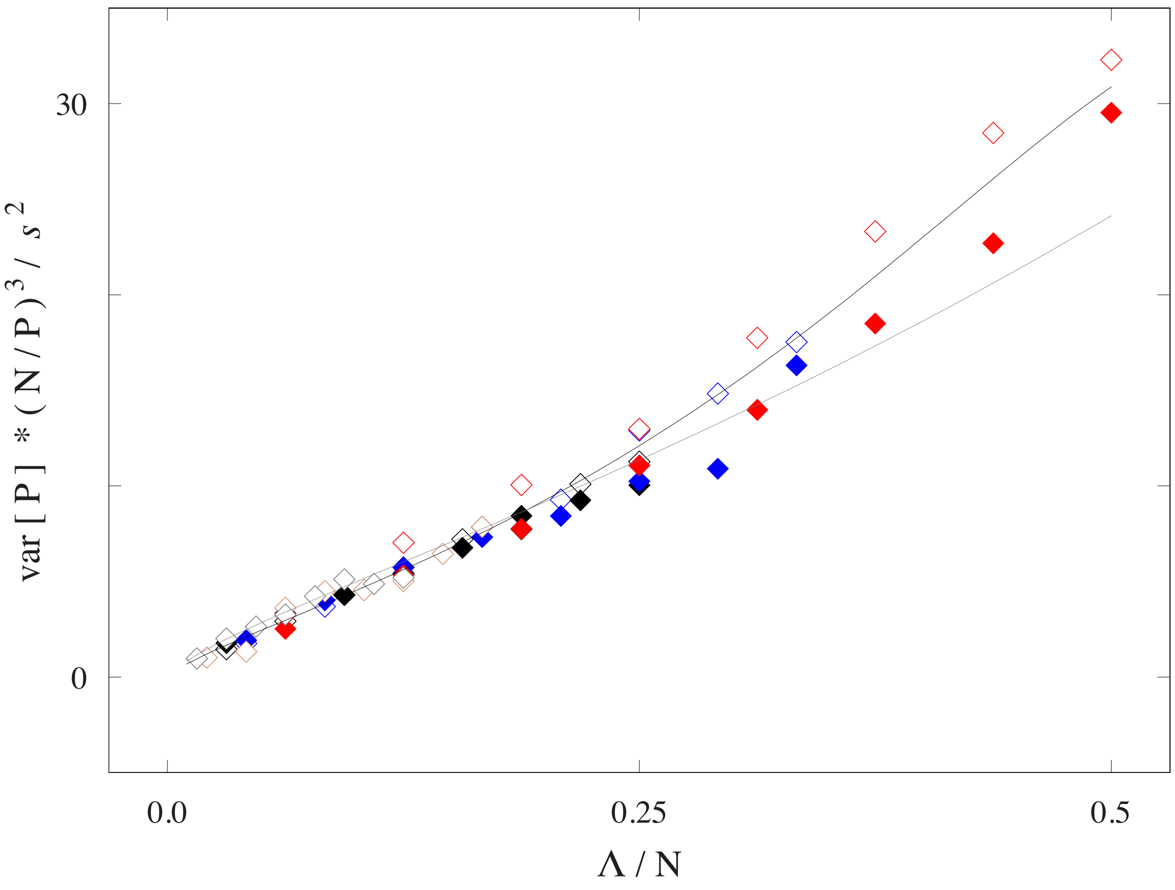


FIGURE 3.37. Rotated–window period variances reproduced from Fig. 3.26 with additional empirical Lorentzian theory (grey line) calculated from (3.21).

sampled length since it was measured over an interface segment of 350 Å. We use the theoretical relationship (3.13) to predict s^2 for 300 Å and for 400 Å sampling lengths. The period fluctuations in the InAlAs / InGaAs superlattice (Fig. 3.38) agree extremely well with our expected curves calculated from (3.18) and (3.21). The 400 Å window lies directly between the Gaussian and Lorentzian theories and 300 Å window lies within one standard deviation.

To justify the extra work that was expended understanding and correcting for windows that lie at 45° to the crystal axes, we consider how well the vertical window fluctuations in Fig. 3.17 describe the two InAlAs / InGaAs data points, which were just placed on the graph in Fig. 3.38. Fig. 3.39 shows the 400 Å and 300 Å window sizes are three standard deviations away from the Gaussian theory, reinforcing the need to correctly account for the orientation of the sampling window in these calculations.

Finally we return to the question that motivated all of this, namely what interface correlations might be present in the InAs / InAsSb superlattice. We've established that the period fluctuations themselves aren't sufficient to accurately determine both the roughness amplitude and the correlation length, we instead endeavor to put limits on both quantities. We aren't able to define interface profiles in the InAs / InAsSb superlattice and therefore don't have a direct measurement of s^2; we instead return to scaling the period variances by σ^2 in Fig. 3.40 and in the discussion of limiting roughness that follows.

Like Fig. 3.15, Fig. 3.40 appears to approach a constant value for large correlation lengths; since we don't have a theoretical counterpart to (3.9) we aren't able

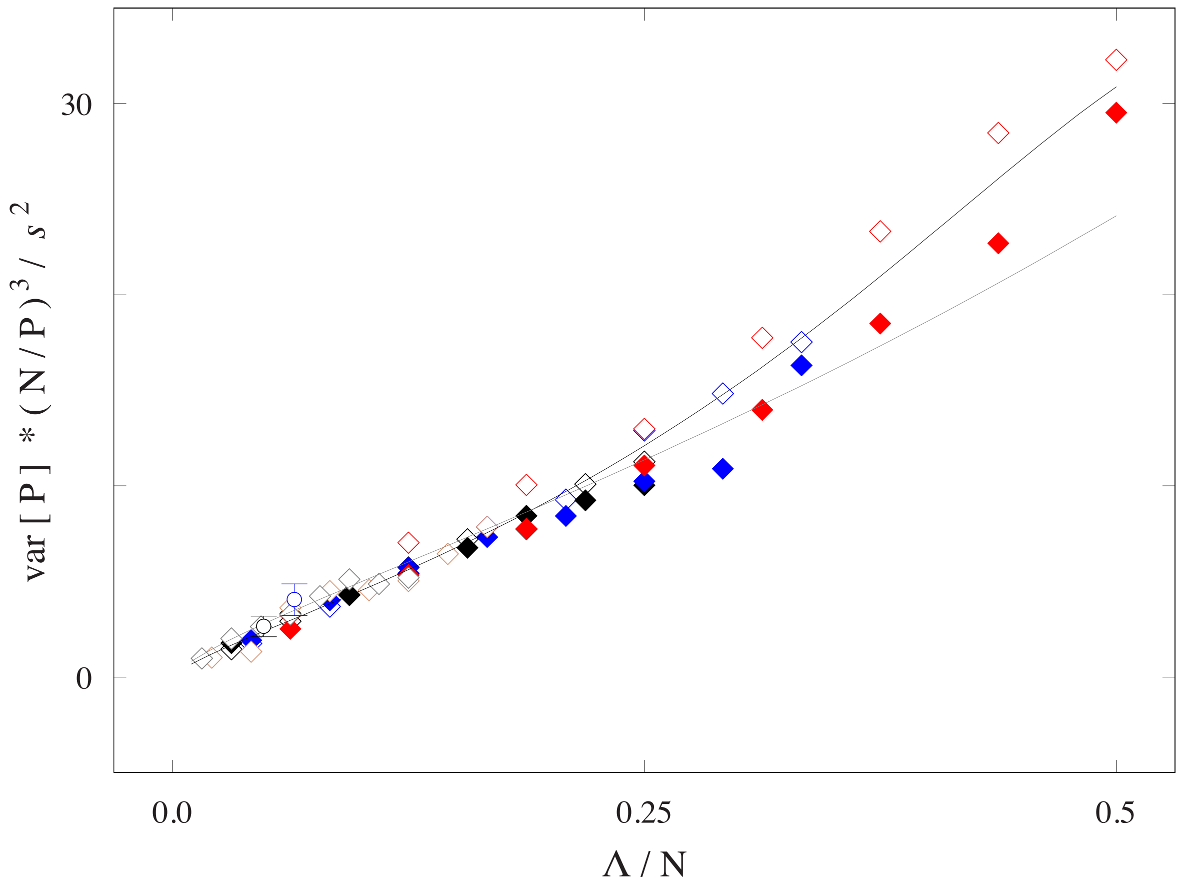


FIGURE 3.38. Rotated–window period variances with Gaussian and Lorentzian theories reproduced from Fig. 3.37 with experimental values independently measured on InAlAs / InGaAs superlattice (open circles).

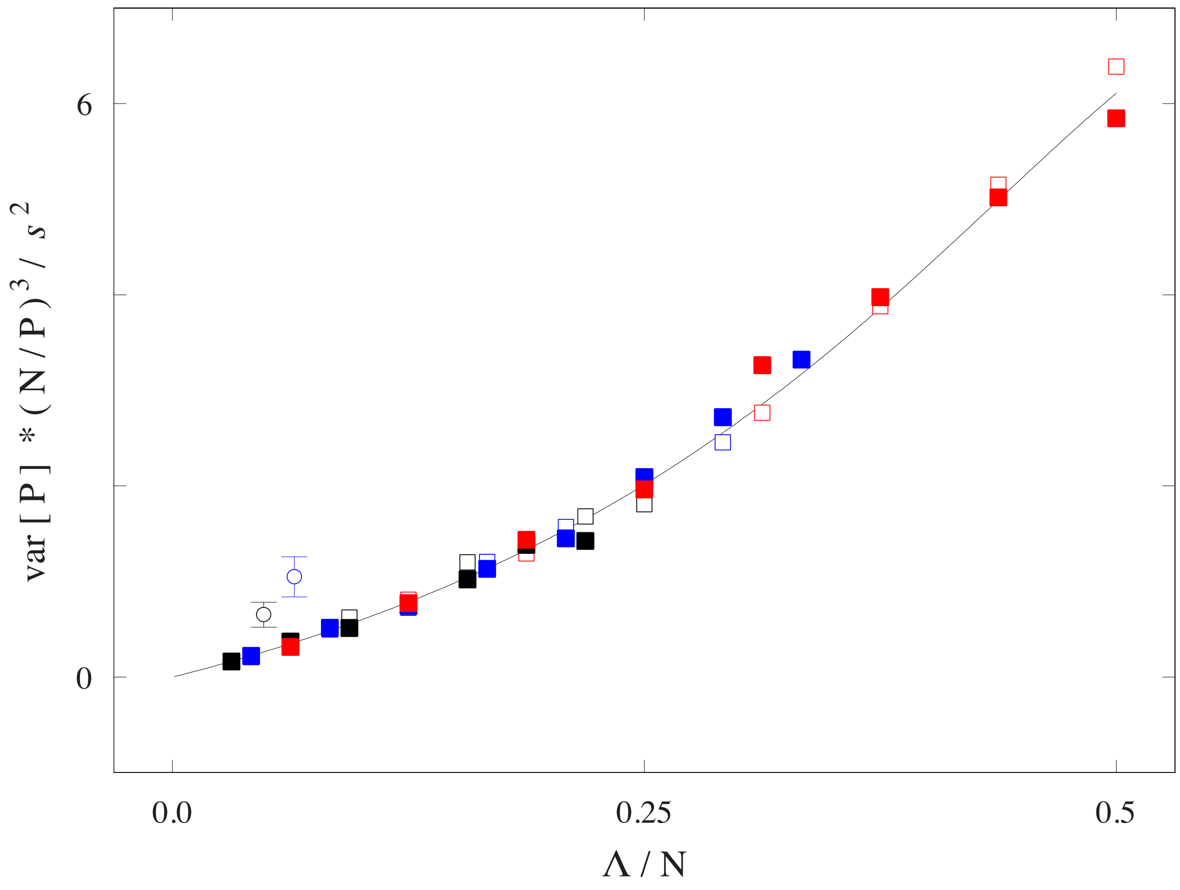


FIGURE 3.39. When experimental InAlAs / InGaAs values from Fig. 3.38 are compared to vertical–window period variances from Fig. 3.17 they no longer agree, thus demonstrating the need for rotated–window treatment of period variances.

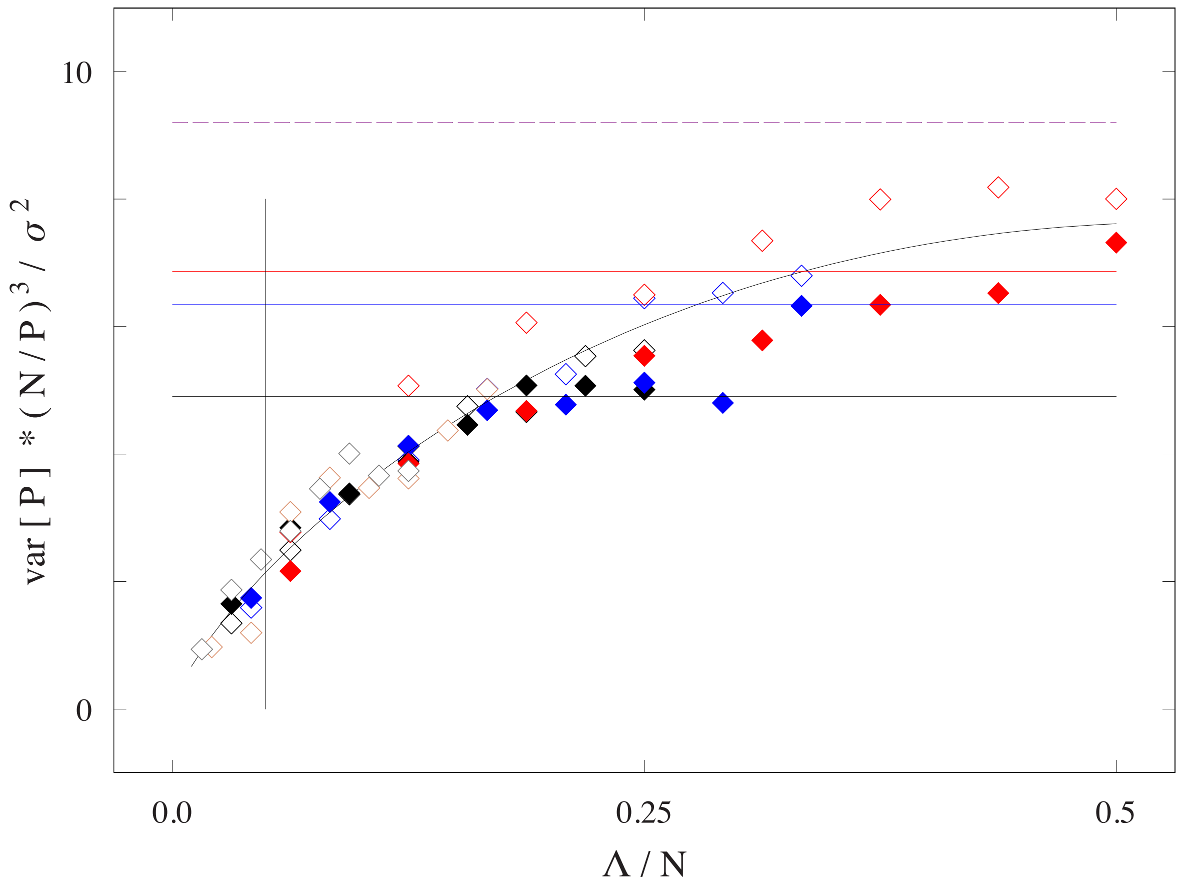


FIGURE 3.40. Rotated–window period variances scaled to parent roughness asymptote at 9.5 (dashed line) setting a minimum roughness (σ=0.6 ML) that explains period fluctuations in the InAs / InAsSb superlattice. Were the correlation length to be 20 Å (vertical black line), same as the mean length in InAlAs / InGaAs, then the roughness would be ~ 1.4 ML. If the roughness were instead 0.85 ML as in the InAlAs / InGaAs material, the correlation length would be ~ 70Å (reflected differently in each window size shown by the horizontal lines).

to treat the rotated windows analytically this time and instead find an average of 9.2 from simulations of a 1000Å correlation length over 2 periods, 3 roughness amplitudes, and 3 window sizes. This leads to a minimum σ of 0.6 ML. If we instead assume the InAs / InAsSb superlattice has the same correlation length as the average one (19.2 Å) found in the InAlAs / InGaAs superlattice, we find that σ would be 1.4 ML. Similar constraints can be placed on the correlation function by considering the example from earlier in the chapter, if $\sigma = 1$ ML then $\Lambda = 36$ Å, or if the roughness amplitude was the same as in the InAlAs / InGaAs superlattice, $\sigma = 0.85$ ML, then $\Lambda = 70$ Å

Summary

We calculated the period in 40 nm x 40 nm windows by referencing superlattice satellite peaks to the local [001] reciprocal lattice vector. A small, but reproducible drift in laterally–averaged periods was observed over similar vertical repeats in orthogonal cross sections on different dies. The drift was confirmed via fits to the laterally–averaged [001]–monolayer–indexed antimony fraction compiled from nominally the same area when the period is included as a fit parameter to the well–established segregation model. A careful analysis of the HRXRD spectrum also indicates the presence of the two STM–identified periods. Image–to–image period fluctuations as the structure was surveyed in the lateral (i.e. [1–10] or [110] directions) are larger than can be explained by uncorrelated interfaces. Treating the period as the difference between mean interface locations separated by *M* repeats (enclosed by 2*M*+1 interfaces) lets us construct a theory relating the period variance to the correlation length and window size in both vertical (crystal–aligned) and rotated (*x*–*y* image–aligned) sampling windows. An InAlAs / InGaAs superlattice, where interface roughness amplitudes and correlation lengths are directly measured, provides the experimental benchmark proving that the period fluctuations are due to interface roughness. We found that period fluctuations alone do not provide strong enough constraints to extract roughness amplitudes and correlation lengths, however limits in the InAs / InAsSb superlattice place the roughness amplitude larger than 0.6 ML and the correlation length likely between 20 and 70Å. More importantly, we now have a nomograph that, should we be able to independently measure an interface s^2, directly determines the interface roughness correlation length.

CHAPTER IV

SHORT–RANGE ALLOY ORDER

Introduction

Having established, in Chapter III, that the interfaces between InAs/InAsSb superlattice constituents are correlated, we return on the atomic scale and ask whether the distribution of individual antimony atoms throughout the InAsSb alloy layer, itself, might also be correlated. Early STM experiments [26] pinpointed local arrangements of antimony atoms consistent with CuPt order, but unable to assess their likelihood relative to a random alloy of similar composition. Long–range CuPt order has been definitively identified in InAsSb alloy films [68] with TEM, and companion studies have shown this order depends, at least in part, on the amount of [001] strain in the epitaxial film [28].

To investigate whether a structure displays CuPt order we calculate a correlation function, which is directly related to the likelihood for a specified arrangement of atoms relative to the same arrangement occurring in a random distribution. The pair correlation function is the second in a family of particle distribution functions (the first being the average density) and represents two–body interactions. It is commonly used to describe gasses [69], where it is a function of a continuous separation vector; for semiconductor alloys, however, the crystalline lattice must be taken into account by restricting the separation vector connecting any two sites in the anion sublattice to integer multiples of [001] and <110> lattice vectors and linear combinations thereof. In order to calculate the correlation function between pairs of antimony atoms in the InAs / InAsSb superlattice we need to be able to assign coordinates to each atom in the image. With a perfect tripod scanner the *x* and *y* image coordinates could be used to calculate the desired separation vectors. Given the image distortions documented in Fig. 2.13 and Fig. 2.14, we must instead convert to [001] column and <110> row coordinates.

We begin this chapter by examining the experimental evidence for antimony–pair correlations. We then detail an algorithm which can automatically assign [001] column and <110> row coordinates to each atom in a given image window effectively sidestepping the effect of bowing on these rows and columns. We explain how these coordinates are used to calculate a pair correlation function in the context of a uniform, isotropic bulk alloy, and then adapt this formalism to tackle a spatially–graded superlattice. To understand the origin of the antimony pair correlations so observed we partition the antimony population into sub–ensembles based on two experimentally accessible parameters: MBE shutter timing and [001] layer strain. We investigate the behavior of these sub–ensemble pair correlation functions in both (–1–10) and (1–10) cross sections, and find their logarithm exhibits a linear dependence on both strain and antimony fraction. These quantities are inextricably linked in any coherently–strained structure, but physically interpretable fits favor antimony fraction as the controlling variable. Comparison with the STM data from a fully–relaxed bulk alloy film, whose strain and antimony fraction are independent of one another, decisively supports this interpretation.

Experimental Evidence for Alloy Correlations

We begin this chapter by reflecting on the experimental evidence we have for atomic ordering in the InAs / InAsSb superlattice. Consider first a portion of the lateral survey from Fig. 2.26 and reproduced in Fig. 4.1. The superlattice is, by definition, modulated in the [001] growth direction, however looking along the [1–10] rows the InAsSb doesn't appear as featureless as a random distribution would imply. There are gaps of several lattice sites where no antimony is to be found. Elsewhere it seems there is much more antimony than the 33% targeted.

Focusing in on representative images from orthogonal cross sections (Fig. 4.2) we find an abundance of antimony atoms occupying next–nearest–neighbor lattice sites (encircled with blue) in (–1–10) cross section, and comparatively few antimony atoms occupying the nearest–neighbor sites (encircled with red). The situation is different in (1–10) projection, where antimony is more likely to incorporate at nearest–neighbor sites and comparatively few antimony atoms incorporate at next–nearest–neighbor sites.

The proper quantification of these insights will be presented later in the chapter, but as a naïve initial estimate we calculate the image autocorrelation function. To emphasize the antimony atoms we first process the images in Fig. 4.2, replacing any grey–scale value below a chosen threshold with the image mean (128 grey). The grey–scale values of 185 and 170 for (–1–10) and (1–10) respectively were chosen empirically with initial guesses based on a change in curvature of the image histogram, followed by fine tuning so that the antimony atoms look similar in both projections. Each resulting thresholded image (Fig. 4.3) consists of antimony atoms (primarily, but not exclusively

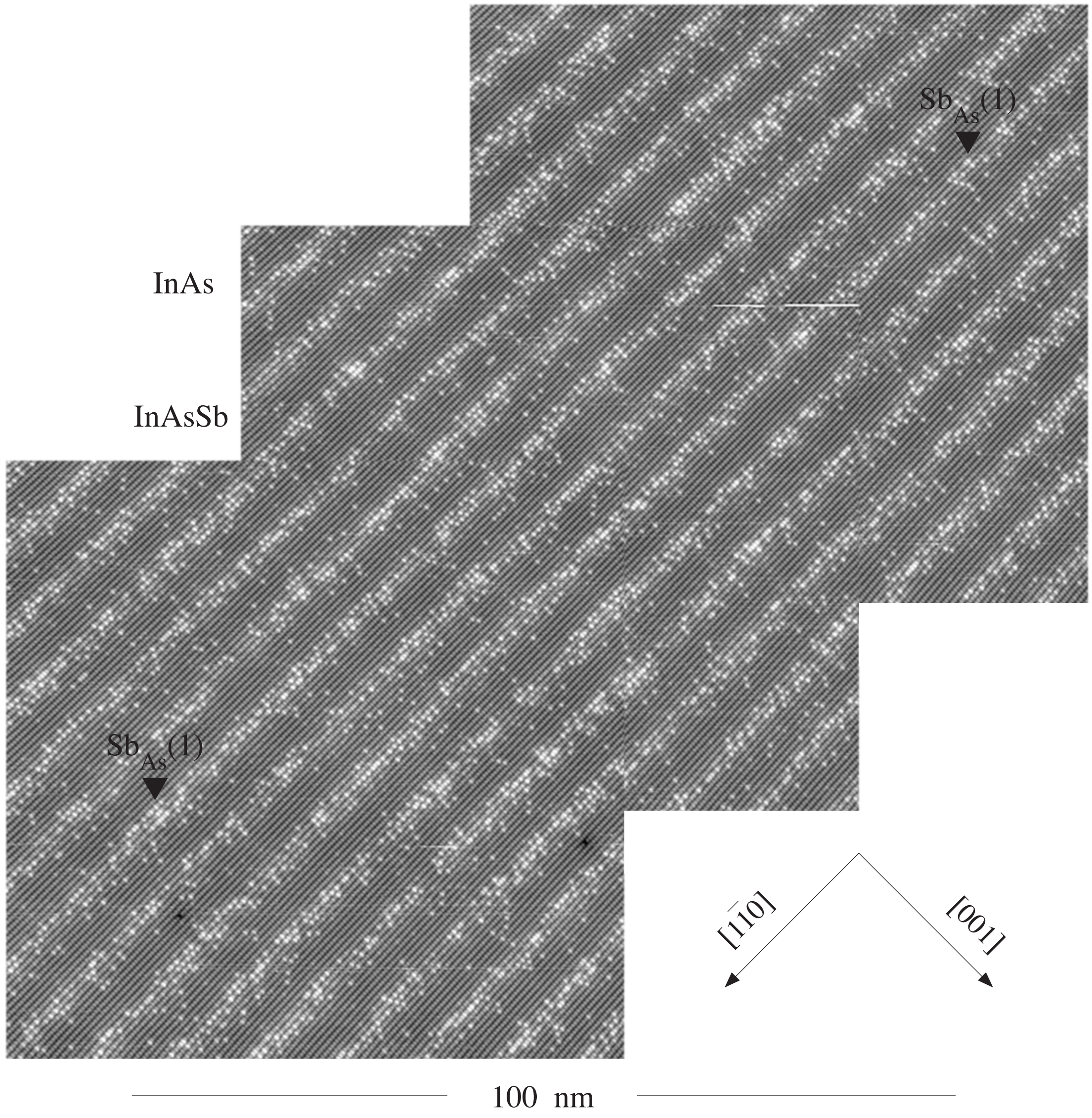


FIGURE 4.1. Section of lateral survey in Fig. 2.26. Bright layers are InAsSb, with carets indicating individual top–layer antimony–for–arsenic replacement. Growth direction is from top–left to bottom–right.

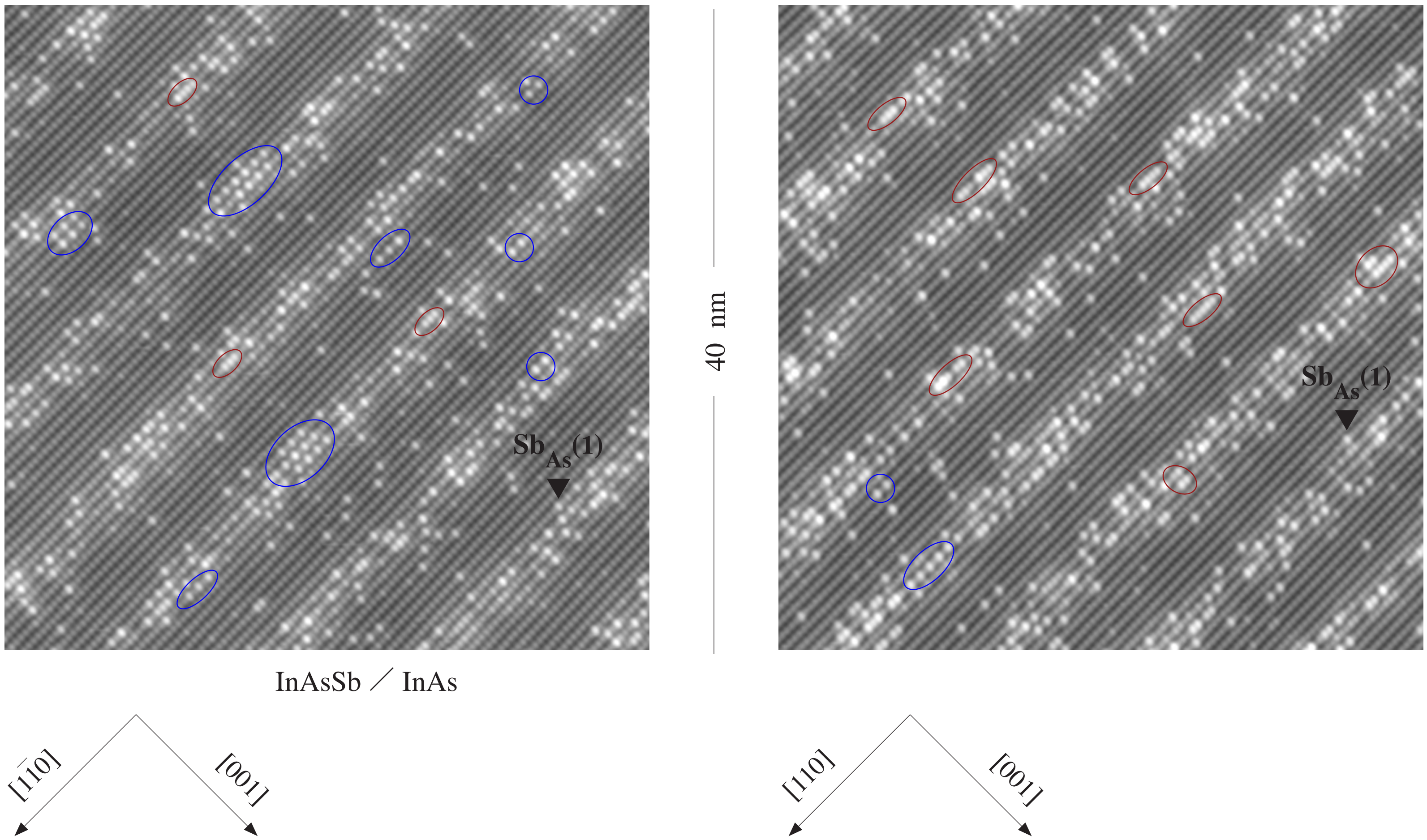


FIGURE 4.2. Individual images from Fig. 4.1, (–1–10) cross section is on the left and (1–10) is on the right. Bright sites are top–layer antimony atoms. Antimony–for–arsenic substitutions appear preferentially situated at next–nearest–neighbor anion sites (encircled in blue) when viewed in (–1–10) cross section, whereas nearest–neighbor anion sites (encircled in red), are preferred in (1–10) cross section.

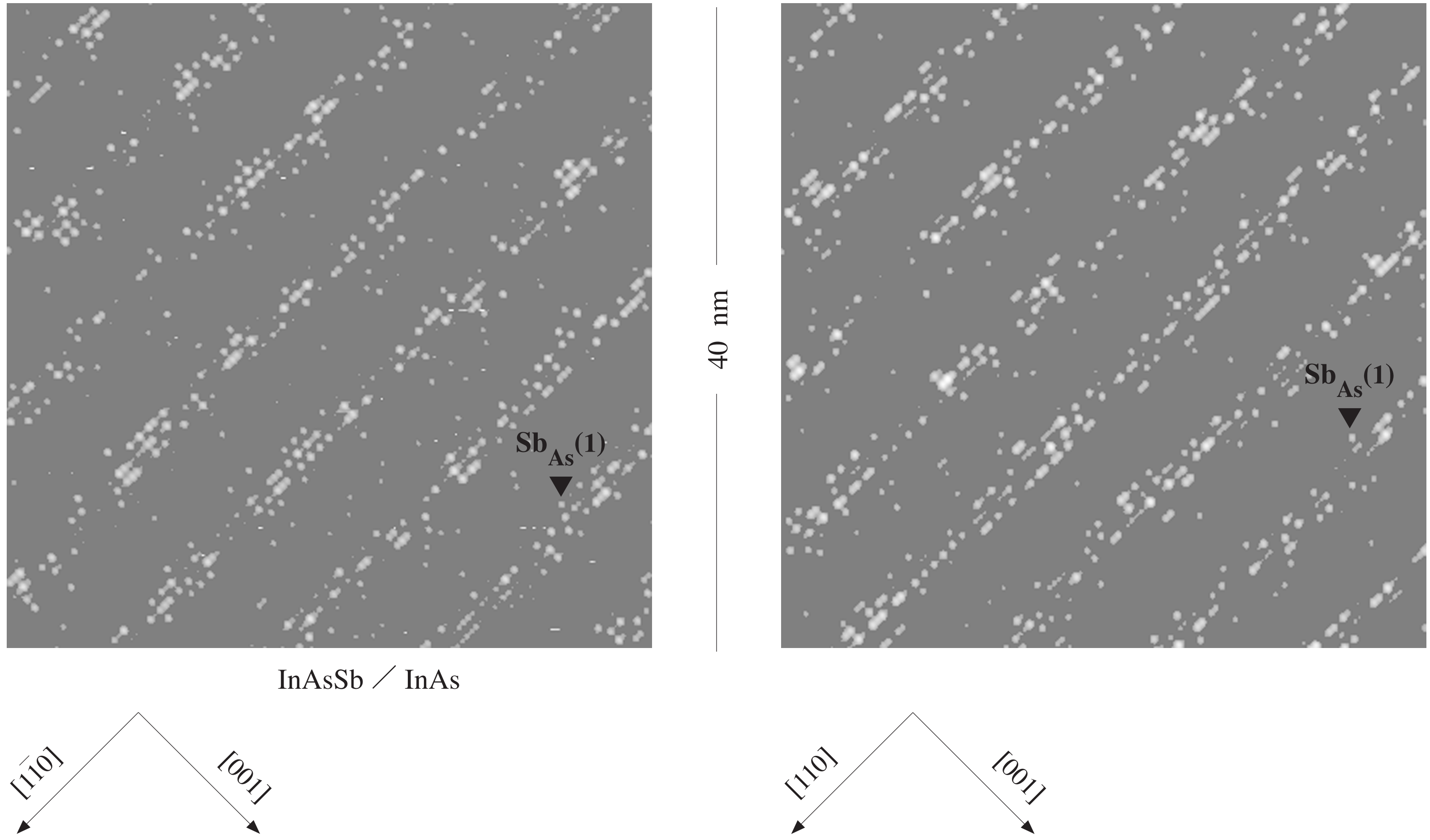


FIGURE 4.3. Individual images from Fig. 4.2 thresholded to emphasize top–layer antimony atoms. Grey values below the image threshold (175 and 180 in left and right images respectively) are set to middle grey (128).

limited to, top–layer) surrounded by a featureless plane. The high–density InAsSb layers, as well as the nearest–neighbor and next–nearest–neighbor spacing identified in Fig. 4.2 are apparent in these images. That the antimony atoms in the InAsSb layer are brighter than the antimony atoms in the InAs layer, is due to a small, but not negligible, electronic contribution to the contrast in the original (unthresholded) STM images.

We form a survey–averaged image autocorrelation (Fig. 4.4) by correlating image–by–image before averaging across the survey. Like the DFTs explored in Chapters II and III, the origin is located at the center of the autocorrelation. The feature that stands out the most in these 2D correlation maps are the strong bands corresponding to the InAsSb layers. More subtle is the evident structure in the [1–10] main diagonal in (–1–10) cross section (left) that is missing from the corresponding [110] diagonal in (1–10) cross section (right). The small–separation differences are more readily visualized by taking a <110> section through the origin. The subtle structure in Fig. 4.4 is now very obvious from the <110> autocorrelation sections in Fig. 4.5. The next–nearest–neighbor lattice spacing of antimony atoms encircled in blue in Fig. 4.2 shows up as strong correlations at even multiples of the <110> lattice constant in (–1–10) cross section (left), and weak correlations at odd multiples. In (1–10) cross section (right), on the other hand, there are increased correlations for nearest–neighbor separations, but a very weak response for all other lattice sites.

These correlations can also be viewed through the lens of DFTs calculated from images like those in Fig. 4.2 and then averaged over an entire survey. This is especially useful since the autocorrelation map (Fig. 4.4) is the Fourier transform of the DFT power

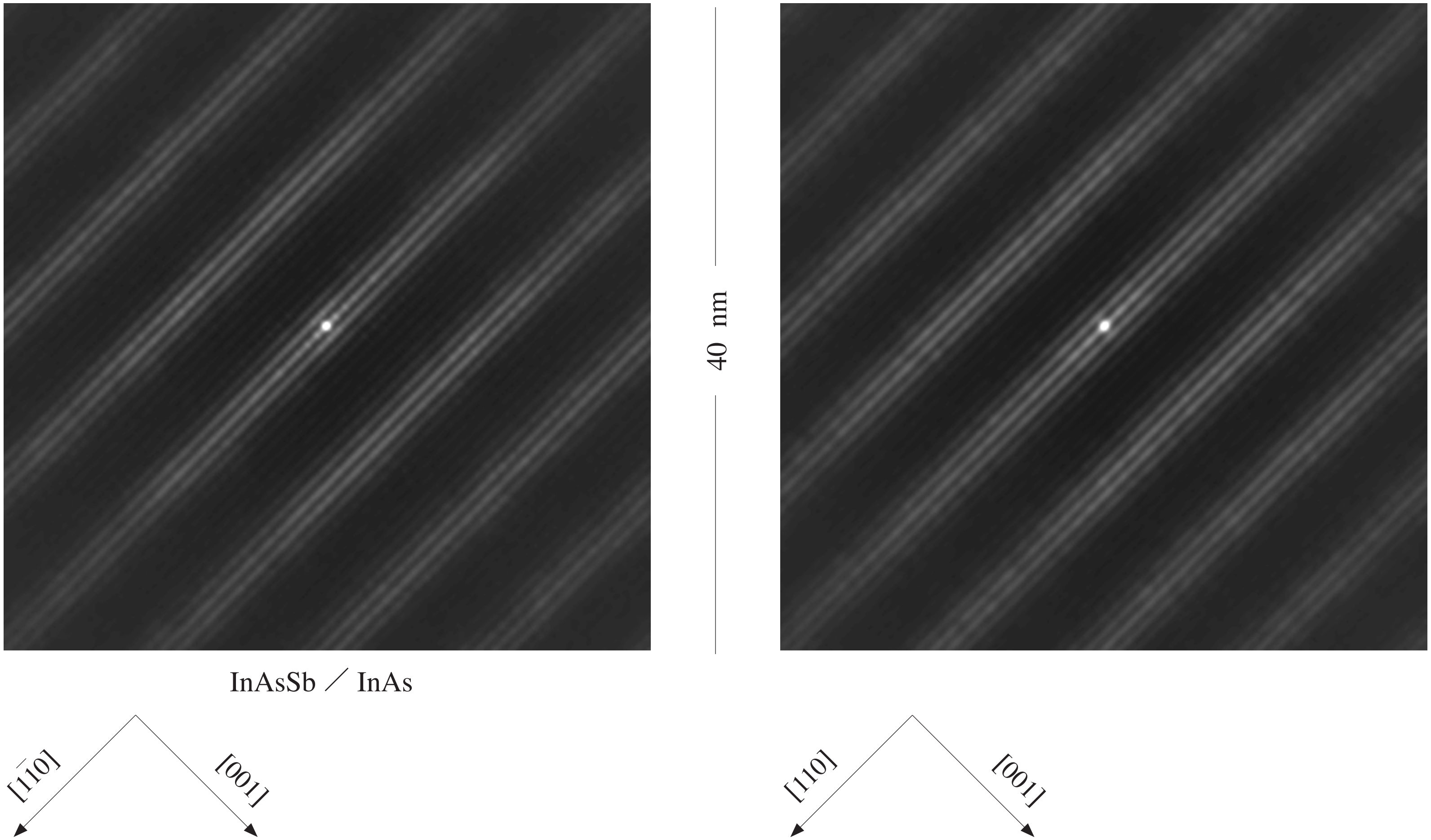


FIGURE 4.4. Two–dimensional autocorrelation maps corresponding to the thresholded images in Fig 4.3. The InAsSb layers show up as bands perpendicular to the growth direction. Along the main diagonal there is a bright spot at every lattice site, but in (–1–10) cross section (left) every other site appears slightly brighter, an effect which is absent on the (1–10) surface (right).

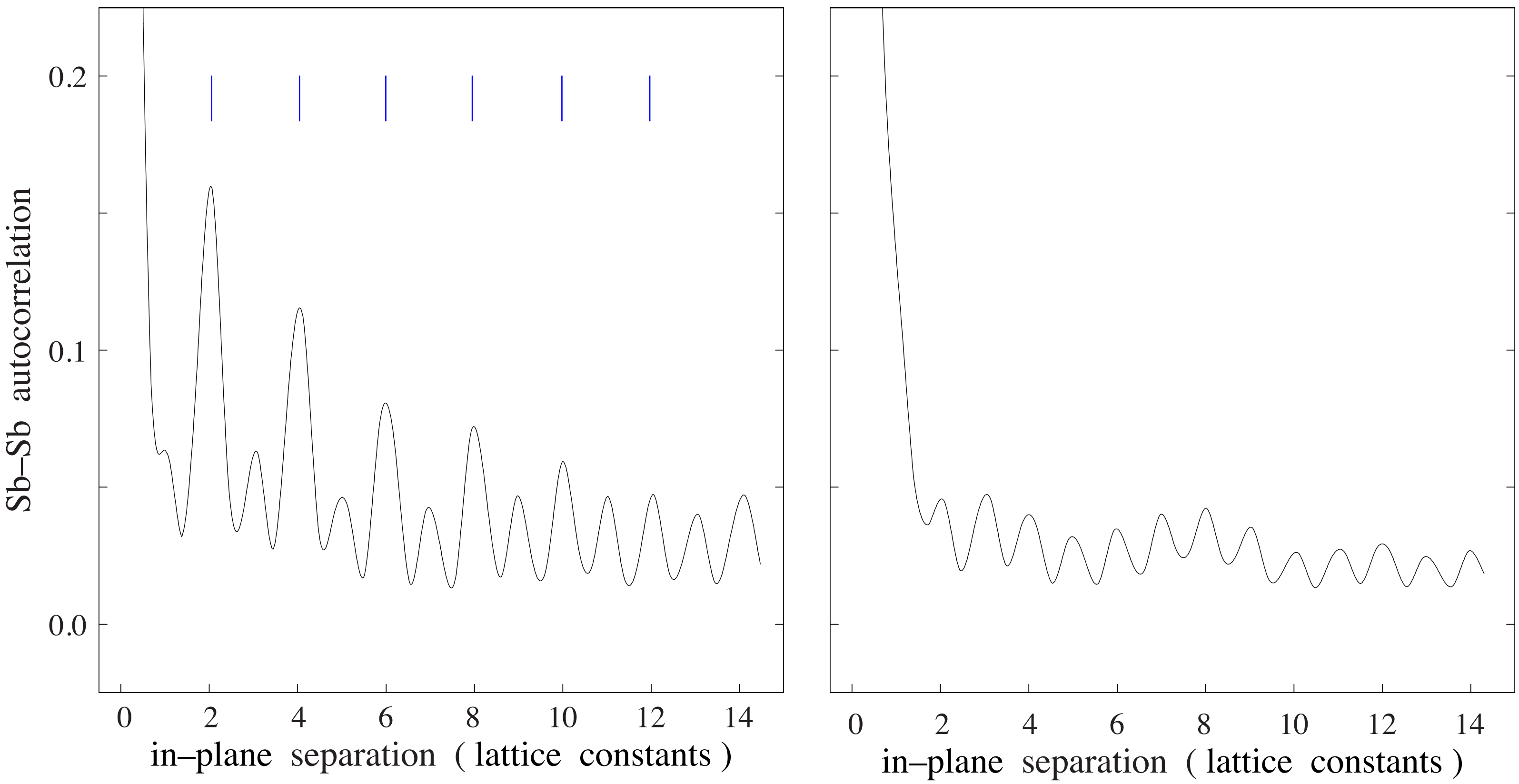


FIGURE 4.5. Sections through the <110> main diagonal of thresholded autocorrelation in Fig. 4.4. There is a damped but unmistakable correlation (marked by blue tics) at every–other–site in (–1–10) cross section. While in (1–10) cross section, right, the correlations appear constant (or random) for separations greater than one lattice constant.

spectrum (Fig. 4.6) [55]. Next–nearest–neighbor lattice spacing in (–1–10) cross section (circled in blue in Fig. 4.2 and marked by blue tics in Fig. 4.5) reveals itself as a streak of excess power located at half of the <110> reciprocal lattice vector (encircled in blue, left); this streak is absent in (1–10) cross section (right). Together these results imply that over the course of a lateral survey (~ 1 micron in length) next–nearest–neighbor pairing is not statistically significant on the (1–10), but is on the (–1–10) surface.

It is tempting to use the autocorrelation sections in Fig. 4.5 as a stand–in for a formal correlation function, and while they are useful for qualitative comparisons there are several drawbacks to this approach. First is that the image correlations cannot be directly translated into a probability of finding antimony atoms separated by given lattice vectors. Second, there is an electronic contribution to the contrast, so a single threshold emphasizes atoms in the InAsSb layer, at the expense of atoms in the InAs layer; antimony atoms in the superlattice have been identified via a trained eye to circumvent this, and we will show a representative example of this in the next section. Finally, because the image autocorrelation is *x* and *y* pixel based, all of the distortion detailed in Chapter 2 will at the very least cause correlations to leak into adjacent separations. This last drawback is addressed next.

Automatic Coordinate Assignment

The image in Fig. 4.2, left, is enlarged by a factor of ~ 3 in Fig. 4.7 to better visualize previously identified top–layer antimony atoms [20,23] that have been called out with blue dots. The pair distribution is a function of the separation vector between

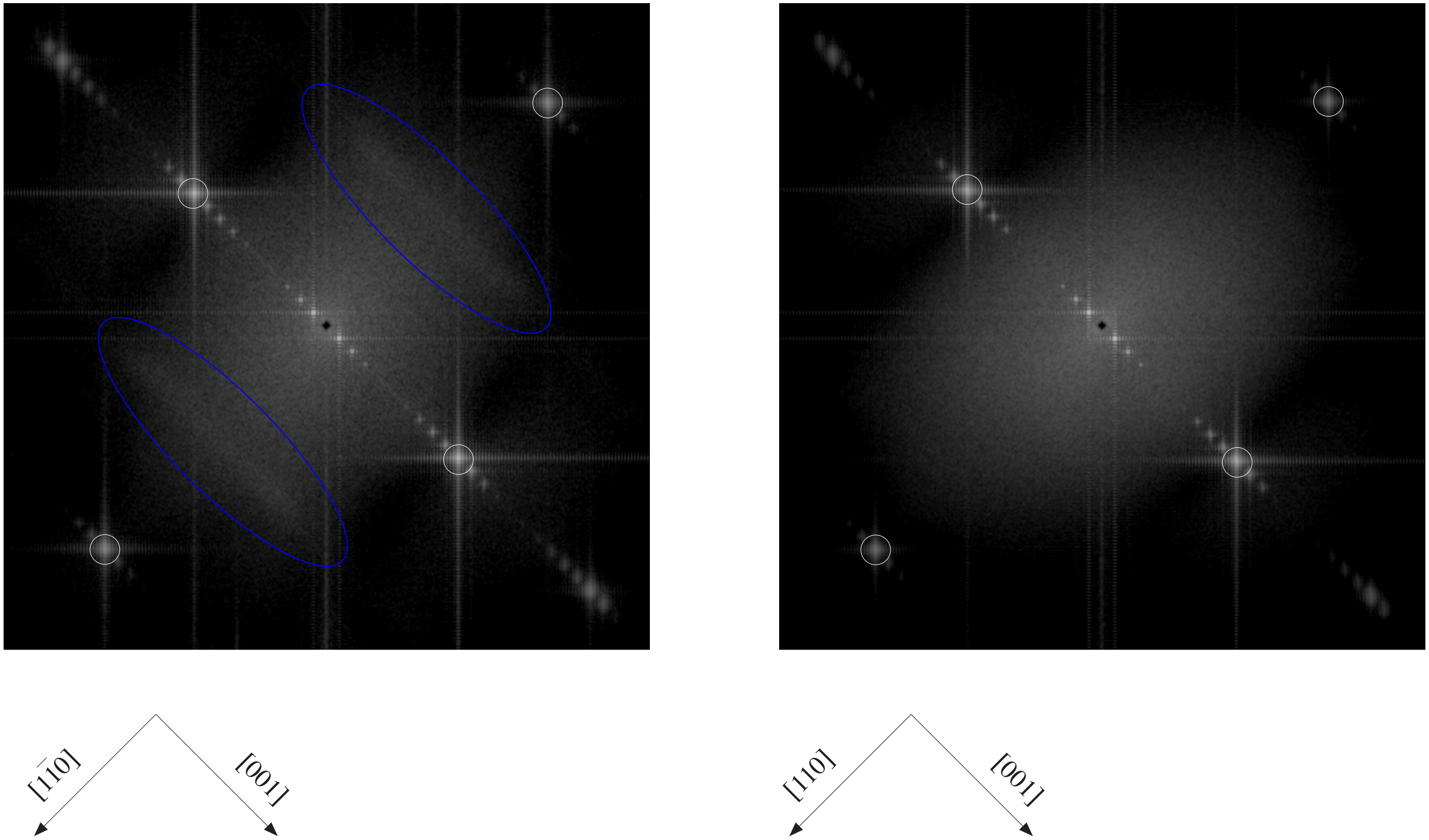


FIGURE 4.6. Ensemble reciprocal–space maps generated from (–1–10) and (1–10) cropped images (left and right respectively), with reciprocal lattice vectors (encircled in white) in the growth and in–plane directions indicating the surface atomic mesh. Excess power (encircled in blue) located at one–half the in–plane reciprocal lattice vector in (–1–10) cross section denotes the every–other–atom separation observed in Fig. 4.2 (left); this power is missing from the (1–10) reciprocal–space map as expected from Fig. 4.2 (right).

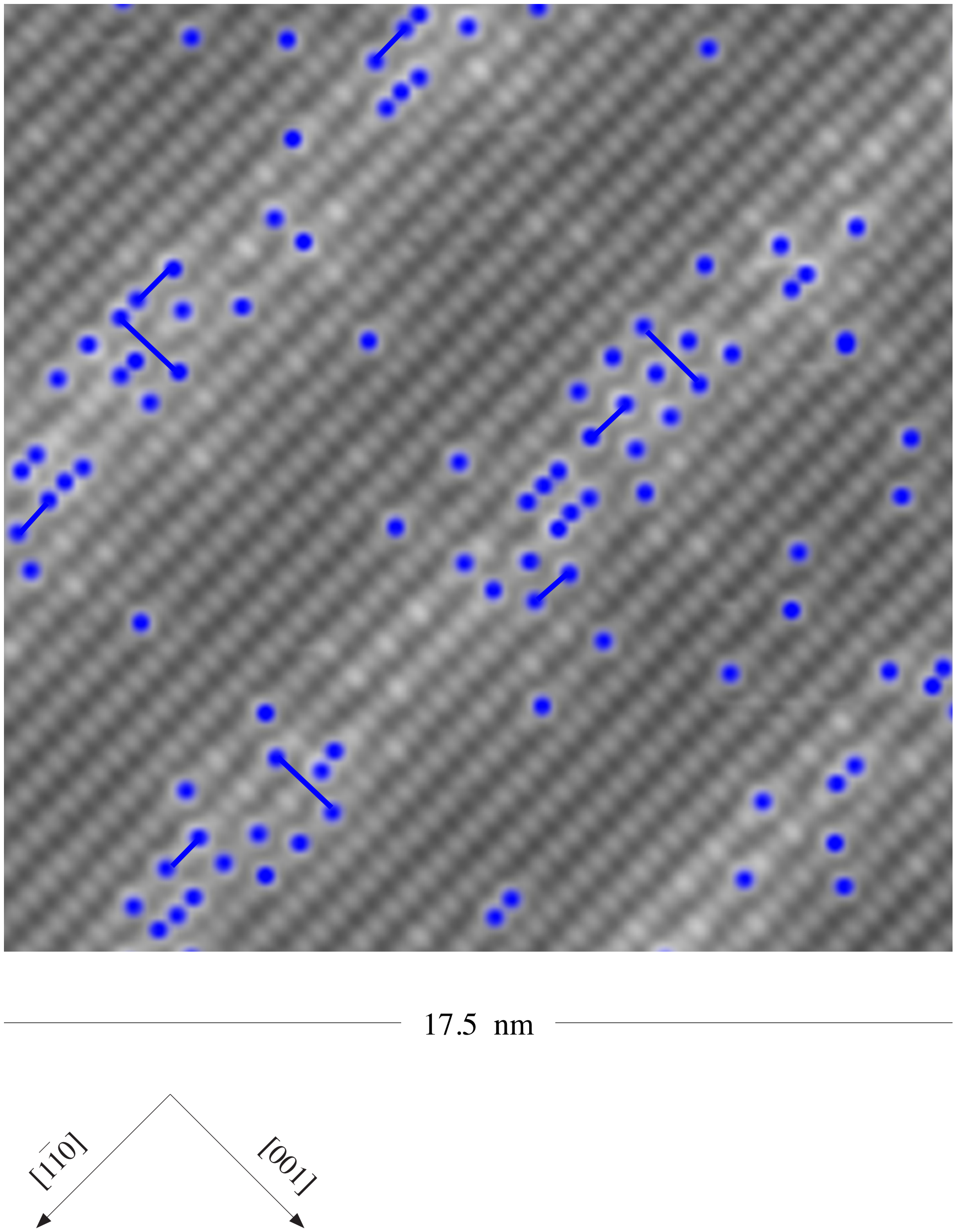


FIGURE 4.7. Enlarged image from Fig. 4.2 (left) with individual top–layer antimony atoms identified by blue dots. To quantify the two–body interactions between Sb atoms the separation vectors (indicated here by the "bonds" between antimony atoms) must be computed for any given pair. Examples of every–other–site separation vectors in the [001] and [1–10] directions are overlaid on the image.

pairs of antimony atoms calculated as a difference of locations in the [001] column and <110> row coordinate system[1]; representative 2 lattice site separation vectors are marked with blue bonds in Fig. 4.7. The problem, then, reduces to transforming the *x* (fast–scan) and *y* (slow–scan) indexing of antimony atoms (i.e. blue dots) to an index based on [001] columns and <110> rows.

As the bowed columns in Fig. 2.14 illustrate, any attempt at indexing in real–space is ill–fated. We instead endeavor to create image–by–image real–space masks, which can be used to separately identify rows and columns (two masks per image), derived from the reciprocal–space map of the full image. This seems like a step backwards after going to great lengths to crop the images and obtain cleaned–up power spectra such as those seen in Fig. 4.6, however the very properties of the full images which make the power spectra unusable (bowing resulting in smeared RLV spots) for the purposes of Chapter III are the very properties that we are trying to emulate with our real–space masks. We can use the distorted RLVs such as those seen in Fig. 4.8 together with phase information[2], which relates power to a location in the image, to emulate rows, which are bowed in precisely the same manner as the atomic rows in the STM images.

---

[1] Each <110> row is actually a (001) plane of atoms, and likewise each [001] column is actually a <110> plane of atoms. We choose to label the rows and columns by the indicated surface directions within each plane instead of the conventional miller indices since the directions are easier to keep track of when looking at 2D STM images.

[2] The reciprocal–space phase information is much more difficult to visualize than the reciprocal–space power and for that reason these phases are routinely “thrown away” in the process of computing the DFT.

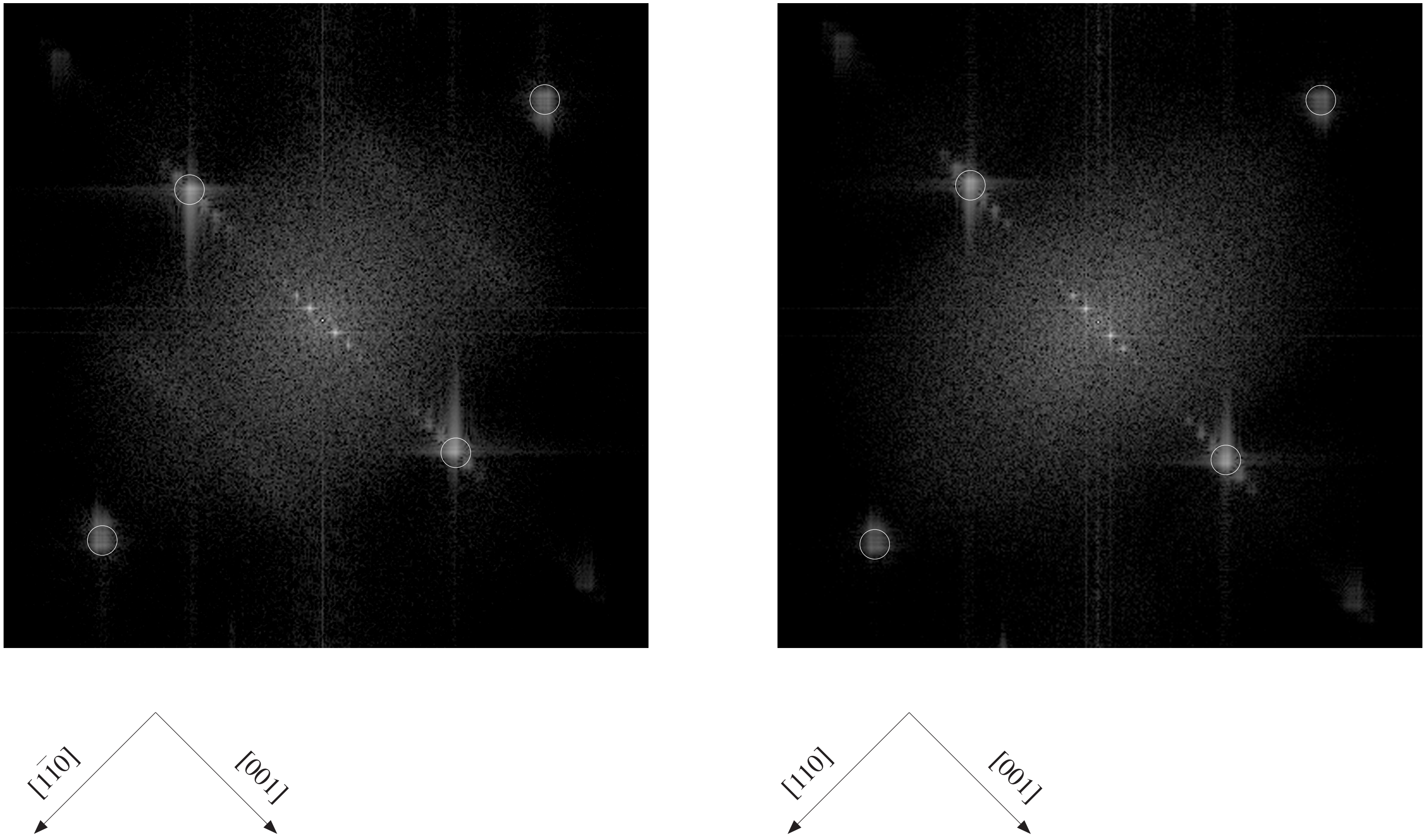


FIGURE 4.8. Individual reciprocal–space maps generated from full–size (–1–10) and (1–10) images (left and right respectively) embody the image distortion discussed in Chapter II. The complex–valued Fourier transform (which was squared to create the maps shown) will be used to create full–size lookup tables for [001] column and <110> row coordinates.

We then filter the DFT (with phase intact) so that only the two <001> spots[3] remain (Fig. 4.9, left). A 2D notch filter is constructed of a Gaussian function centered over each spot, with a width fine–tuned to be as restrictive as possible while still capturing the bowing that we wish to include in our indexing. Multiplying both real and imaginary parts of the DFT by the notch filter retains a small region around the ± [001] RLVs. The inverse transform of a pair (+/-) of delta functions is a sine wave, and likewise the inverse transform of this filtered (complex) DFT (Fig. 4.9, right) is approximately a sine wave modulation in the [001] direction and uniform in the perpendicular <110> direction except for the bowing associated with <110> rows. The brighter rows correspond to the atomic crests while the darker regions between the rows correspond to the atomic troughs. Likewise, filtering the DFT in the <110> direction (Fig. 4.10, left) yields a sine wave in the <110> that also contains the bowing from the [001] columns (Fig. 4.10, right).

Thresholding the inverse transforms (Fig. 4.9 and Fig. 4.10, right) at the zero crossing creates black and white masks of the rows (Fig. 4.11, left) and columns (Fig. 4.12, left), which are then sequentially labeled starting at the lower–left corner of the image for columns or the upper–left corner of the image for rows (i.e. left edge of the two corresponding image diagonals). The two different origins are necessitated by the 45° rotation of the [001] and <110> crystal axes relative to the *x* and *y* image axes[4]. Were

---

[3] As a reminder the DFT is reflection symmetric through the origin, so a single frequency is represented by two values in reciprocal space.

[4] The sought after $g_2(\mathbf{r}_2 - \mathbf{r}_1)$ is a function of separation vector, and therefore inversion symmetric, however our choice of origin flips the vector aligned with the rows (relative to our standard crystal axes) while leaving the vector aligned with the columns

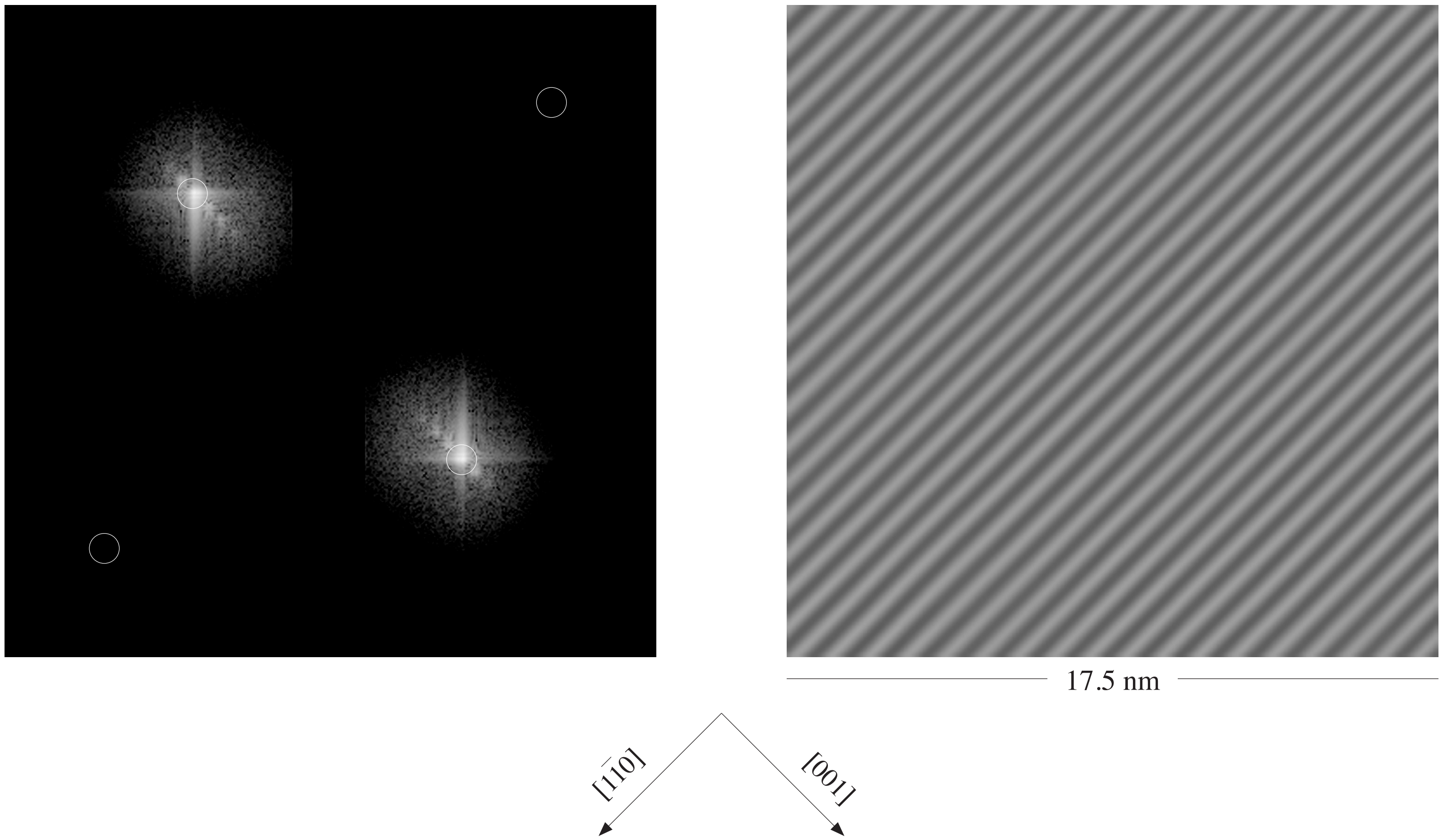


FIGURE 4.9. Complex–valued Fourier transform, which created Fig. 4.8 (left), is filtered to eliminate any power not localized around the <001> RLV peaks, shown on the left as a power spectrum. The inverse of the filtered Fourier transform yields continuous [1–10] rows along which arsenic and antimony atoms were situated in Fig. 4.7.

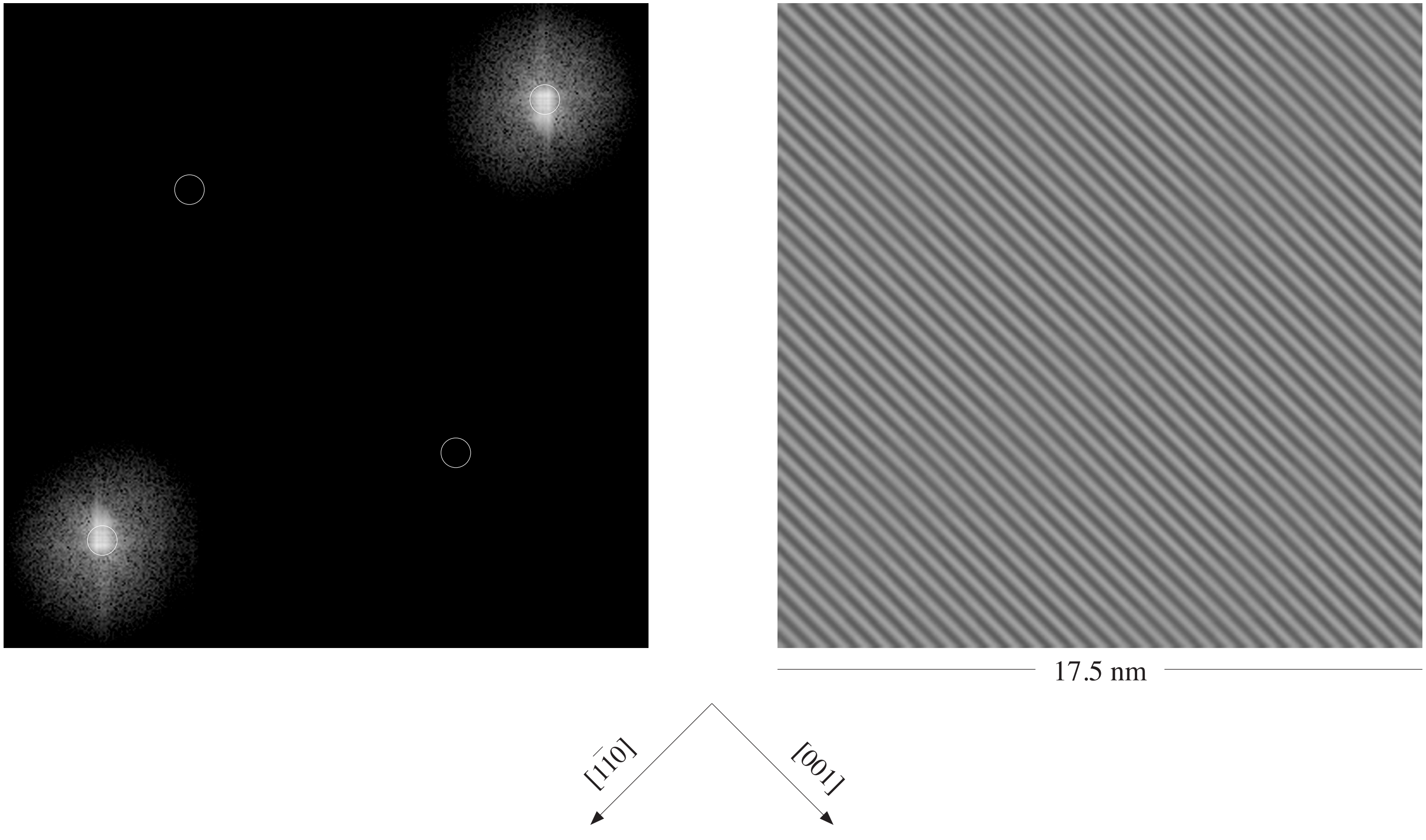


FIGURE 4.10. Complex–valued Fourier transform, which created Fig. 4.8 (left), is filtered to eliminate any power not localized around the <110> RLV peaks, shown on the left as a power spectrum. The inverse of the filtered Fourier transform yields continuous [001] columns along which arsenic and antimony atoms were situated in Fig. 4.7.

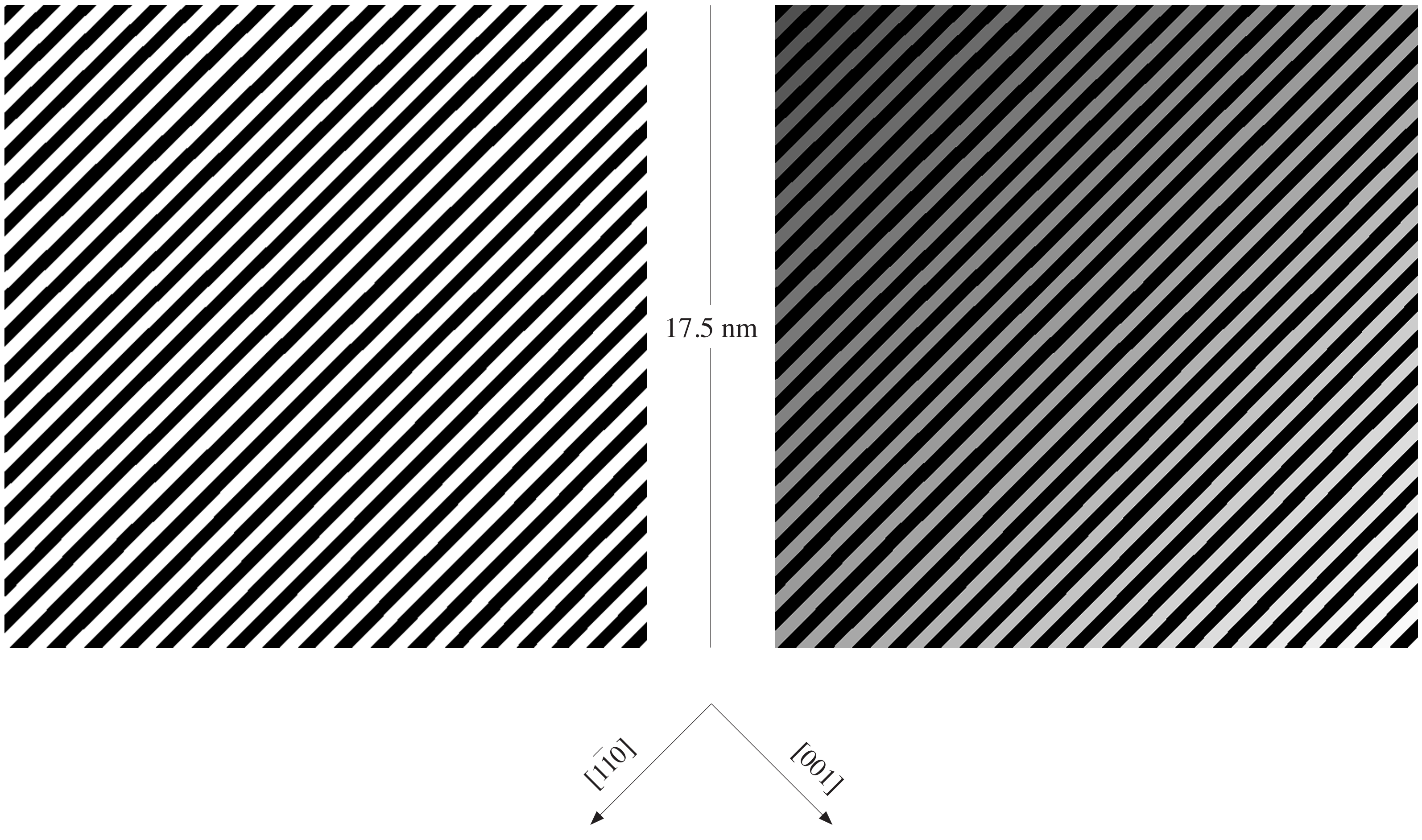


FIGURE 4.11. Inverse transform in Fig. 4.9 is thresholded to obtain a binary mask (left) of the [1–10] rows (white). This mask can then be turned into a lookup table (right) by sequentially labeling each row starting in the upper left corner. Grey values in lookup table have been adjusted for visibility.

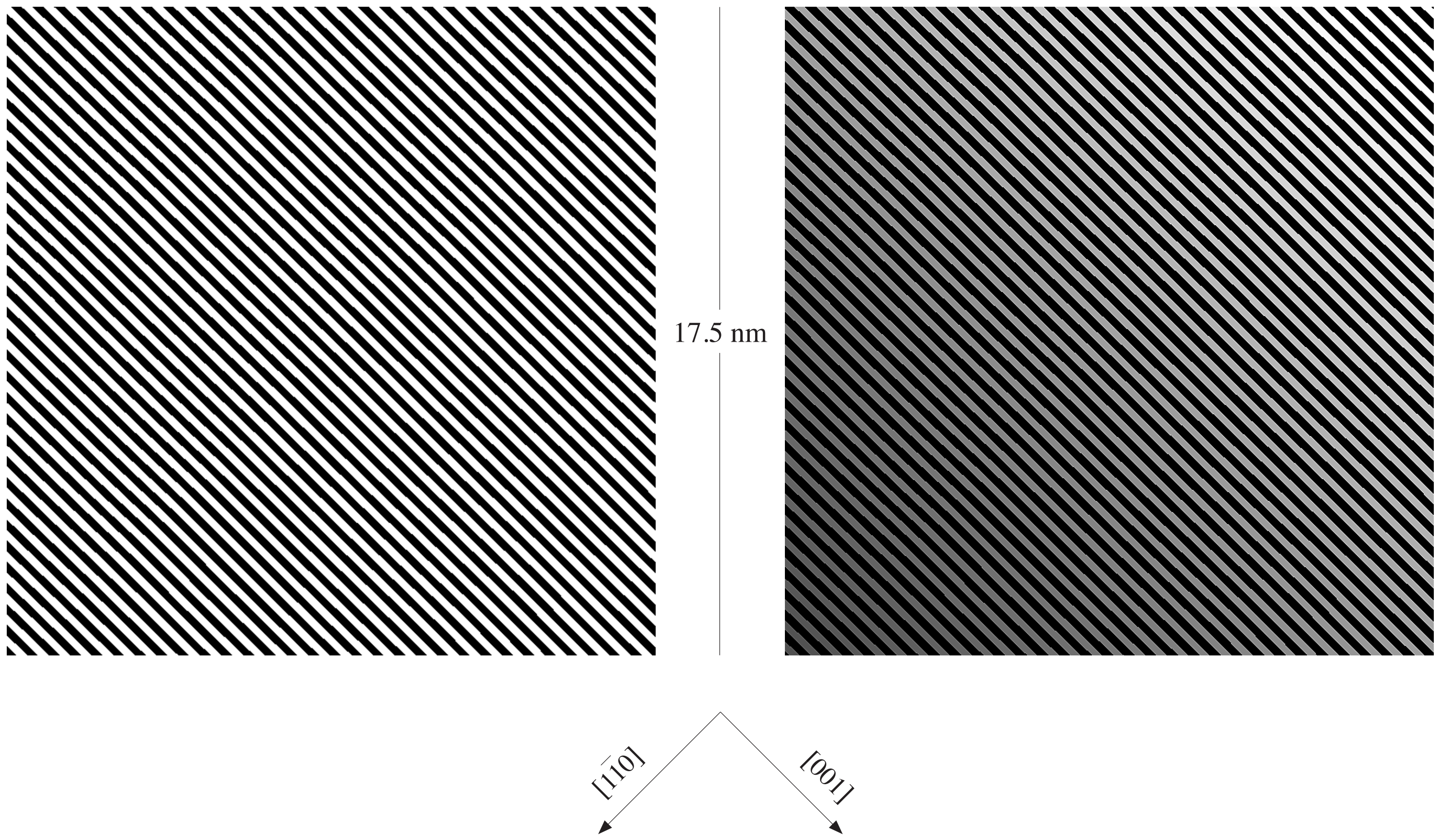


FIGURE 4.12. Inverse transform in Fig. 4.10 is thresholded to obtain a binary mask (left) of the [001] columns (white). This mask can then be turned into a lookup table (right) by sequentially labeling each column starting in the lower left corner. Grey values in lookup table have been adjusted for visibility.

the two sets of axes aligned, both rows and columns would originate in the standard lower–left corner. These sequentially labeled masks give us two lookup tables, where we can take x and y values and lookup the corresponding <110> row in Fig. 4.11 (right) or [001] column in Fig. 4.12 (right).

Pair Correlation Function: General Properties

Having established an algorithm for converting the coordinates of antimony atoms (identified in Fig. 4.7) into crystallographic <110> rows (Fig. 4.11, right) and [001] columns (Fig. 4.12, right), we turn now to computing correlations from the distribution of these atoms. The formal pair correlation function (whose natural logarithm is the pair potential in units of $-k_BT$ [69]), is given by

$$g_2(\mathbf{r}_2 - \mathbf{r}_1) = \frac{1}{x_{Sb}^2}\left[\frac{N_{Sb\text{-}Sb\ pairs}(\mathbf{r}_2 - \mathbf{r}_1)}{N_{anion\ pairs}(\mathbf{r}_2 - \mathbf{r}_1)}\right], \tag{4.1}$$

where the number of Sb–Sb pairs for each separation vector $|\mathbf{r}_2 - \mathbf{r}_1|$ is summed over all images in a given ensemble and normalized to the expected number for a random distribution of given density. As we'll see shortly, this normalization factor is given by the squared antimony fraction (averaged over a suitably large area) multiplied by the number of available anion pairs at said separation vector.

---

unchanged. This change is irrelevant to the superlattice (where separation vectors are limited to lie along the rows), however in the bulk alloy (where separation vectors are permitted to be linear combinations of rows and columns) care must be exercised.

To better connect the results of this calculation with the site occupancy patterns observed in our STM images, we consider two limiting cases: random and perfectly ordered. If antimony atoms are truly randomly distributed, then the site occupancy is isotropic, and pair separation vectors in all directions are equivalent. We may therefore look at just one of the available (arbitrarily chosen) directions in our example as representative of all others.

To create a row of randomly distributed antimony atoms we first generate a string of uniform random numbers between 0 and 1 (Fig. 4.13, left), then convert this continuous distribution to a bimodal distribution by assigning all values above a threshold to 1 and all values below the threshold to 0 where 1 is associated with antimony and 0 with arsenic as depicted in Fig. 4.13 (right) for two different impurity fractions. We can adjust the "density" of antimony atoms that this random string represents by setting the threshold to one minus the desired antimony fraction.

Enumerating the possible anion pairs (Fig. 4.14, left) separated by one lattice site over 1000 different random distributions each containing nominally 10% antimony spread over 100 lattice sites yields, $1000\,(\,100-1) = 99{,}000$, available pairs with $|\mathbf{r}_2 - \mathbf{r}_1| = 1$[5]. Since $|\mathbf{r}_2 - \mathbf{r}_1| = |\mathbf{r}_1 - \mathbf{r}_2|$, this function is mirror symmetric about the vertical axis. Tabulating the number Sb–Sb pairs in the same ensemble is done by marching through each antimony atom and calculating the separation between every

---

[5] We counted pairs of lattice sites to arrive at this number, however $N-1$ is the general result for $|\mathbf{r}_2 - \mathbf{r}_1| = 1$ in any row of length N lattice sites; whereas for $|\mathbf{r}_2 - \mathbf{r}_1| = n$ the general (combinatoric) result is $N-n$, i.e. the number of available anion pairs is a linearly decreasing function of their separation in lattice constants.

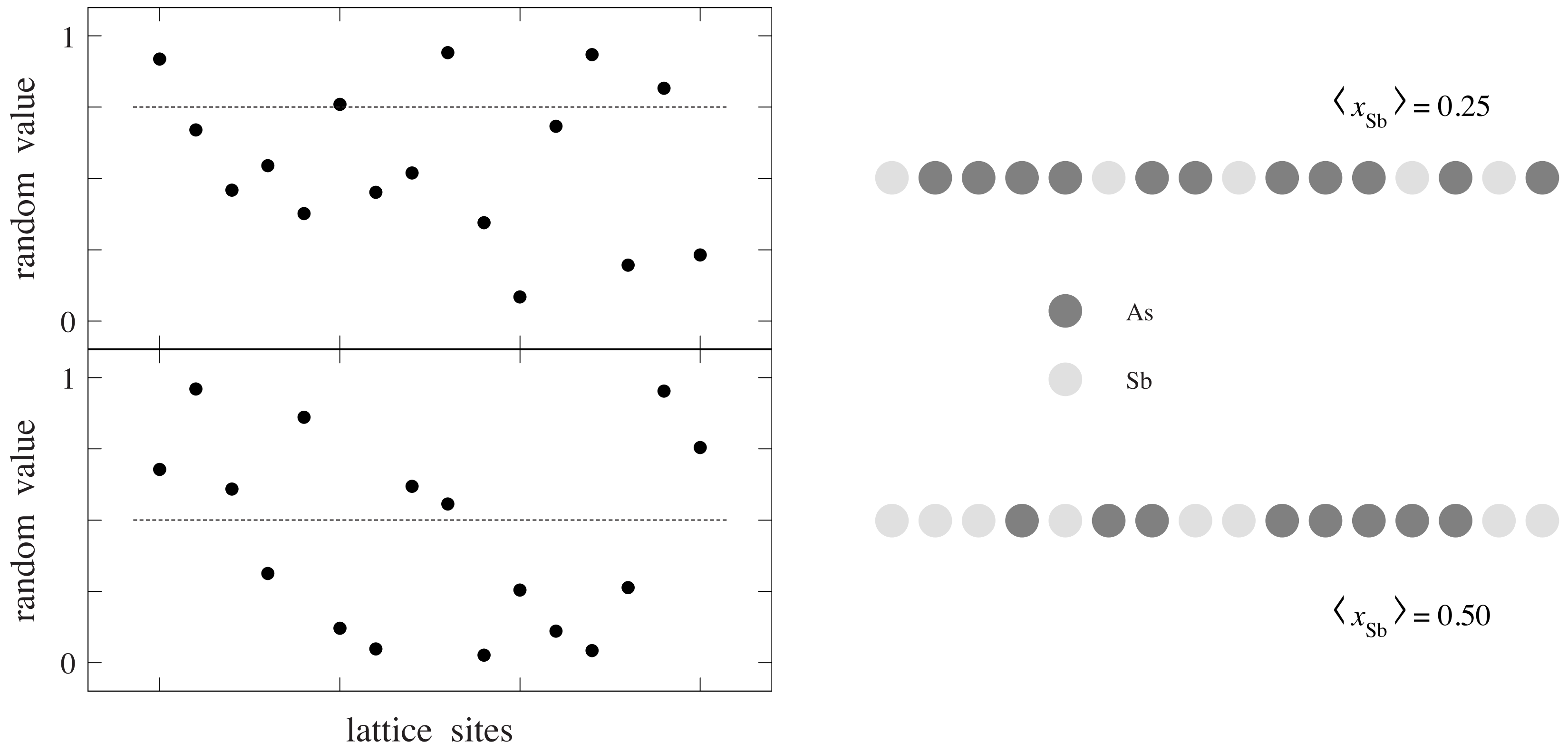


FIGURE 4.13. Random numbers ranging between 0 and 1 are thresholded to obtain a random distribution of antimony atoms (right) for two sets of random values (left). The threshold value is set to one minus the intended antimony fraction.

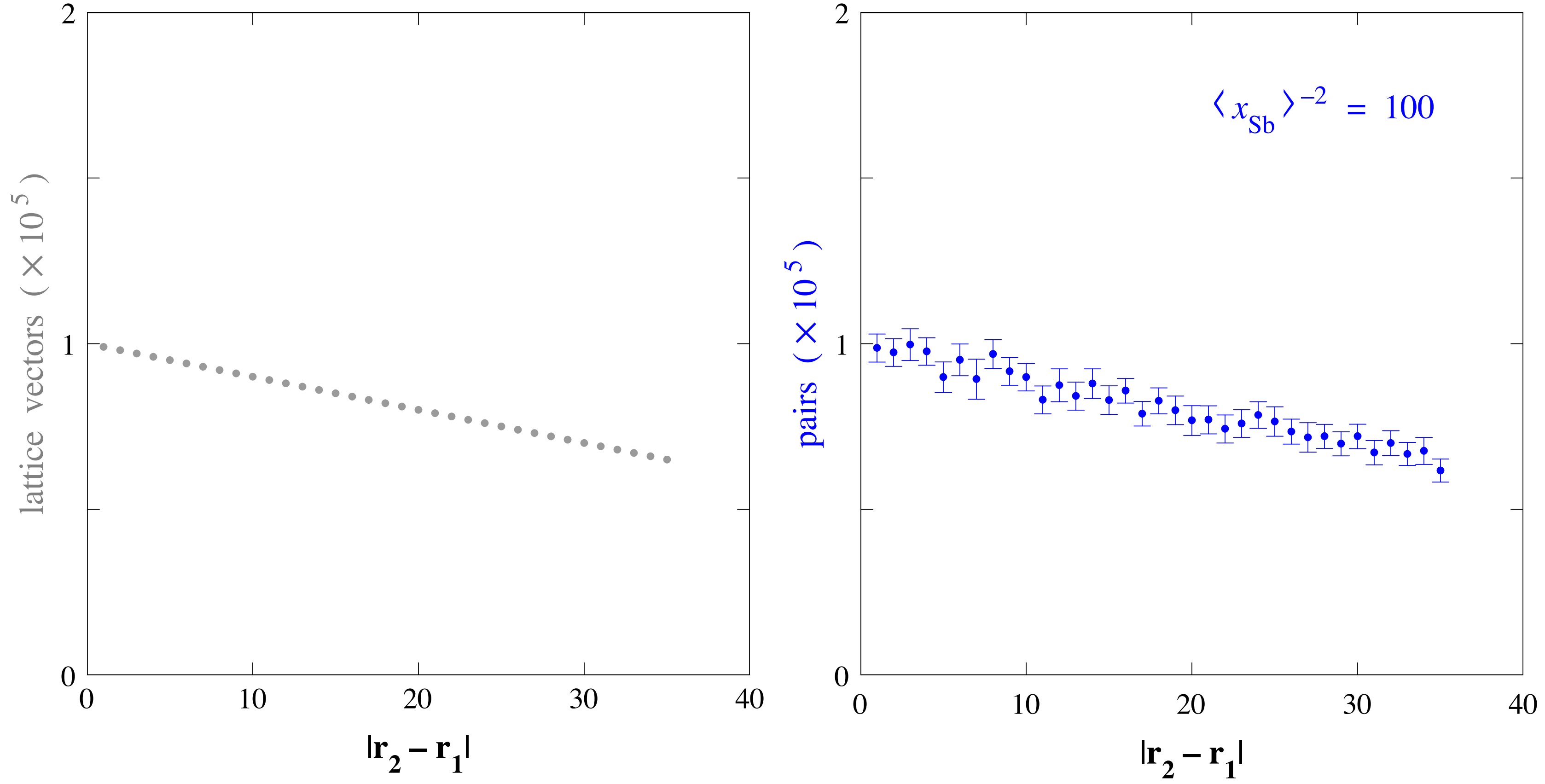


FIGURE 4.14. The number of randomly distributed antimony pairs rescaled by the inverse impurity fraction squared agrees (right) with the available lattice vectors (left). The linear decline with separation is expected for data of finite extent.

other antimony atom in the 100 total lattice sites. Adding these up across the 1000 unique distributions yields 987 nearest–neighbor pairs (Fig. 4.14, right), explicitly demonstrating that the antimony fraction squared times the total number of lattice pairs at a given separation vector is the infinite length stand–in for the number of Sb–Sb pairs expected for a random distribution. The virtue of this normalization is that the correlations are directly related to probabilities with random (or uncorrelated) equal to one; then anything below one is less likely than random, and anything above one is more likely than random. We will return to this normalization when we consider the appropriate calculation for a superlattice whose antimony fraction is modulated in the growth direction, however we turn first to a consideration of perfect alloy order as the limiting counterpoint to the random distribution just examined.

Perfect CuPt–B ordering consists of alternating (1–11) or (–111) planes populated with one or the other type of anion (either all antimony or all arsenic), creating a monolayer superlattice in the [1–11] or [–111] direction. This situation is clearly obtained only when $x_{Sb} = 0.5$. Since STM is surface sensitive, these (–111) planes cannot be probed directly, but the unique patterns that these planes display as they intersect either the (–1–10) or (1–10) cleavage surfaces are schematically illustrated in Fig. 4.15. We consider here the (–111) variant of CuPt–B ordering, which results in chains of antimony atoms aligned in the [1–12] direction and alternating arsenic and antimony atoms on the (–1–10) cleavage surface in both [001] and [1–10] directions. As a direct consequence of the surface strobing every second bulk monolayer in {110} cleavage, the (1–11) and (–111) variants are indistinguishable from the surface. For

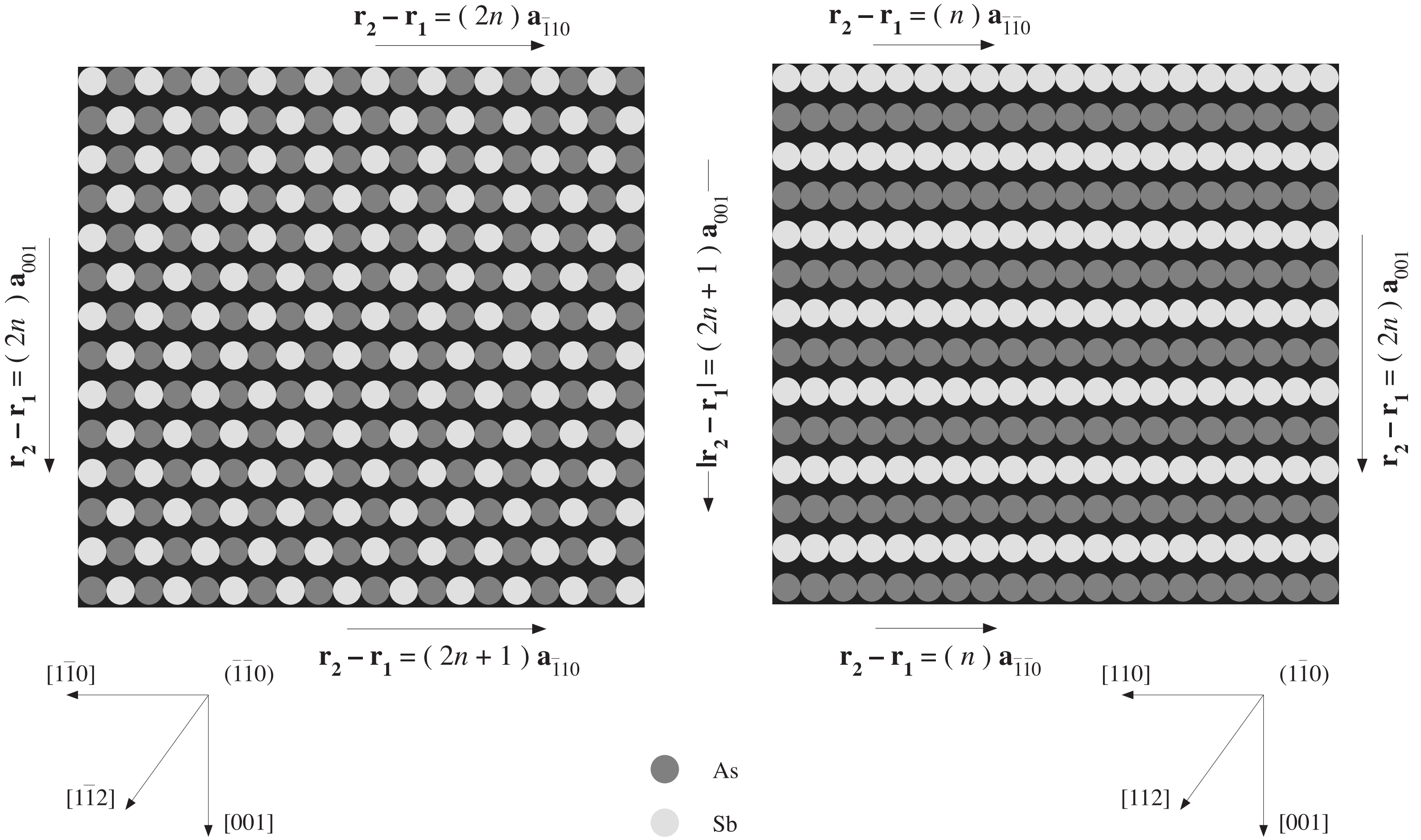


FIGURE 4.15. Surface projections of perfect CuPt–B ordering in (–1–10) and (1–10) cross section (left and right respectively). Alternating planes of (–111) atoms result in [001] alternation of antimony and arsenic atoms in both projections; the two projections differ along the <110> rows however, in (–1–10) projection the atoms also alternate in the [1–10] direction whereas in the (1–10) projection chains of like atoms align. Crystal axes are rotated 45° clockwise relative to STM images. Coordinate origin is lower left for <110> lattice constants, upper left for <001> lattice constants.

example, if the planes of arsenic or antimony are (–111), then in the bulk these anions alternate in the [–111] direction, but on the (–1–10) surface we miss the second anion type, so both [1–12] and [–112] directions appear to lie in ordered planes even though only one of them truly does. Although not considered here, the A–type of CuPt ordering, which consists of (111) or (11–1) planes of anions reveals itself in the orthogonal cross section, simply interchanging left and right schematics in Fig. 4.15. CuPt–A is observed much less frequently in nature, although preliminary evidence has shown that CuPt–A order might occur in InAsSb materials [26].

There are only two unique sections through the 2D surface projected correlation function for a bulk structure displaying perfect order. The observed 1D correlation function depends on which surface crystallographic direction the separation vector is aligned with. Once this surface direction is chosen, the correlations are a function only of the length of the separation vector. All of the plots shown in the remainder of this chapter assume that this direction has been specified, and only show the correlations for a single direction, furthermore since the correlations are symmetric, only positive vectors are plotted. Considering first the [1–10] direction in (–1–10) cross section, Fig. 4.15 (left) any separation vector whose length is an odd number, $|\mathbf{r}_2 - \mathbf{r}_1| = (\,2n + 1\,)$, of lattice sites will not join Sb–Sb pairs; the corresponding pair correlation function is then equal to 0, as illustrated in Fig. 4.16 (left). On the other hand, half the even length vectors, $|\mathbf{r}_2 - \mathbf{r}_1| = (\,2n\,)$, connect Sb–Sb pairs while the other half connect As–As pairs. Normalizing to the 50 / 50 random alloy, which has equal number of As–Sb, Sb–As, As–As, and Sb–Sb pairings (and therefore an Sb–Sb pair probability of one quarter),

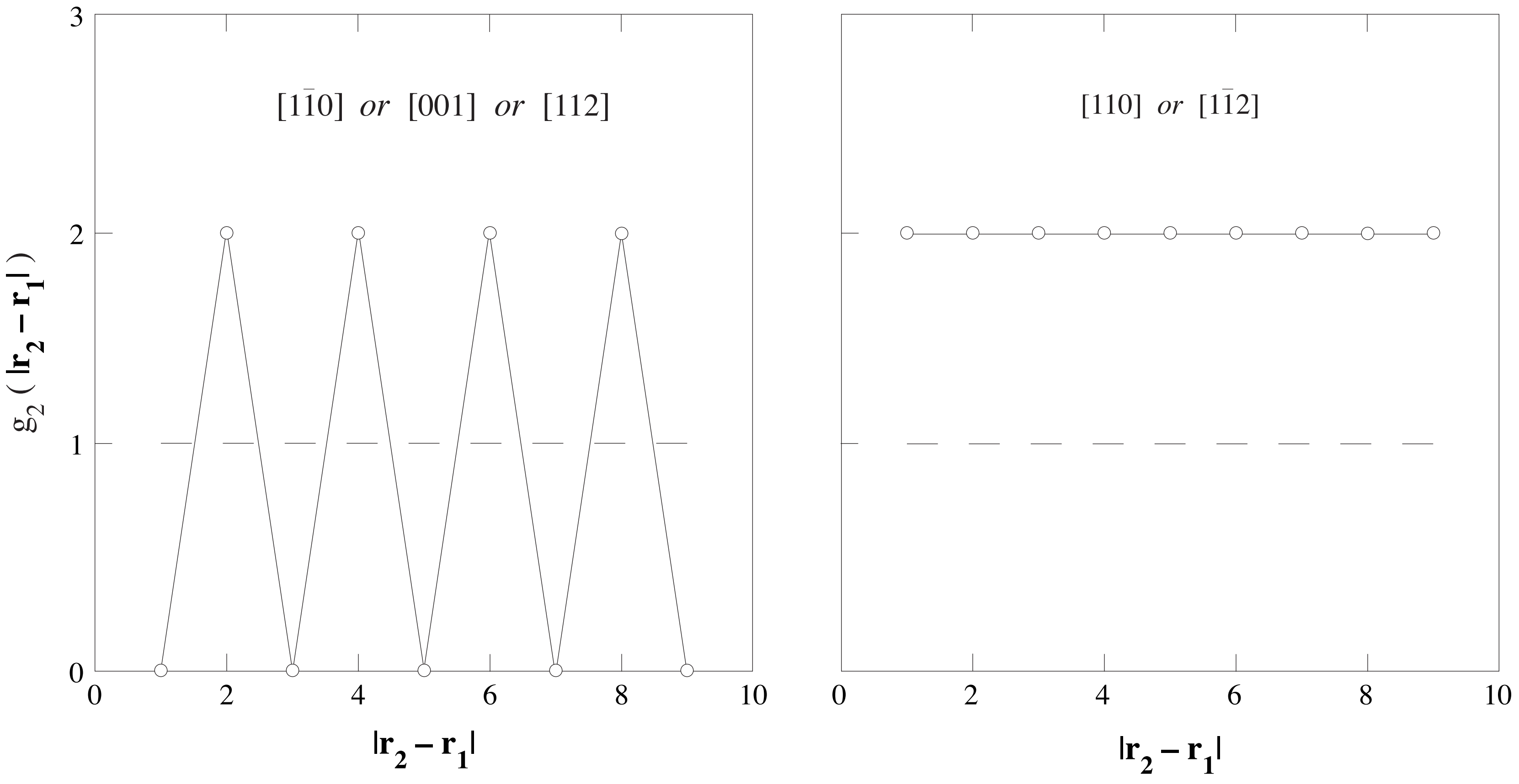


FIGURE 4.16. Pair correlation functions calculated from orthogonal surface projections of perfect (–111) CuPt–B order illustrated schematically in Fig. 4.15. Correlations in the growth plane manifest very differently on the two cleavage surfaces. Dashed line indicates a random distribution.

yields a correlation of 2. So the correlation function in the [1–10] direction alternates between 0 and 2 for perfect anion ordering, and this alternation persists for the largest separation vectors allowed by the image size[6]. Due to symmetries in the surface projections of the ordered planes considered in Fig. 4.15, this same correlation function is obtained in the [001] (observed in both left and right panels) and [112] (observed in the right panel) directions.

Considering separation vectors that lie in the [110] direction, in (1–10) cross section (Fig. 4.16, right) Sb–Sb pairs occur for either any combination of lattice sites, or no combination of lattice sites, therefore averaged across the entire image (with mean impurity fraction squared of 0.25) the correlation function remains fixed at 2 (Fig. 4.16, right). Again due to symmetries in Fig. 4.15 we expect this same constant result in the [1–12] (observed in the left panel) direction in addition to the [110] (observed in the right panel) direction.

Even though the potential associated with a given correlation function is meaningful in gasses [69], it isn't clear whether the correlations in a crystalline lattice are governed by a corresponding potential. For example, correlations of 0, found in perfect CuPt–B order, imply an infinitely repulsive potential ($\ln(0) = -\infty$), which isn't physical. We will nevertheless use the potential framework in the last section of this chapter to compare results across sub–ensembles of the data.

---

[6] Although Fig. 4.16, left, is described in terms of probabilities, the same result is obtained by considering the ratio of the total number of antimony pairs to available pairs and normalizing to the antimony fraction squared.

Pair Correlation Function: Bulk Alloy Data

Before particularizing to the InAs / InAsSb superlattice, where modulation in the growth direction poses an addition complexity, we introduce the correlation function calculated from experimental data by considering a bulk InAsSb alloy which is nominally uniform in the [001] growth direction. We will return to the specifics of this bulk alloy growth later, when needed to place the corresponding results in proper context, but for now focus only on the structure imaged in (–1–10) cross section (Fig. 4.15, left). The structure was surveyed laterally following a single point in time during the growth as was done with the InAs / InAsSb superlattice (Fig. 4.1), and a portion of that survey is reproduced in Fig. 4.17. Local alloy order, similar to that in the superlattice, is observed along the [1–10] direction together with relatively large voids free of antimony atoms. Enlarging a single image (Fig. 4.18, left) shows the same every–other–atom antimony site occupancy preferred along [1–10] rows seen in the superlattice (Fig. 4.2, left), and the survey averaged DFT shares the same [1–10] half order streak seen in the superlattice as well (Fig. 4.6, left). Because of these similarities in both real– and reciprocal–space, we expect the correlations between substitutional antimony atoms in both the bulk alloy and the strain–balanced superlattice to be comparable, a point we will return to later.

Individual antimony atoms are identified in the bulk by adjusting a threshold so that the called out atoms originate almost entirely from the surface layer, this is in contrast to the superlattice where each top–layer antimony atom must be hand identified (see Fig. 4.7). These atoms are then assigned [001] (column) and [1–10] (row)

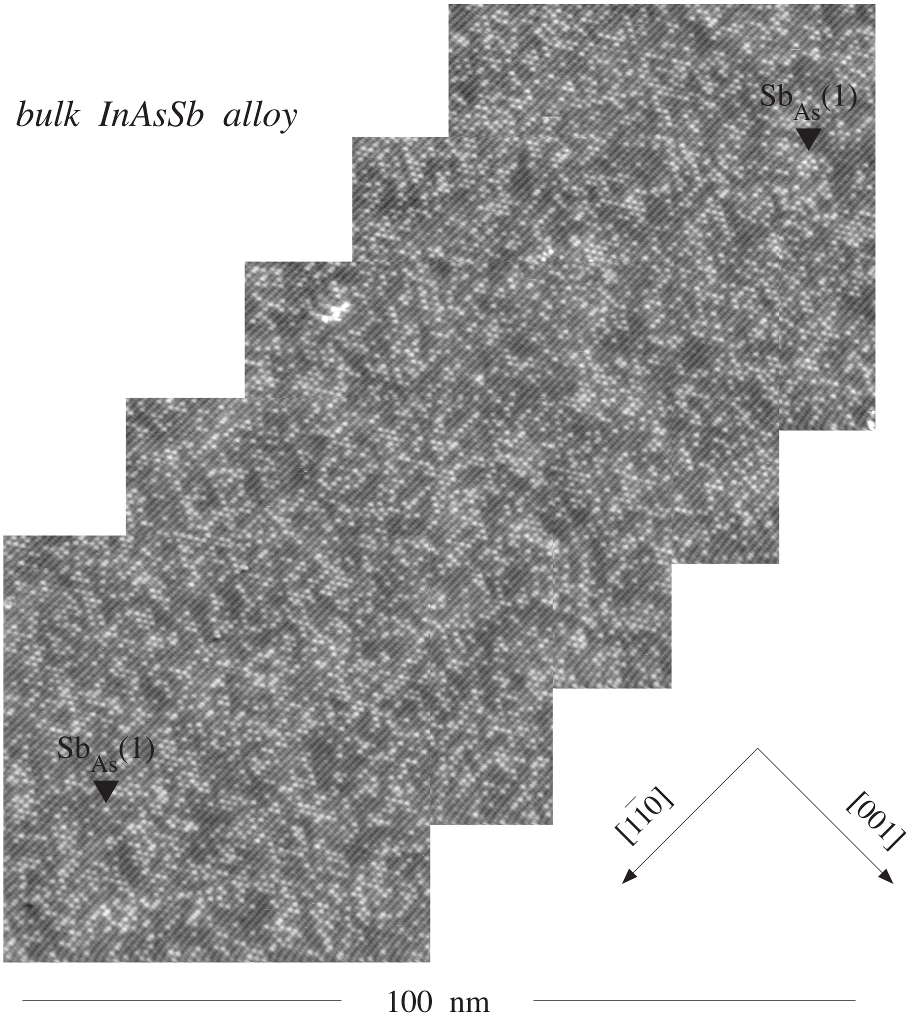


FIGURE 4.17. Lateral survey in (–1–10) cross section over bulk $InAsSb_{0.2}$, keeping a single point in time during the growth fixed. Top–layer antimony–for–arsenic replacement is indicated with carets. Growth direction is from top–left to bottom–right.

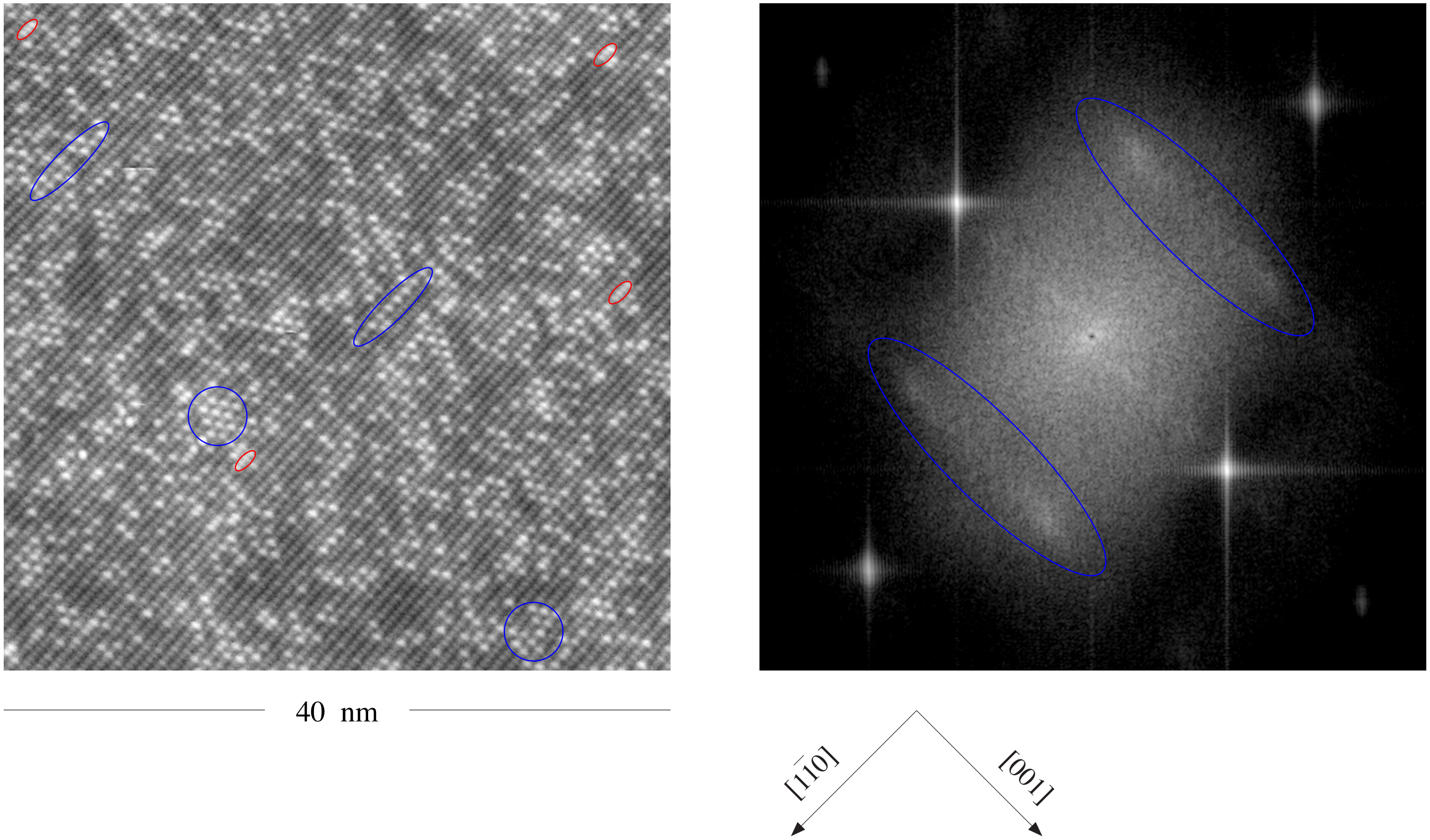


FIGURE 4.18. Individual image of the anion sublattice (left) taken from the (–1–10) survey shown in Fig. 4.17, accompanied by the survey average reciprocal space map (right) showing <110> half order streak (encircled in blue) associated with every–other atom spacing in the image (examples also encircled in blue). Less prevalent in the image are antimony atoms occupying nearest–neighbor lattice sites (examples encircled in red). Relatively large gaps where no antimony incorporated are also notable in the image. Growth direction is from top–left to bottom–right.

coordinates as previously described, a separation vector is calculated from these coordinates for all combinations of two antimony atoms, and the number of times each separation vector appears is tallied. The result is a discrete, 2D histogram, with bins located at integer numbers of lattice sites in the [001] and <110> directions, which when normalized to random via (4.1) is the 2D pair correlation function. Since each atom could be either $\mathbf{r}_1$ or $\mathbf{r}_2$ the correlation function will be mirror symmetric through the origin. Appropriate sections can then be taken in physically relevant directions.

Fig. 4.19 illustrates the results for [1–10] and [1–12], directions representative of the left and right panels in Fig. 4.16 respectively. The [1–10] correlation function[7] displays a damped version of every–other–lattice–site pattern in Fig. 4.19 (left) approaching unity – the value expected of a random distribution –after approximately 15 lattice sites; the amplitude of the correlations is also weaker than expected for a perfectly ordered 50 / 50 bulk alloy. The observed next–nearest–neighbor value of 1.5 corresponds to a 50% higher likelihood than random for antimony atoms separated by 2 lattice sites, similarly the nearest–neighbor value of $0.7 = 1/1.4$ corresponds to a 40% lower likelihood (compared to random) for these atoms to be situated next to one another. A bit surprisingly the high and low probability branch amplitudes are not symmetric as they are in Fig. 4.16 (left) and at this time we don't know why this might be the case.

The [1–12] correlation function (Fig. 4.19, right) starts out greater than one (nearest neighbor pairs are 20% more likely than random), but quickly returns to random

---

[7] The illustrated uncertainties in both directions are smaller than they should be due to for double counting where images overlap in Fig. 4.17. Corrected errors could increase by as much as a factor of $\sqrt{2}$ over those shown in Fig. 4.19.

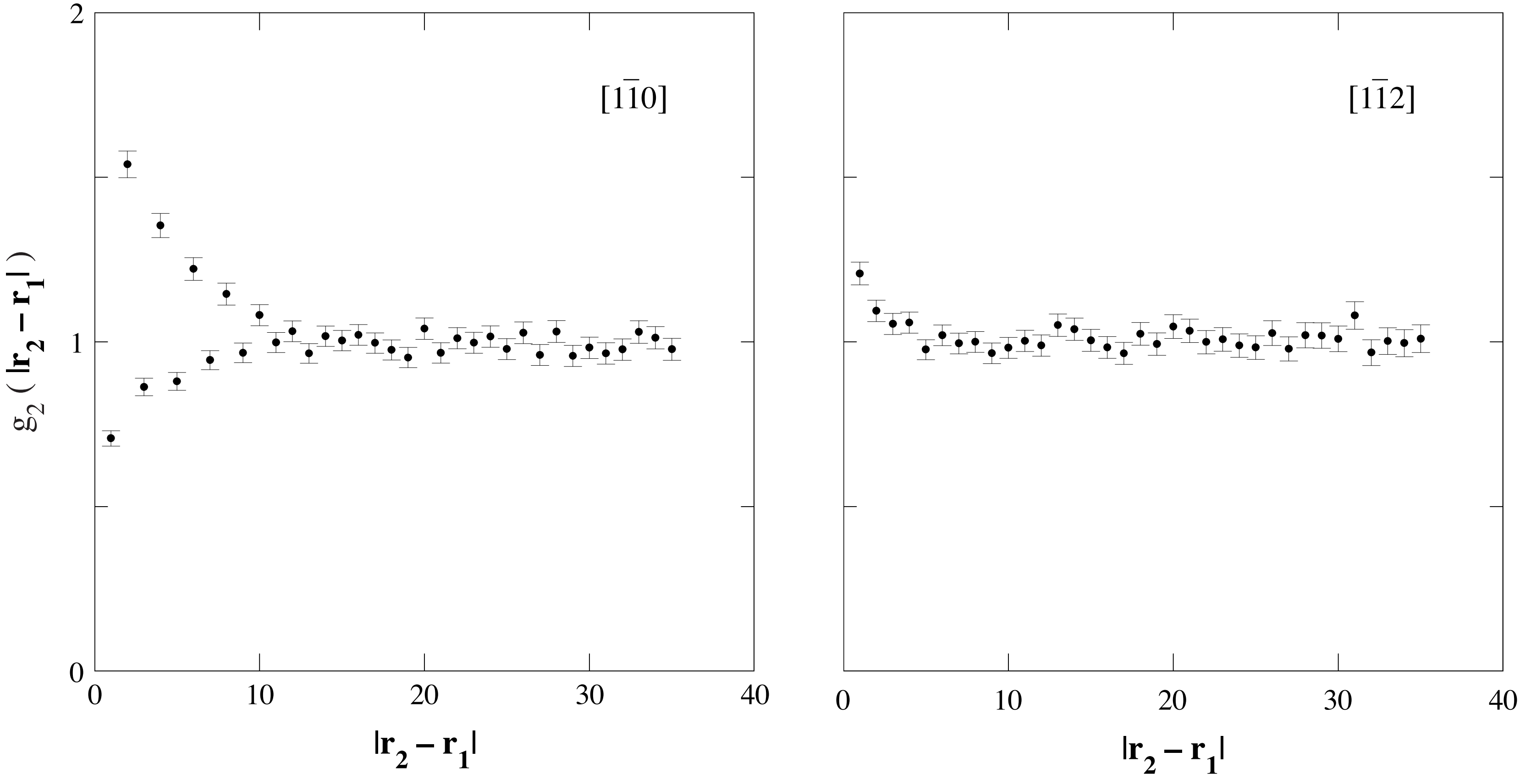


FIGURE 4.19. Correlation functions calculated on (–1–10) cross section of bulk $InAsSb_{0.2}$. Principle cleavage surface directions reveal short–ranged next–nearest–neighbor correlations along [1–10], and short–ranged nearest–neighbor correlations along [1–12].

after approximately 4 lattice sites, and again displays weaker correlations than those expected for perfect ordering (Fig. 4.16).

Pair Correlation Function: Superlattice Data

To apply this formalism to the superlattice we need to first make several changes to the algorithm discussed above on account of the [001]–modulated nature of the structure. We restrict the separation vectors to lie in the growth plane since the antimony fraction clearly varies in the growth direction, this leaves the pair correlations as functions only of the distance between any two antimony atoms (in <110> lattice constants). We also use a predefined counting window [20] such as the one illustrated in Fig. 4.20 (left) to ensure that each <110> row is equally sampled in the [001] growth direction, and care must also taken to eliminate overlap between superlattice images ensuring each window represents an independent measurement. A similar mask, applied to the bulk, is likewise shown in Fig. 4.20, and is used in what follows to test for any difference between bulk and superlattice computational algorithms.

The bulk correlation functions presented in the last section were normalized to a random distribution with corresponding antimony fraction via (4.1); as we'll see this normalization does not work when the structure in question is modulated in the growth direction. We then have two options for calculating a [1–10] or [110] correlation function. We can calculate the correlation function within a given <110> row of presumably constant antimony fraction using Eq. (4.1), and then average over the <110> rows of [001] modulated antimony fraction; alternatively we can pool the number of

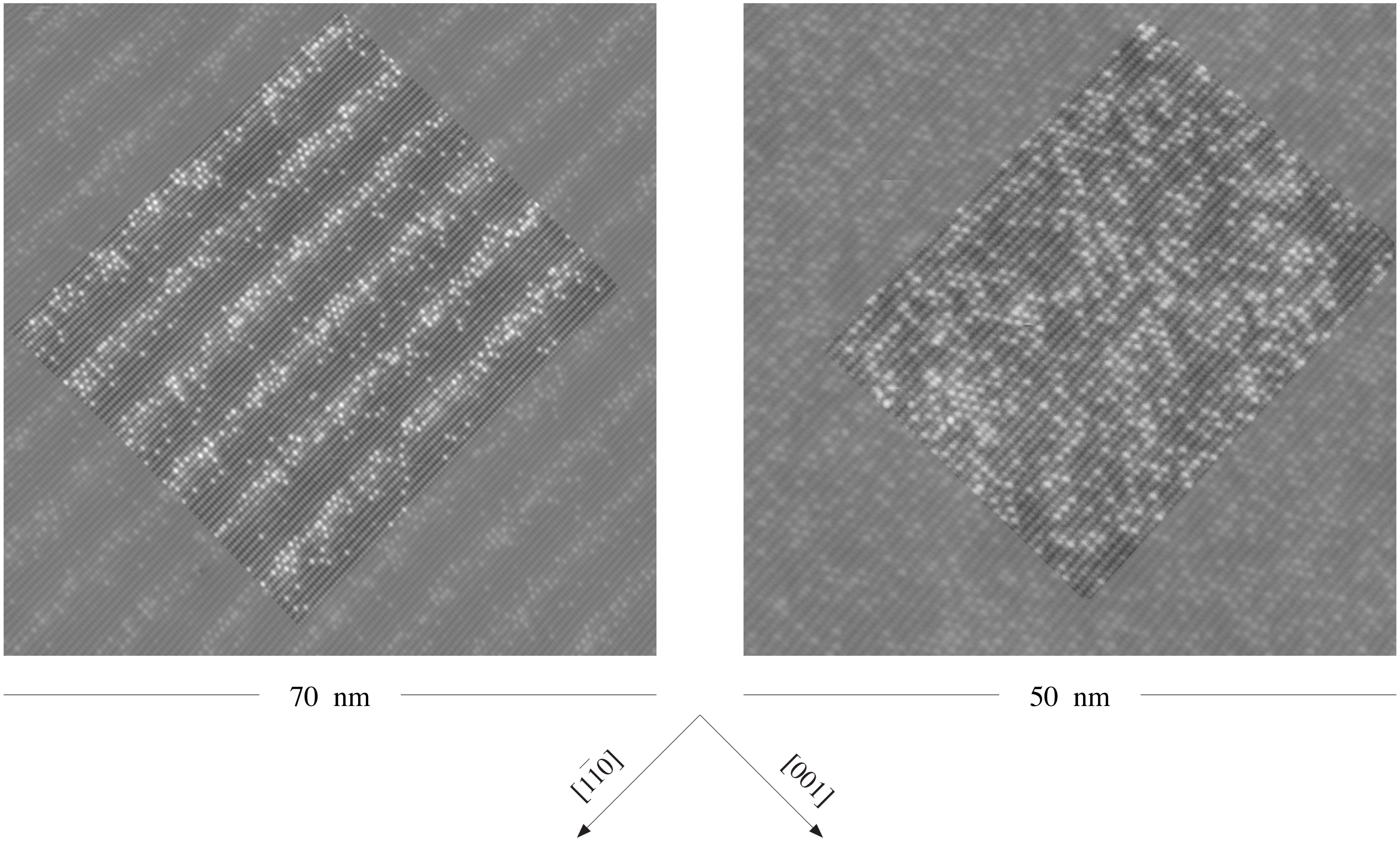


FIGURE 4.20. Single superlattice image shown in Fig. 4.2 reproduced in full together with single bulk alloy image shown in 4.17 reproduced in full. Calculation of pair correlation function in superlattice is restricted to the window previously used for counting to ensure equisampling in the growth direction and no overlap between adjacent images in the growth plane. A similar mask was applied to the bulk images to test the influence of an algorithm change on the resulting correlation function.

antimony pairs for any <110> separation vector and normalize to an appropriately redefined random distribution. Provided such a new normalization can be found a more accurate calculation is always achieved with finite data by averaging before normalizing instead of normalizing before averaging on account of error propagation [61].

We take inspiration from (4.1), which normalized the number of pairs at a given separation vector to the number expected for a random distribution at a given antimony fraction and now normalize the number of pairs summed over all (modulated) monolayers to the total expected pairs corresponding to random distributions with the same antimony fractions. Thus our pair correlation function is redefined as

$$\mathrm{g}_2(|\mathbf{r}_2-\mathbf{r}_1|)=\frac{1}{\sum_i x_i^2}\left[\frac{\sum_i N^i_{Sb\text{-}Sb\,pairs}\,(|\mathbf{r}_2-\mathbf{r}_1|)}{N_{anion\,pairs}(|\mathbf{r}_2-\mathbf{r}_1|)}\right], \tag{4.2a}$$

with *i* indexing the monolayer which has the (presumably constant) antimony fraction x_i.

Numerical experiments confirm the desired normalization in the presence of a modulated antimony fraction is indeed the mean squared antimony fraction multiplied by the number of anion pairs at the given separation vector. As a concrete example, consider the exponentially decaying fraction illustrated in Fig. 4.21 (left), the observed pairs normalized to the squared mean antimony fraction, x_{Sb}^2(Fig. 4.21, right, red), exceed the lattice vectors, however, the observed pairs normalized to the mean squared fraction, $\sum_i x_i^2$ (Fig. 4.21, right, blue), coincide with the lattice vectors. The observed ratio of 2 between x_{Sb}^2 and $\sum_i x_i^2$ in this case is dictated by the explicit functional form of

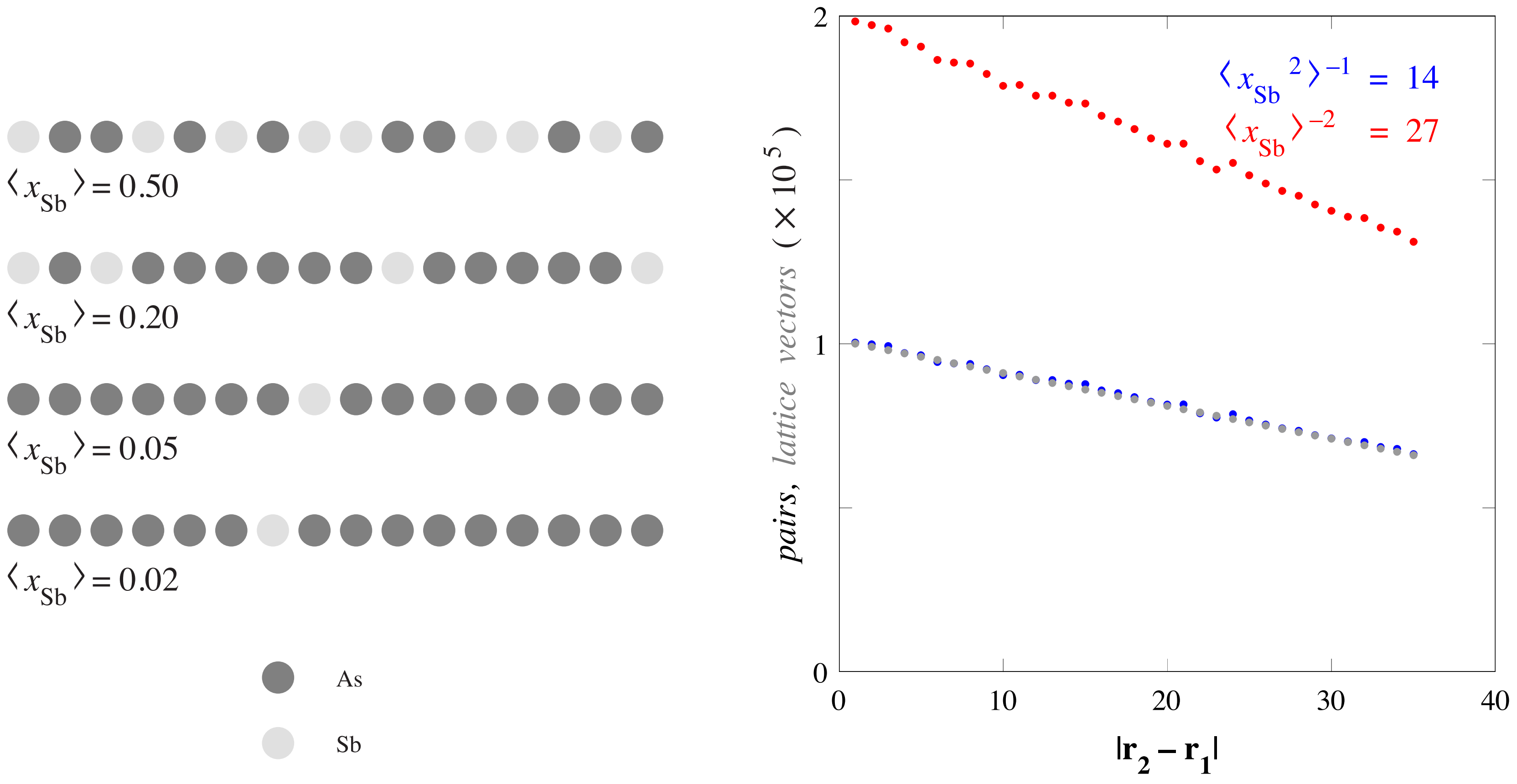


FIGURE 4.21. Schematic illustrating four impurity fractions approximating an exponentially decaying antimony profile (left) assuming random impurity incorporation in each row. Normalizing the antimony pairs to the squared mean antimony fraction exceeds the number of lattice vectors (right). Changing the normalization to the mean of the squared antimony fraction, on the other hand, makes the ratio of antimony pairs to lattice vectors equal to unity, as expected for a random distribution. The exponential decay length was chosen to mimic the bulk antimony profile of our superlattice.

the antimony profile chosen to mimic the segregation profile reconstructed from experimental data, but other function forms (e.g. linear ramp, saw tooth) further substantiate that $\sum_i x_i^2$ provides the desired normalization.

It's interesting to consider the mathematical basis behind (4.2a), by rewriting the term inside the brackets we get

$$
\begin{aligned}
\mathrm{g}_2(|\mathbf{r}_2-\mathbf{r}_1|) &= \frac{1}{\sum_i x_i^2}\left[\sum_i \frac{N^i_{Sb\text{-}Sb\ pairs}(|\mathbf{r}_2-\mathbf{r}_1|)}{N_{anion\ pairs}(|\mathbf{r}_2-\mathbf{r}_1|)}\right] \\
&= \frac{1}{\sum_i x_i^2}\sum_i \frac{N^i_{Sb\text{-}Sb\ pairs}(|\mathbf{r}_2-\mathbf{r}_1|)}{x_i^2\, N_{anion\ pairs}(|\mathbf{r}_2-\mathbf{r}_1|)}\, x_i^2 \qquad (4.2b) \\
&= \frac{1}{\sum_i x_i^2}\sum_i \mathrm{g}_2^i(|\mathbf{r}_2-\mathbf{r}_1|)\, x_i^2 \quad ,
\end{aligned}
$$

with the identification that (4.1) calculated for a single monolayer can be written as

$$
\mathrm{g}_2^i(|\mathbf{r}_2-\mathbf{r}_1|) = \frac{1}{x_i^2}\left[\frac{N^i_{Sb\text{-}Sb\ pairs}(|\mathbf{r}_2-\mathbf{r}_1|)}{N_{anion\ pairs}(|\mathbf{r}_2-\mathbf{r}_1|)}\right] \quad . \qquad (4.2c)
$$

So, our correlation function, (4.2a), in a modulated structure is the expectation value of a single monolayer correlation function calculated over the population of squared antimony fractions. As expected, (4.2a) reduces to (4.1) in the isotropic case (bulk alloy) where x is independent of monolayer, since x^2 can be brought in front of the sums. And as we will see shortly either computational approach produces the same correlation function for bulk experimental data.

It is also useful to briefly look back at Fig. 4.15 and consider the correlation functions as calculated from (4.2a) for the case of perfect order. The [1–10] correlation function (Fig. 4.16, left) remains the same alternating between 0 and 2 for odd and even separation vectors respectively. The [110] correlation function (Fig. 4.16, right), on the other hand will be a constant 1 (or random) instead of 2 (or correlated) which is clearly incorrect. This anomaly stems from the change in normalization from $< x_{Sb} >^2 = {}^1\!/_4$ to $< x_{Sb}^2 >= {}^1\!/_2$, because the antimony fraction for entire rows alternate between 0 and 1. Since a random distribution is undefined whenever all anions in the row are of the same type (either antimony or arsenic). This example is pathological, and illustrates that suitable care must be exercised in applying (4.2a). This is of no practical concern for the superlattice data considered here, however, since cross–incorporation sets a nonvanishing minimum antimony fraction, and the target concentration of 33% antimony sets a corresponding maximum antimony fraction.

To test the sensitivity of the correlation function to these changes we compare full (Fig. 4.18, left) and masked (Fig. 4.20, right) bulk survey images analyzed using both (4.1) and (4.2a) in Fig. 4.22. The only difference between black points in the left and right panels is the introduction of an image mask on the right, with the corresponding reduction in statistics due to fewer counted atoms; aside from statistics the two are very similar. Fig. 4.22 (right) also shows a change from (4.1), which is

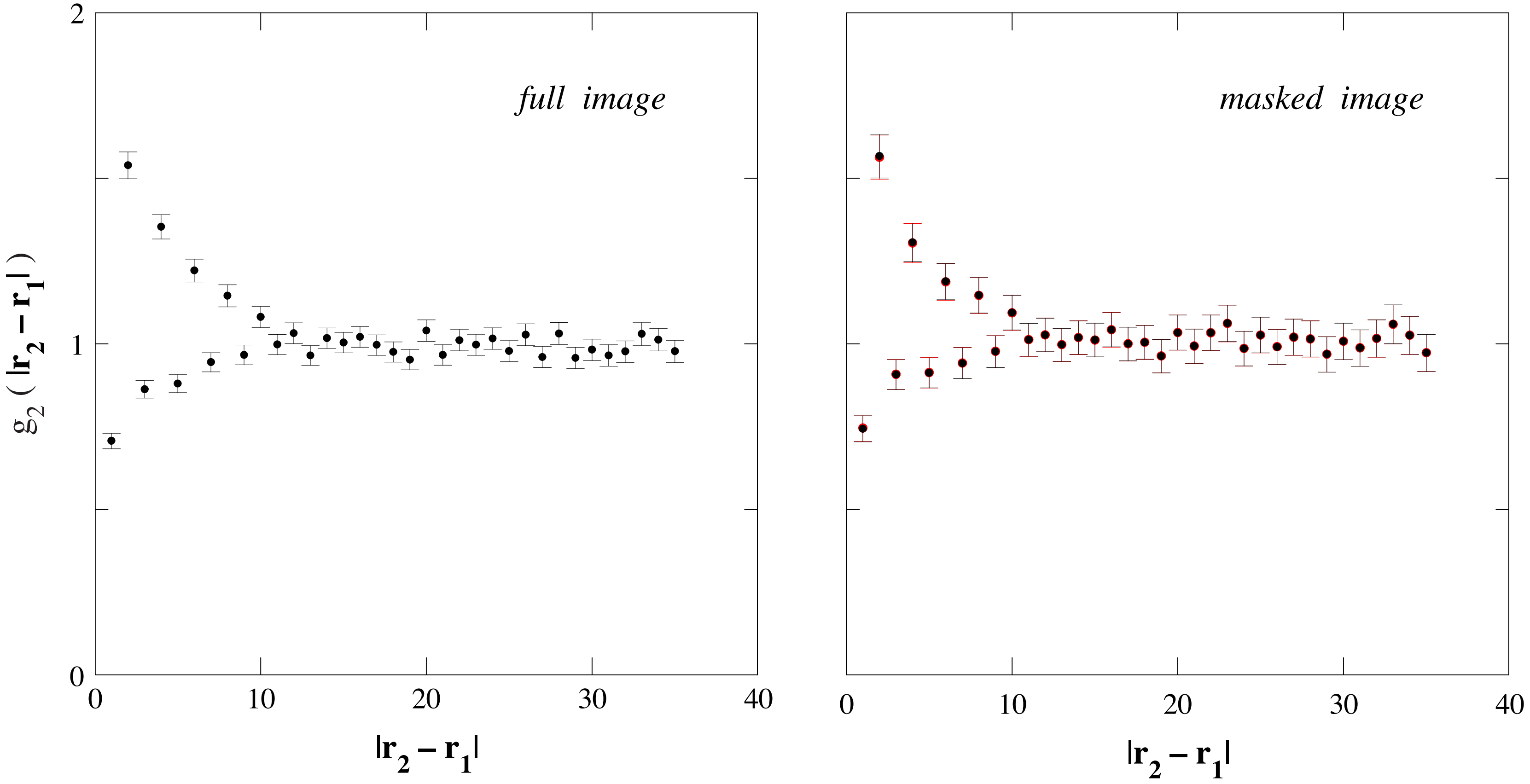


FIGURE 4.22. The antimony–antimony correlation function along [1–10] calculated from (4.1) over full and masked bulk alloy images using Equation 4.1 (left and right in black respectively). Calculating the masked image correlation function using Equation 4.2 (right, red, enlarged slightly for visual purposes) yields an essentially indistinguishable result.

plotted with black points, to (4.2a), plotted with red points, does not effect the resulting correlation function.

There is one final detail, which must be addressed before proceeding with the analysis of experimental data from the superlattice. Interface roughness was shown in Chapter III to be a source of disorder contributing to lateral period fluctuations. This roughness can also introduce the uncontrolled mixing of cleavage–exposed <110> rows with different antimony fractions. To make the point, consider the hypothetical InAs–like / InSb–like structure (illustrated in Fig. 4.23, left) with an abrupt discontinuity in an otherwise uniform and random antimony distribution within both InAs–like ($x_{left} < x_{right}$) and InSb–like ($x_{right} > x_{left}$) regions. We next focus on the row of randomly distributed antimony atoms that mixes antimony fractions and furthermore assume for simplicity that the two fractions are localized to opposite ends of the row in question. We now label the number of antimony pairs as a function of separation vector by whether it refers to the left half of the row, $N_{left}(|\mathbf{r}_2 - \mathbf{r}_1|)$, or the right half of the row $N_{right}(|\mathbf{r}_2 - \mathbf{r}_1|)$; and using Equation 4.2a find

$$\mathrm{g}_2^{left}(|\mathbf{r}_2 - \mathbf{r}_1|) = \frac{1}{x_{left}^2}\left[\frac{N_{left}(|\mathbf{r}_2 - \mathbf{r}_1|)}{\frac{1}{2}\,N_{anion\,pairs}(|\mathbf{r}_2 - \mathbf{r}_1|)}\right] \approx 1\,, \qquad (4.3)$$

and

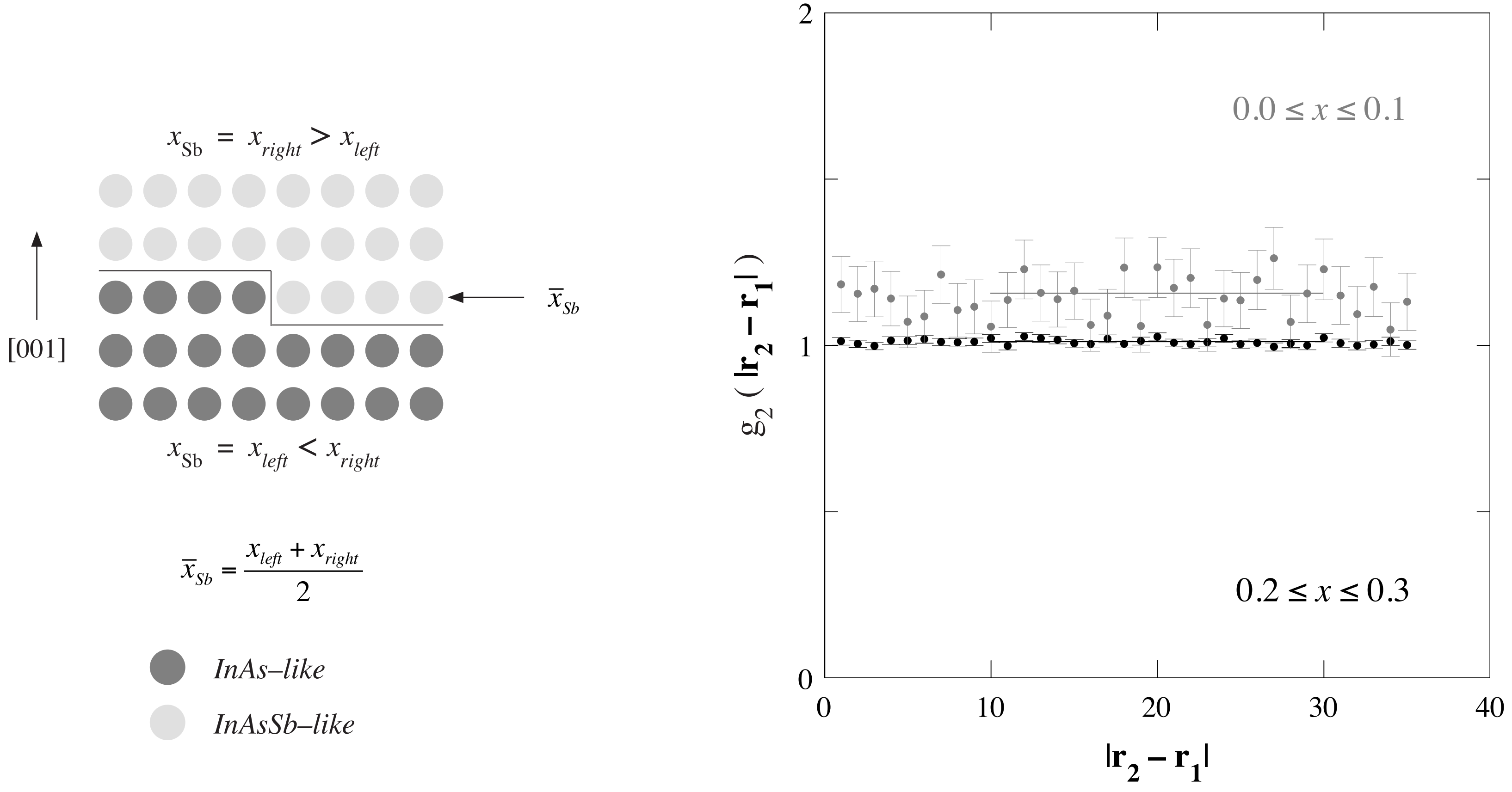


FIGURE 4.23. Monolayer roughness illustrated schematically on the left creates rows where the mean density is very different than portions of said row. A correlation function (right) calculated over a random distribution with exponentially modulated densities as a function of [001] growth direction, and which has had monolayer roughness deliberately introduced, asymptotically approaches an value greater than 1 (<110> uniform). This offset (~15% for small impurity fractions and ~1% for large impurity fractions) must be subtracted to accurately gauge correlations.

$$g_2^{right}(|\mathbf{r}_2 - \mathbf{r}_1|) = \frac{1}{x_{right}^2}\left[\frac{N_{right}(|\mathbf{r}_2 - \mathbf{r}_1|)}{\frac{1}{2}\,N_{anion\,pairs}(|\mathbf{r}_2 - \mathbf{r}_1|)}\right] \approx 1\,. \tag{4.4}$$

The corresponding correlation function for the full row given by (4.2a) (ignoring pairs that straddle the dividing line between the two halves) is

$$g_2(|\mathbf{r}_2 - \mathbf{r}_1|) \approx \frac{1}{\left(\frac{x_{left}+x_{right}}{2}\right)^2}\left[\frac{N_{left}(|\mathbf{r}_2 - \mathbf{r}_1|) + N_{right}(|\mathbf{r}_2 - \mathbf{r}_1|)}{N_{anion\,pairs}(|\mathbf{r}_2 - \mathbf{r}_1|)}\right]. \tag{4.5}$$

We can rewrite (4.3), to isolate $N_{left}(|\mathbf{r}_2 - \mathbf{r}_1|)$,

$$N_{left}(|\mathbf{r}_2 - \mathbf{r}_1|) \approx x_{left}^2\,\frac{N_{anion\,pairs}(|\mathbf{r}_2 - \mathbf{r}_1|)}{2}, \tag{4.6}$$

and (4.4) to isolate $N_{right}(|\mathbf{r}_2 - \mathbf{r}_1|)$,

$$N_{right}(|\mathbf{r}_2 - \mathbf{r}_1|) \approx x_{right}^2\,\frac{N_{anion\,pairs}(|\mathbf{r}_2 - \mathbf{r}_1|)}{2}. \tag{4.7}$$

Plugging (4.6) and (4.7) into (4.5) and simplifying, we obtain,

$$
\begin{aligned}
\mathrm{g}_2(|\mathbf{r}_2-\mathbf{r}_1|) &\approx \frac{1}{\left(\frac{x_{left}+x_{right}}{2}\right)^2}\left[\frac{x_{left}^2+x_{right}^2}{2}\right] \\
&\approx 2\frac{x_{left}^2+x_{right}^2}{x_{left}^2+2x_{left}x_{right}+x_{right}^2} \\
&\approx 2\ \frac{x_{left}^2+2x_{left}x_{right}+x_{right}^2-2x_{left}x_{right}}{x_{left}^2+2x_{left}x_{right}+x_{right}^2} \\
&\approx 2\left[1-\frac{2x_{left}x_{right}}{x_1^2+2x_{left}x_{right}+x_{right}^2}\right] \\
&\approx 1+\left[1-\frac{4x_{left}x_{right}}{x_{left}^2+2x_{left}x_{right}+x_{right}^2}\right],
\end{aligned}
\tag{4.8}
$$

where the term enclosed by brackets in the final result is an expected offset above the <110> uniform result of 1. This term in brackets goes to 0 as expected in the specialized case of $x_{left} = x_{right}$.

The idea behind (4.8) may be generalized for roughness that mixes more than two fractions, and / or multiple rows, but this is best facilitated by numerical simulation. As an experimentally relevant example we simulate antimony atoms, which are distributed randomly in the growth plane with a modulated fraction approximating a segregating antimony profile in the growth direction. A 1 ML shift in the [001] direction is introduced half way through the <110> rows. The correlation functions calculated for two antimony–fraction ranges are shown in Fig. 4.23 (right). The average value between separation vectors of 10 and 30 lattice sites is 1.15 ± 0.02 and 1.011 ± 0.002 for the small and large fraction regimes respectively, much higher than the expected value of 1 (<110> uniform). A larger offset for small antimony fractions was also commonly observed in initial analysis of superlattice correlations; correlations for small *x* more or less set an upper bound on observable offsets, but most superlattice offsets were below 5%. In light of this offset, we will subtract out all offsets that remain in survey–averaged superlattice correlation functions. Interestingly this was not a problem in the bulk, because it contained no antimony–fraction–mixing interfaces to cause an offset.

Pair Correlation Function: Superlattice Alloy Order

We now have the needed toolkit in place with which to analyze antimony–antimony correlations in the superlattice. Taking into account mean fraction squared normalization, offset corrections, and counting window, we find the correlation functions shown in Fig. 4.24, each of which is the average of three surveys over the respective cleavage cross section. This cleavage cross section uniquely determines the direction along which the correlation function is calculated, with (–1–10) implying [1–10] and (1–10) implying [110], so we may adopt the more readily recognized cleavage surface as the distinguishing label for our graphs from here on out.

A key feature of these pair correlation functions is the strong anisotropy. The every–other–lattice–site incorporation observed in the STM image in Fig. 4.2 (left) is directly mirrored in the (–1–10) pair correlation function (Fig. 4.24, left), where the next–nearest–neighbor (encircled in blue) antimony pairs occur more frequently than in a random distribution, whereas nearest–neighbor (encircled in red) antimony pairs occur

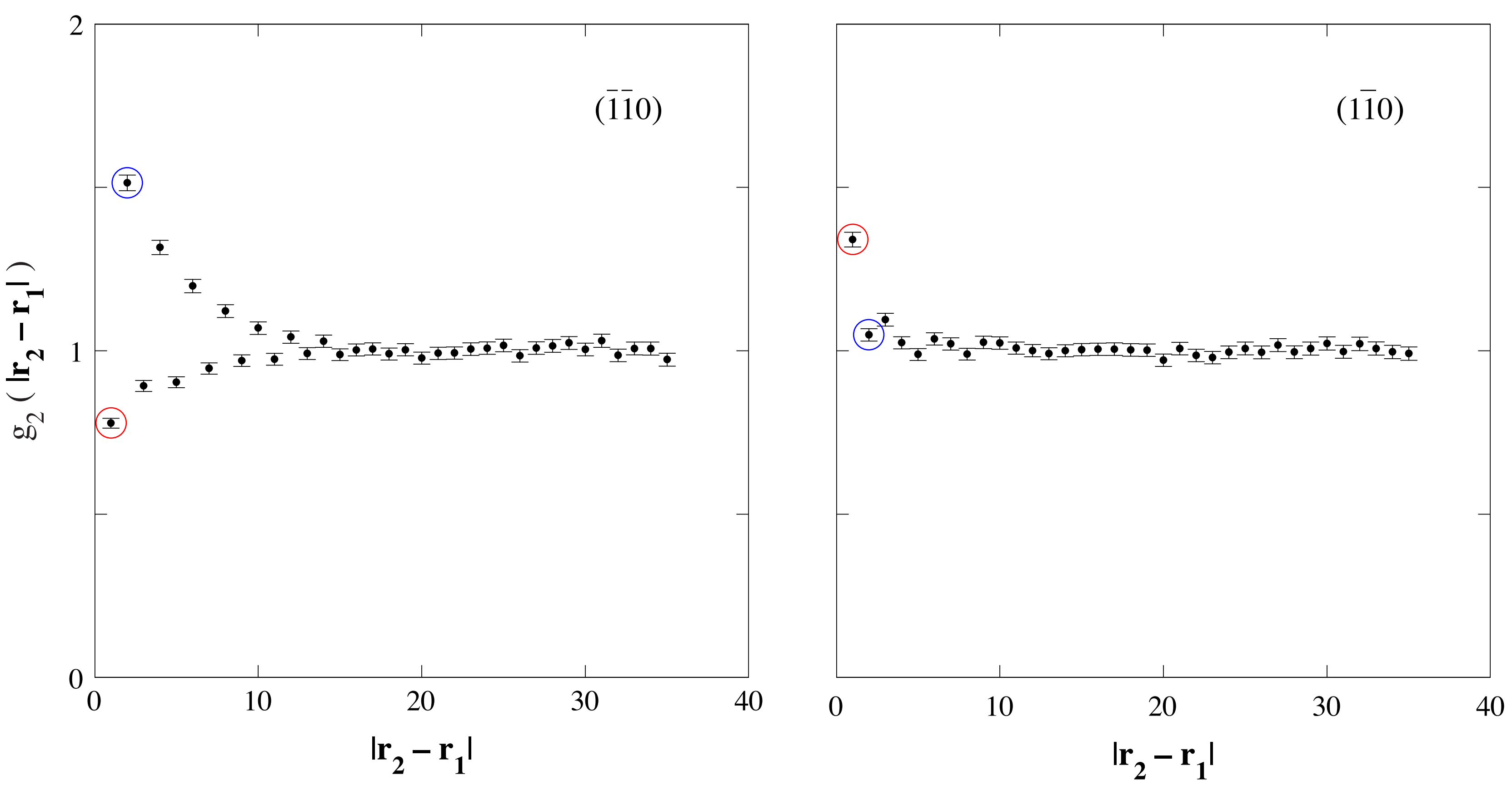


FIGURE 4.24. All–inclusive, superlattice pair correlation functions in (–1–10) and (1–10) cross section (left and right respectively). Next–nearest–neighbor (blue) correlations are prominent in (–1–10) cross section, while in (1–10) cross section the correlations are nearest–neighbor (red). Colors correspond to those used in Fig. 4.2.

less often than a random distribution predicts. Likewise, the chains of antimony atoms observed in the STM image in Fig. 4.2 (right) are echoed by the (1–10) correlation function (Fig. 4.24, right), where nearest–neighbor (encircled in red) antimony pairs are more likely than random, but any two antimony atoms separated by more than one lattice site are essentially randomly distributed. The anisotropic DFT highlighting a streak of excess power in (–1–10) cross section (Fig. 4.6, left) that was absent in (1–10) cross section (Fig. 4.6, right) is similarly represented in the pair correlation functions. The pair correlation functions provide much more detail about the arrangement of atoms than the DFT though, the significance of correlations at any given separation vector can be judged by the extent above or below 1 (random) relative to their statistical error.

Contrasting the nearest–neighbor correlations on the two cleavage surfaces, we note that in (–1–10) cross section nearest–neighbor site occupancy is 0.77 times as likely as random, whereas in the orthogonal (1–10) cross section it is 1.34 times more likely than random; the product of the two (1.03) however is nearly random. These compensating nearest–neighbor probabilities in deficit or excess of random suggest the nearest–neighbor correlations in orthogonal cross sections are complementary, and we explore this possibility in more detail later in the chapter.

Returning to the bulk for a moment, the resemblance between superlattice (Fig. 4.24) and bulk (Fig. 4.19) correlation functions (in the [1–10] direction), is such that fits to the first 15 points of superlattice and bulk agree remarkably well, as demonstrated in Fig. 4.25. The upper and lower branches in both the superlattice and the bulk alloy share a common exponential fall–off of approximately 4 lattice sites, suggesting that the

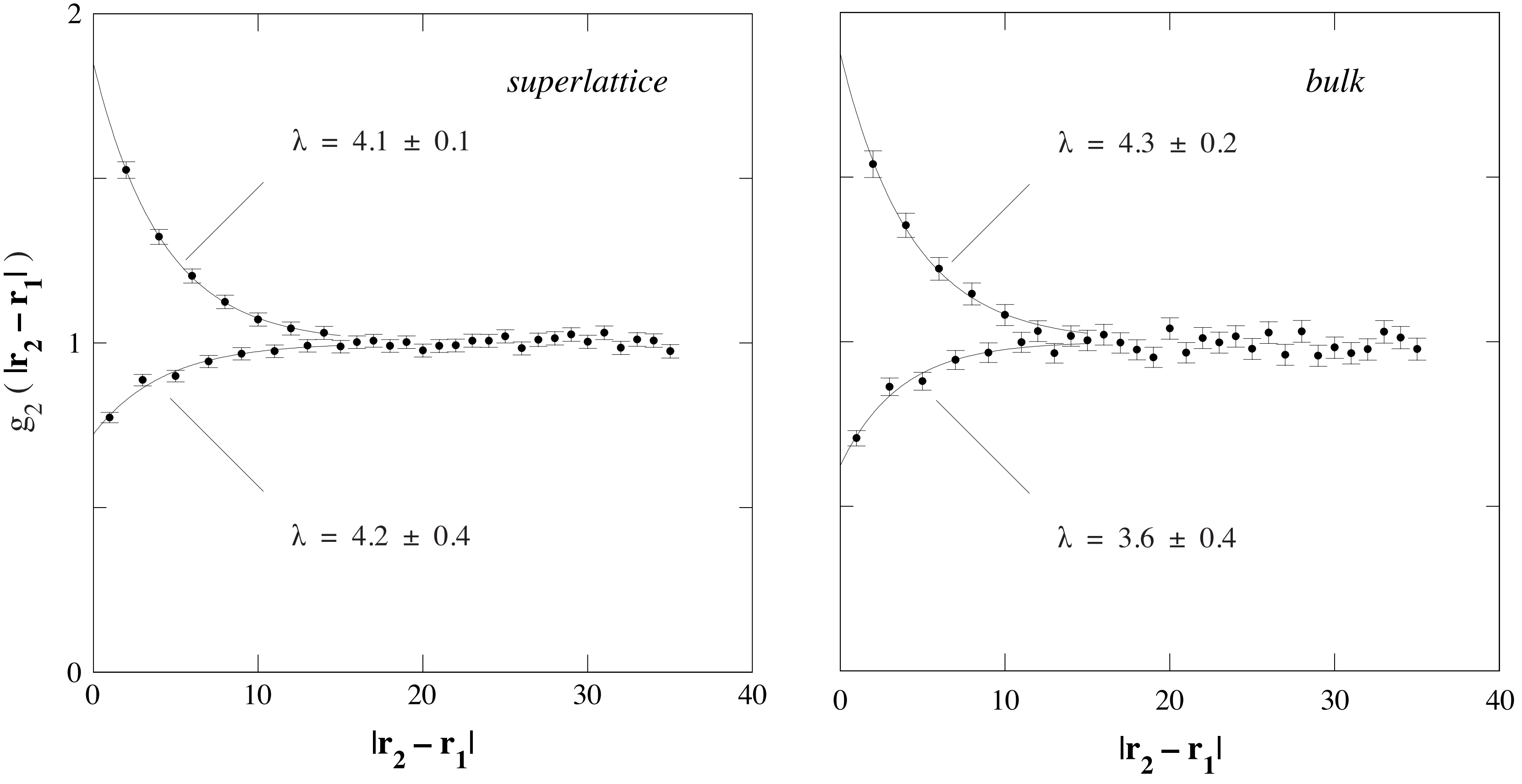


FIGURE 4.25. Pairwise correlation functions in (–1–10) cross section reproduced from Fig. 4.19 (bulk) and Fig. 4.24 (superlattice). First fifteen points of upper (more likely than random) and lower (less likely than random) branches fit independently to exponential functions in both the superlattice (left) and the bulk (right) share common fall–offs of approximately 4 lattice sites.

correlations in the two systems might have a common source; the absence of (1–10) companion data from the bulk, due to the limited availability of material, is unfortunate.

It is tempting to equate these exponential fall–offs with the interface roughness correlation lengths inferred in Chapter III (ranging from 2 nm to 7 nm), however it is unlikely that the alloy order length scale and the roughness disorder length scale are attributable to the same mechanism given the isotropic nature of the interface roughness observed in Chapter III and the extreme anisotropy of the alloy order (Fig. 4.24).

We will circle back to the thought that short–range order in the bulk alloy and the superlattice share a common origin again at the end of the chapter, where correlations in the bulk alloy will help isolate the source for these correlations, but first we explore two experimentally–accessible partitioning schemes of the superlattice data, which will help us pinpoint the sources of this order.

Origins of Superlattice Alloy Order

The monolayer–by–monolayer superlattice composition profile illustrated in Fig. 4.26 displays a strong grading due to segregation of antimony across nominally abrupt InAs / InAsSb interfaces. This antimony accumulates in a surface reservoir of excess, unincorporated antimony while the source is on, with only a fraction of the available antimony incorporating in each bulk monolayer. The prevailing antimony–rich surface reconstructions for common growth conditions, illustrated in Fig. 4.27, both contain 2/3 ML of antimony as an overlayer consisting of dimers aligned with the [110] direction.

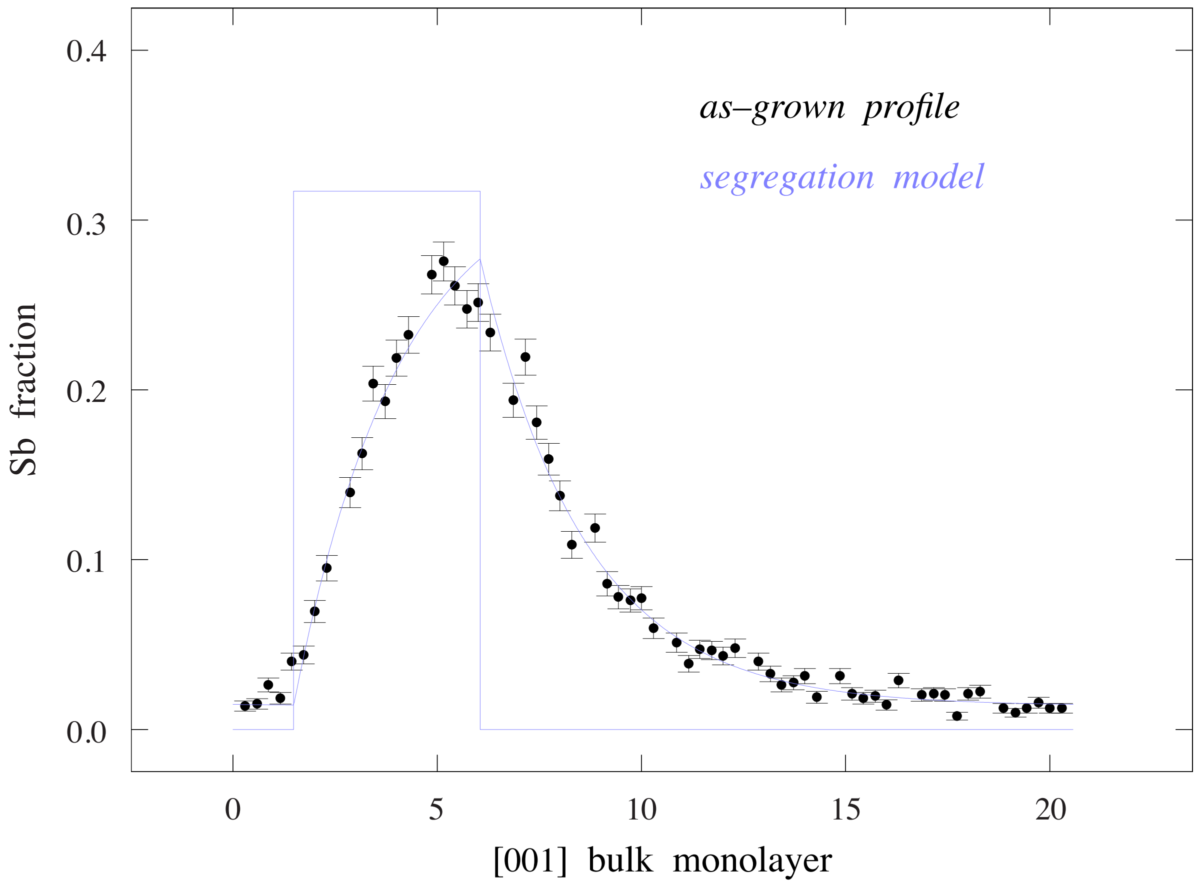


FIGURE 4.26. Monolayer–by–monolayer composition profile reconstructed with STM. Fit to single–source segregation model distinguishes source on (antimony shutter open) from source off (antimony shutter closed). Reprinted with permission from [20].

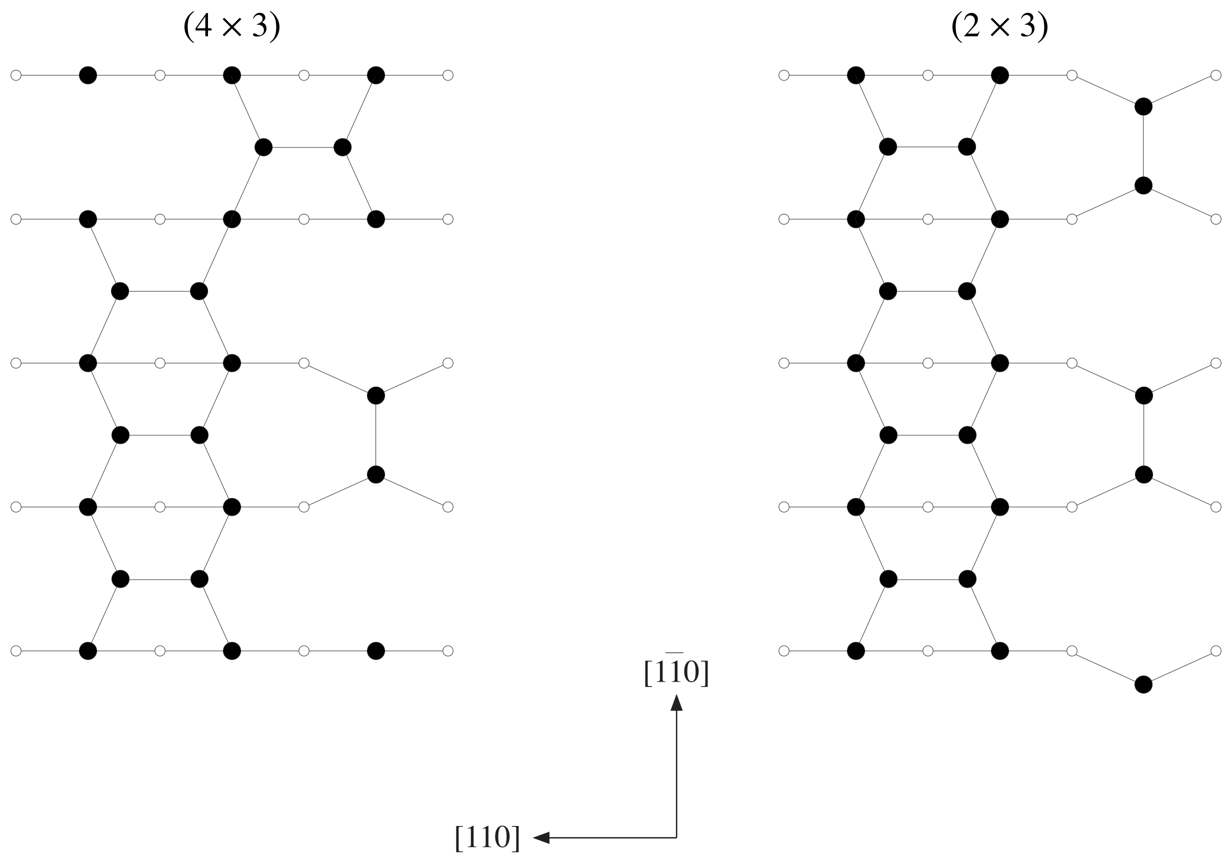


FIGURE 4.27. Common surface reconstructions during antimonide growth with (001) growth surface pointing out of the page. First number represents the periodicity in the [1–10] direction, and second number the periodicity in the [110] direction. One full reconstruction cell is presented for (4x3) and two consecutive (in the [1–10] direction) reconstruction cells are presented for (2x3) reconstruction. Top–layer antimony dimers prefer to align in the [110] direction in both reconstructions. Dimers sitting between [110] rows form a 2/3 ML “floating layer” of unincorporated antimony atoms.

Once the source is turned off the surface reservoir or “floating layer” is slowly depleted as antimony continues to incorporate.

This physical understanding in turn facilitates a detailed mathematical parameterization [23] of the monolayer–by–monolayer incorporation of antimony and the resulting bulk composition profile in terms of an antimony source term originating in the MBE effusion cell illustrated in Fig. 1.3. This source term drives a linear response (segregation) – entirely analogous to the charging and discharging of a RC circuit in response to a voltage pulse – allowing source–on and source–off regimes to be determined from fits to the compositional grading reconstructed with STM [23].

We can thus partition the <110> rows into source–on / source–off ensembles, as illustrated in Fig. 4.28. The source–on and source–off sub–ensembles taken together make up the all–inclusive ensemble, whose correlations on each face were illustrated in Fig. 4.24. The antimony–antimony correlations within these sub–ensembles are qualitatively similar (Fig. 4.29, Fig. 4.30), and bracket the all–inclusive correlations in Fig. 4.24. That the correlations between incorporated antimony atoms are so similar for two very different physical situations indicates the presence of an incoming antimony vapor stream has no bearing. That these correlations persist when the source is off, suggest the observed short–range order likely originates with the floating layer of antimony dimers atop the surface reconstructions illustrated in Fig. 4.27.

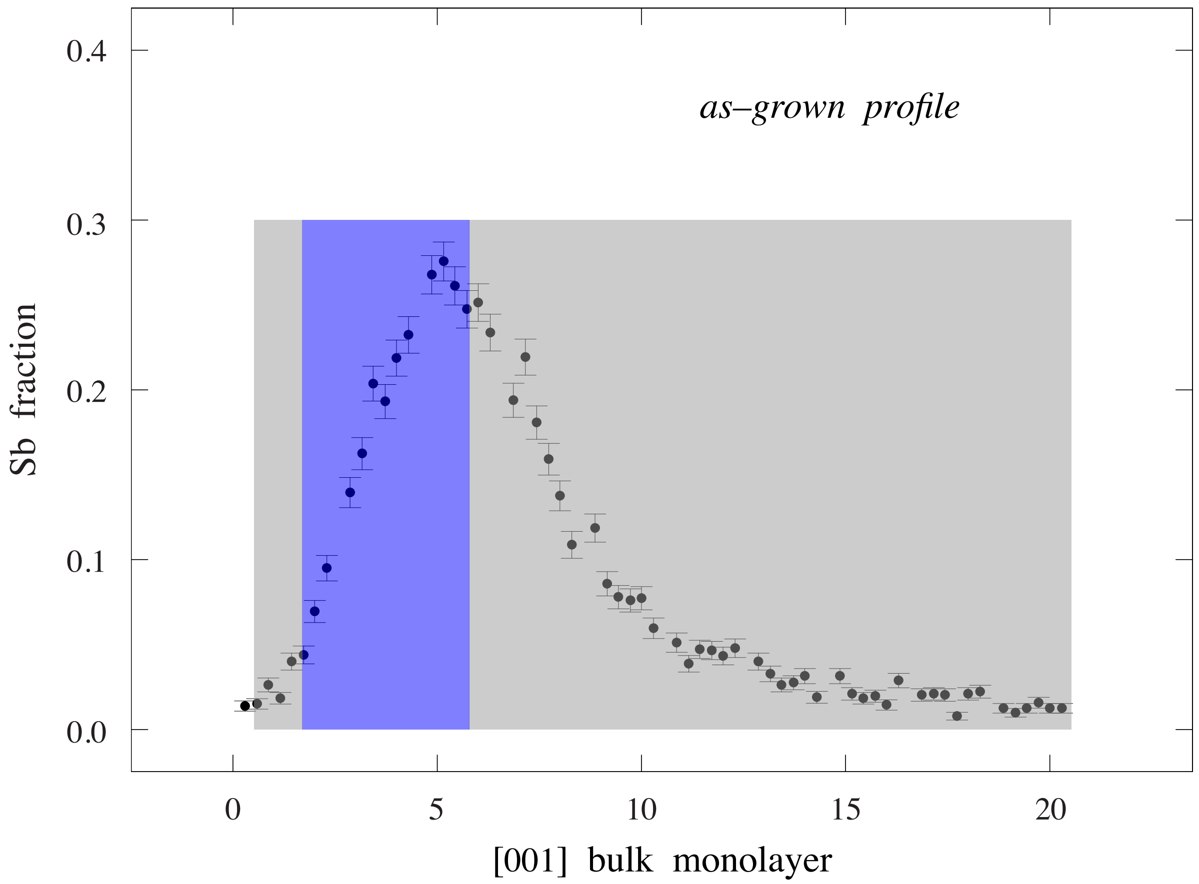


FIGURE 4.28. Partitioning of the as–grown antimony profile based on the single–source fit in Fig. 4.26. Monolayers deposited with antimony vapor source on are highlighted in blue, those deposited with source off are highlighted in grey.

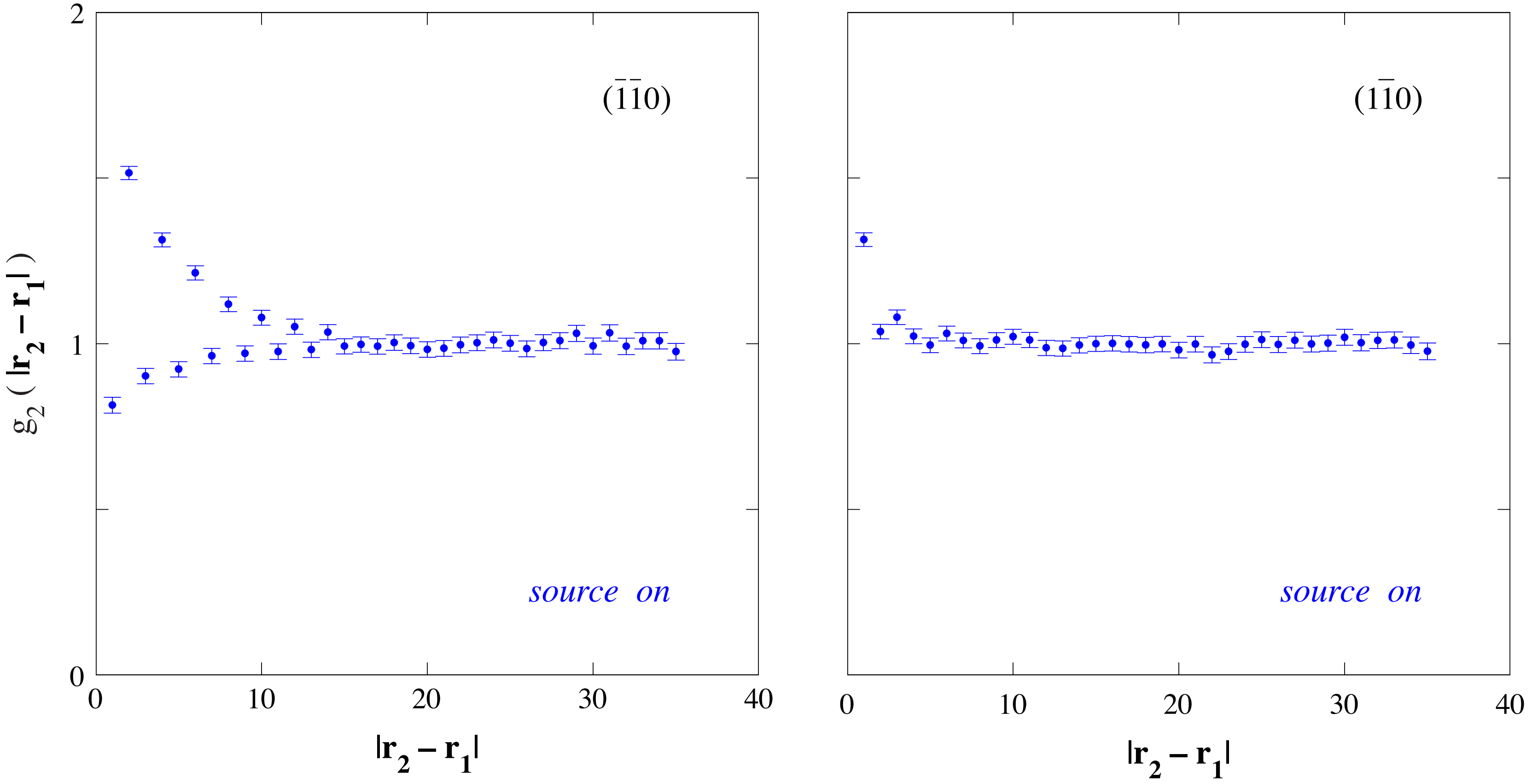


FIGURE 4.29. Antimony pair correlation functions in (–1–10) and (1–10) cross sections (left and right respectively) for the source–on ensemble (indicated in Fig. 4.28) corresponding to bulk monolayers deposited with antimony and arsenic source shutters (Fig. 1.3) open.

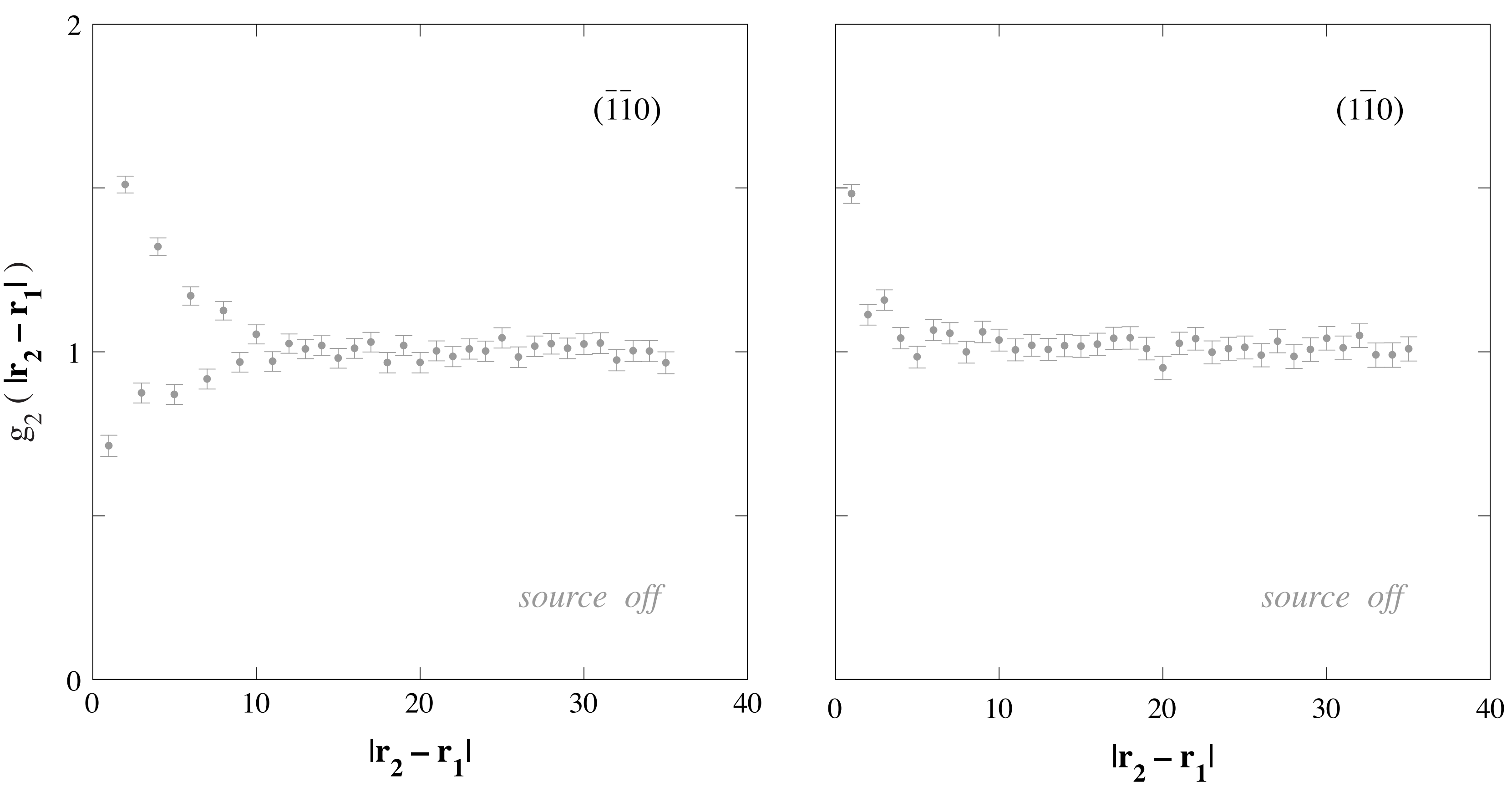


FIGURE 4.30. Antimony pair correlation functions in (–1–10) and (1–10) cross sections (left and right respectively) for the source–off ensemble (indicated in Fig. 4.28) corresponding to bulk monolayers deposited with antimony source shutter (Fig. 1.3) closed and arsenic source shutter open.

The one–to–one connection between alloy composition and [001] strain in this coherently–strained system[8], illustrated in Fig. 4.31, permits another physically meaningful and potentially relevant layer–by–layer partitioning of the antimony pair correlations. A natural division for any strain–balanced sample occurs at 0% strain, thereby dividing the profile into tensile and compressive regimes; pair statistics in the compressive regime (larger antimony fraction equates to more pairs and better statistics) support a further division into weakly and strongly compressive (separated at 1.5% strain).

The correlation functions for these strain–resolved ensembles are all qualitatively similar to one another as well as the all–inclusive ensemble (Fig. 4.24), from which they were drawn. A closer look, however, shows a clear progression in the strength of the correlations. Comparing the strongly–compressive partition shown in Fig. 4.32 to the weakly–compressive partition shown in Fig. 4.33, we observe that the antimony–antimony correlations strengthen as the strain decreases. Likewise, comparing the weakly–compressive partition (Fig. 4.33) with the weakly–tensile one (Fig. 4.34), we again see that antimony–antimony correlations strengthen as the strain decreases. These observations hold true on both (–1–10) and (1–10) surfaces. This same bulk–strain dependence also accounts for the (small) difference between source–on and source–off ensembles in Fig. 4.29 and Fig. 4.30. The source–off ensemble, which shows

---

[8] The InAs / InAsSb strain profile directly mirrors the antimony fraction on account of a near degeneracy [44]in InAs and InSb Poisson ratios that, in turn, causes the [001] lattice constant of coherently–strained InAsSb on GaSb to depend (nearly) linearly on composition. This linear relationship simplifies the conversion from composition to strain, however it is not essential to the arguments that follow.

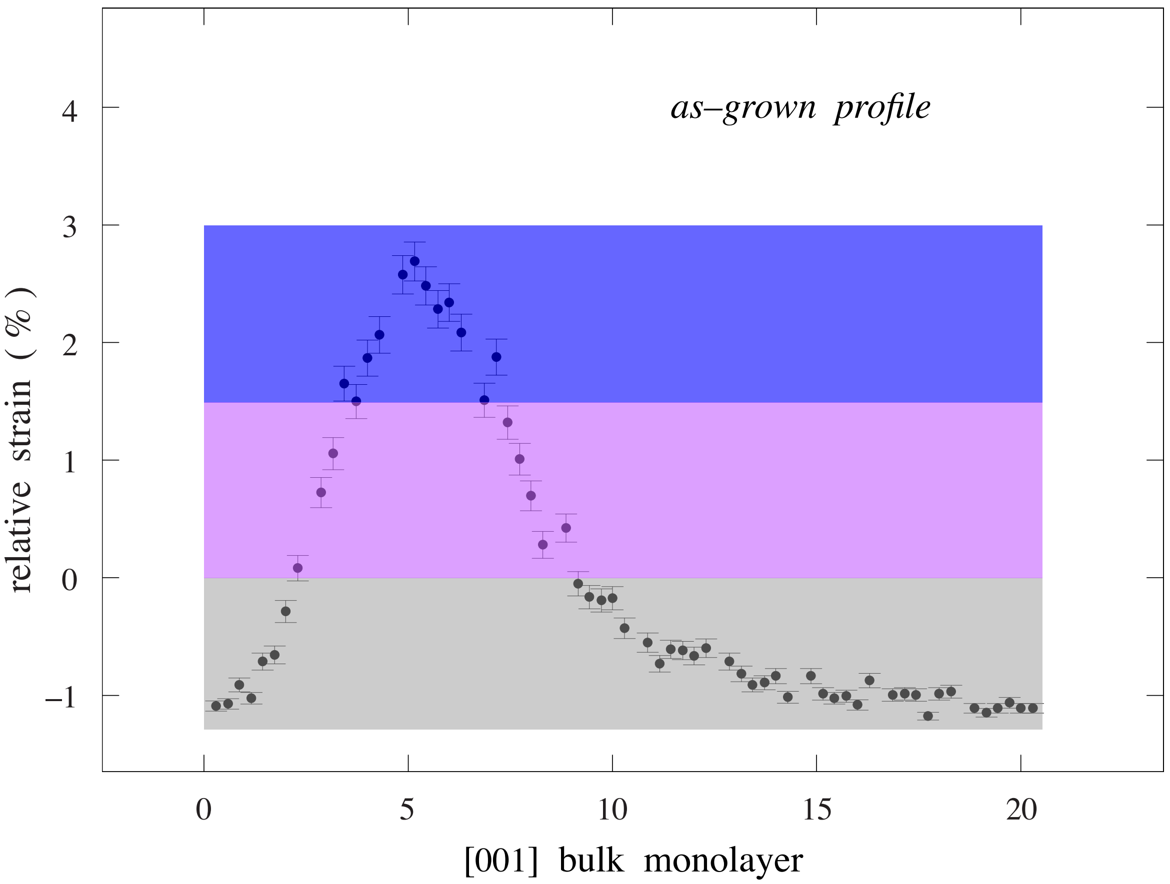


FIGURE 4.31. The one–to–one mapping between composition and monolayer–by–monolayer strain permits translation of the composition profile in Fig. 4.26 into a corresponding [001] strain profile illustrated here. This profile may be partitioned into tensile (grey, strain < 0.0 %), weakly–compressive (magenta, 0.0 % < strain < 1.5%), and strongly–compressive (blue, 1.5% < strain) ensembles. The division between weakly and strongly compressive is set at one half of the maximum observed strain.

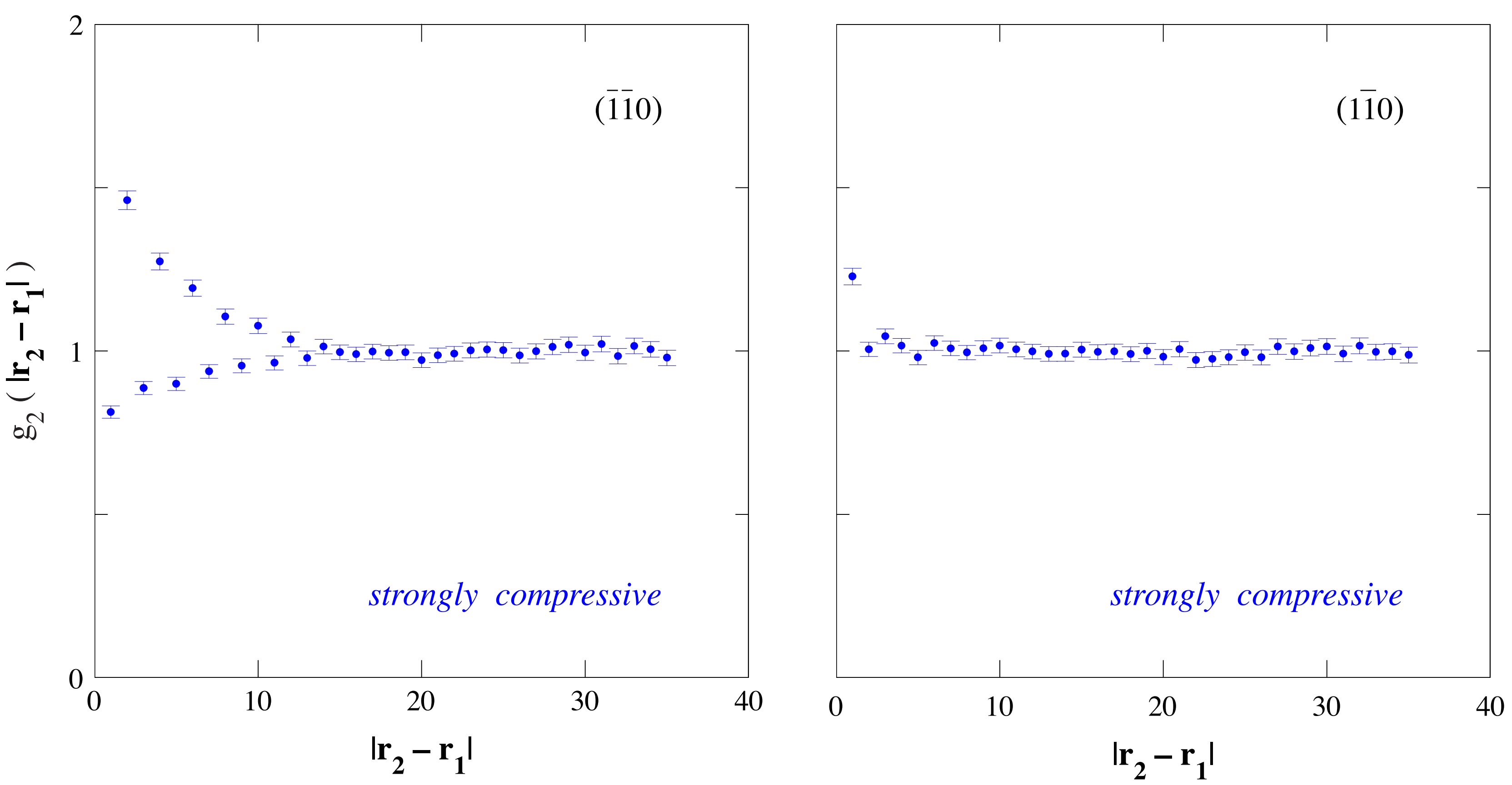


FIGURE 4.32. Antimony pair correlation functions in (–1–10) and (1–10) cross sections (left and right respectively) within bulk monolayers under strong compression (Fig. 4.31, strain > 1.5 %).

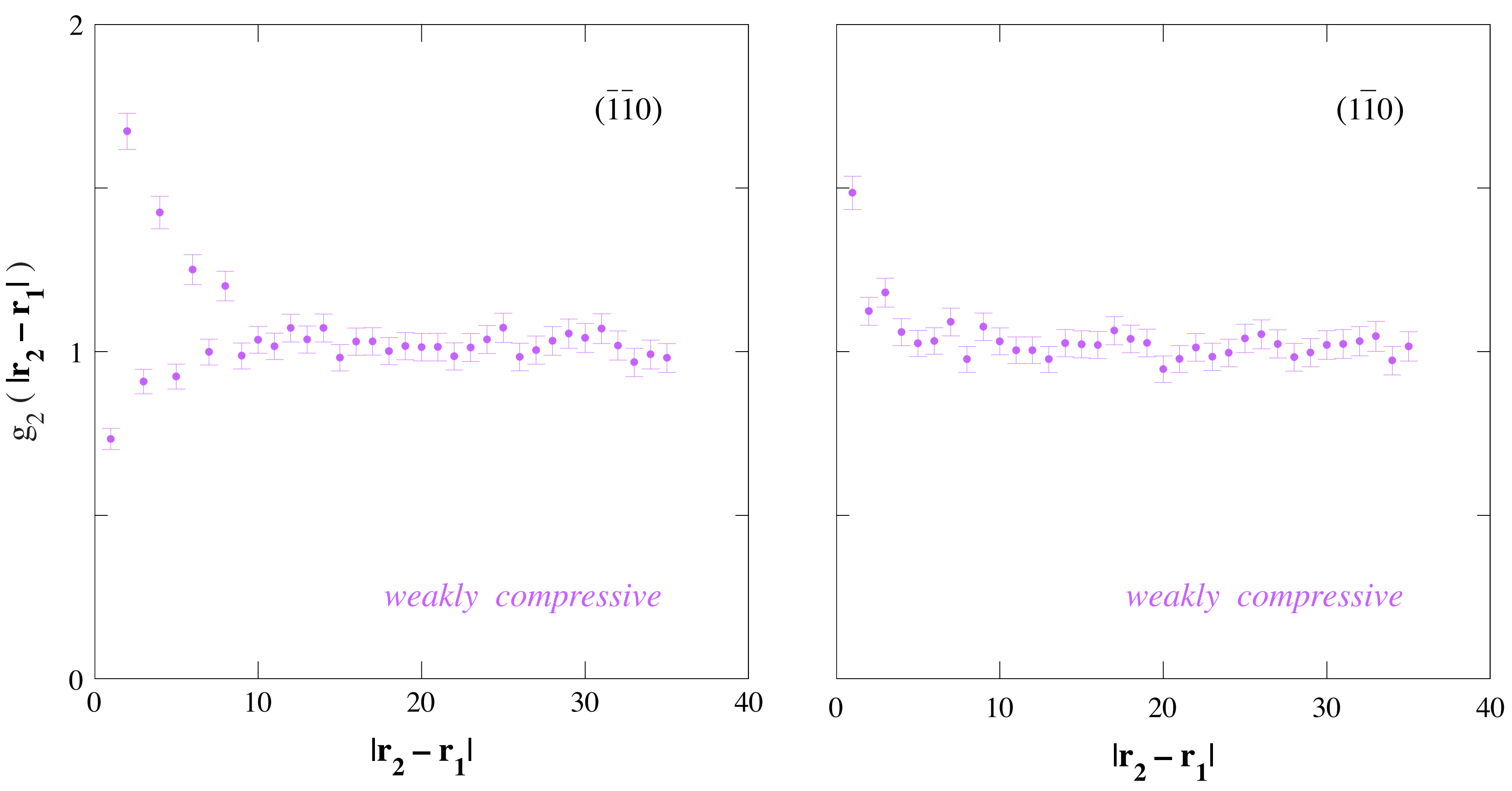


FIGURE 4.33. Antimony pair correlation functions in (–1–10) and (1–10) cross sections (left and right respectively) within bulk monolayers under weak compression (Fig. 4.31, 0.0 % < strain < 1.5%).

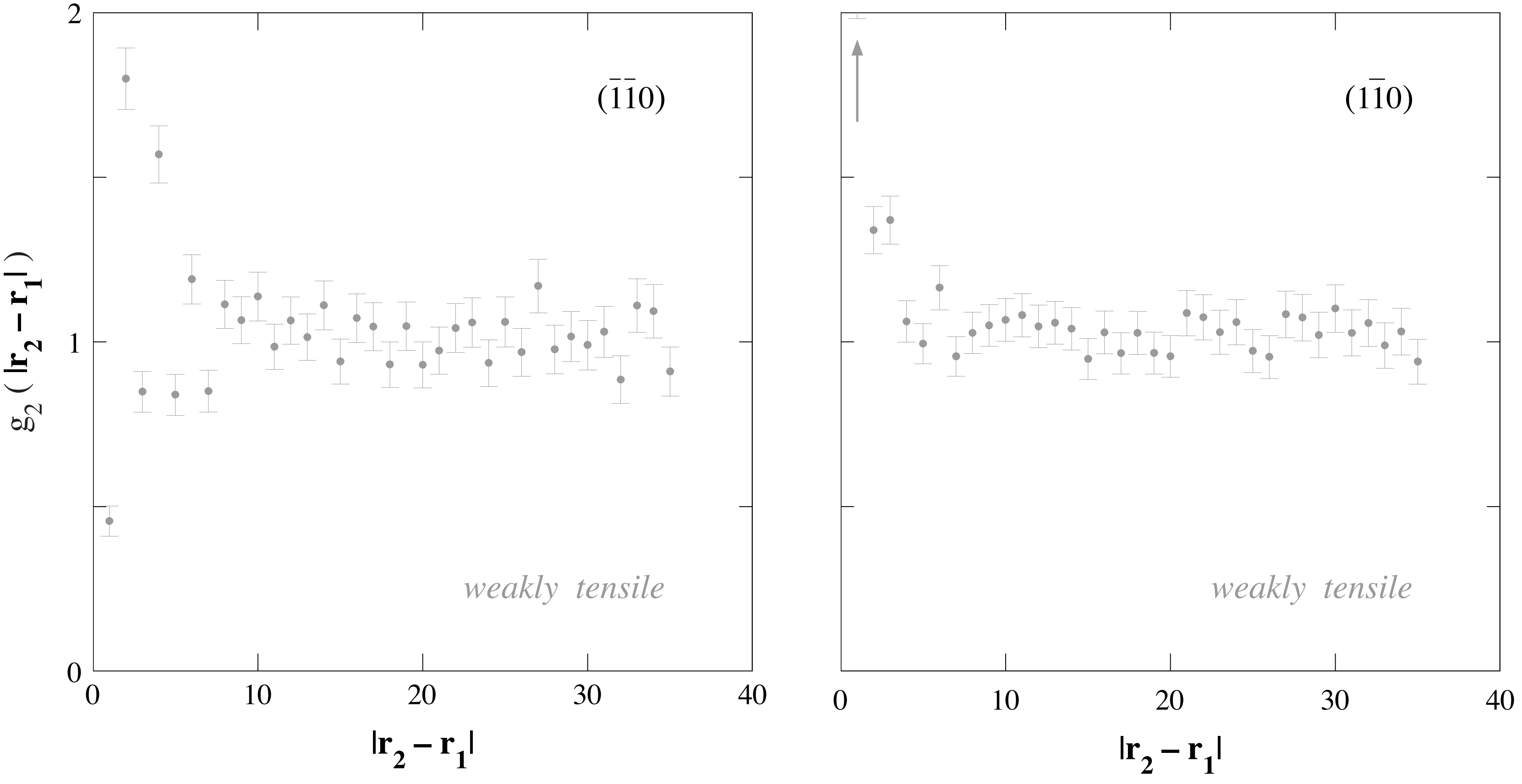


FIGURE 4.34. Antimony pair correlation functions in (−1−10) and (1−10) cross sections (left and right respectively) within bulk monolayers under weak tensile strain (Fig. 4.31, strain < 0.0 %).

(somewhat) stronger correlations, includes more tensile rows, thereby reducing the average strain in that ensemble.

The fall–off exhibited by these sub–ensembles is similar to the fall–off in the all–inclusive. In fact individual fits to (–1–10) correlation functions for separation vectors 2–15 in each sub–ensemble are consistent with an exponential decay of 4.1 lattice sites.

Antimony Pair Interactions: Superlattice

We now explore the compensating behavior of nearest–neighbor correlations in orthogonal cross sections, noted in connection with Fig. 4.24, across each of the experimental ensembles described above. To better visualize this compensating behavior, we turn to the natural logarithm of (–1–10) and (1–10) nearest–neighbor correlations in Fig. 4.35, where compensating terms will be reflection symmetric about zero. When strain–resolved, source–resolved, and all–inclusive ensembles are plotted against their respective [001] strains, calculated from the nearest–neighbor, population–weighted expectation value

$$< \varepsilon >= \frac{\sum_i \varepsilon_i \, N^i_{Sb-Sb\ pairs}(|\mathbf{r}_2 - \mathbf{r}_1| = 1)}{\sum_i N^i_{Sb-Sb\ pairs}(|\mathbf{r}_2 - \mathbf{r}_1| = 1)}, \qquad (4.9)$$

we see a nearly–linear dependence with unmistakable mirror symmetry (dashed line). This symmetry is quantitatively confirmed once the respective logarithms in (–1–10) and

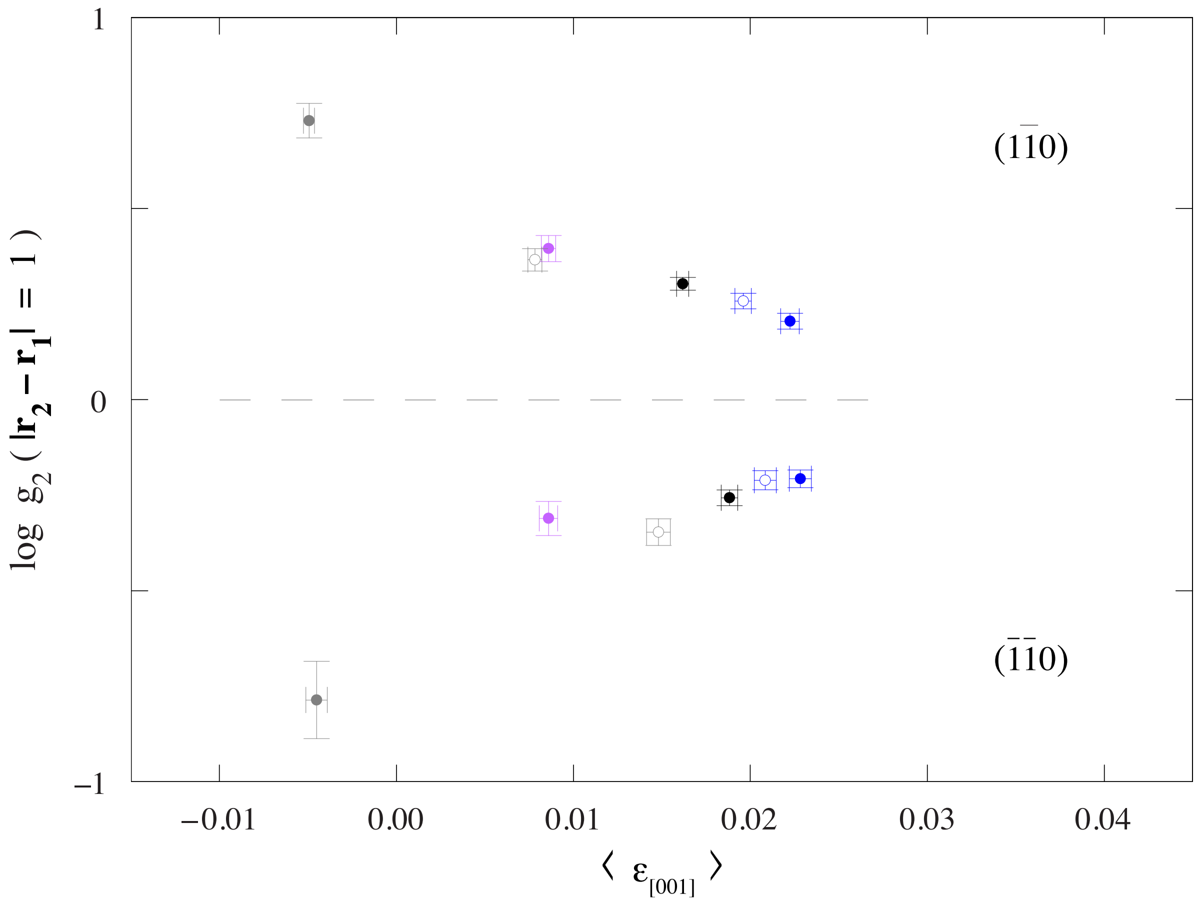


FIGURE 4.35. Logarithm of nearest–neighbor correlations as a function of expected sub–ensemble strain. Horizontal axis, here, corresponds to the vertical axis in Fig. 4.31. Correlations in orthogonal cross sections mirror one another. Solid, colored points are strain–resolved, open points are source–resolved, and solid black points are all–inclusive superlattice ensembles. Vertical and horizontal error bars are propagated from counting statistics.

(1–10) cross section are summed in Fig. 4.36, with the results distributed about zero independent of strain[9].

The three, independent strain–resolved partitions in Fig. 4.35 are simultaneously fit to straight lines with opposite slopes and common strain–intercept to explicitly incorporate the (–1–10) + (1–10) compensation demonstrated in Fig. 4.36. As shown in Fig. 4.37, each of our experimental ensembles lies close to this fit. In the one instance that falls furthest from this line (tensile ensemble), the deviations pair in opposite directions, as they must if they compensate.

The corresponding data for next–nearest–neighbor correlations ($|\mathbf{r}_2 - \mathbf{r}_1| = 2$) is assembled in Fig. 4.38, where (–1–10) and (1–10) correlations are both more likely than random. Independent fits to the strain–resolved ensembles, here, do a strikingly persuasive job of describing source–resolved and all–inclusive ensembles in both cross sections. These linear behaviors hold for pairs separated by 3 ($|\mathbf{r}_2 - \mathbf{r}_1| = 3$, in Fig. 4.39) and 4 ($|\mathbf{r}_2 - \mathbf{r}_1| = 4$, in Fig. 4.40) lattice sites as well. This comparison between cleavage faces fails to be useful at larger separations where the (1–10) correlation function (Fig. 4.24) approaches one.

Let us return, now, to the combined linear fit illustrated in Fig. 4.37 and consider the common horizontal intercept — corresponding to no antimony–antimony correlation whatsoever — at 3.4 ± 0.2 % strain. Why this number? Naïve expectations would place this zero crossing for random incorporation at zero strain, whereas the data in Fig. 4.37

---

[9] The asymmetric (–1–10) and (1–10) errors in Fig. 4.35 are striking in view of the equal number of STM images (53, 52) acquired over each cleavage face, but reflect the disparate numbers of antimony pairs in each cross section due to compensation.

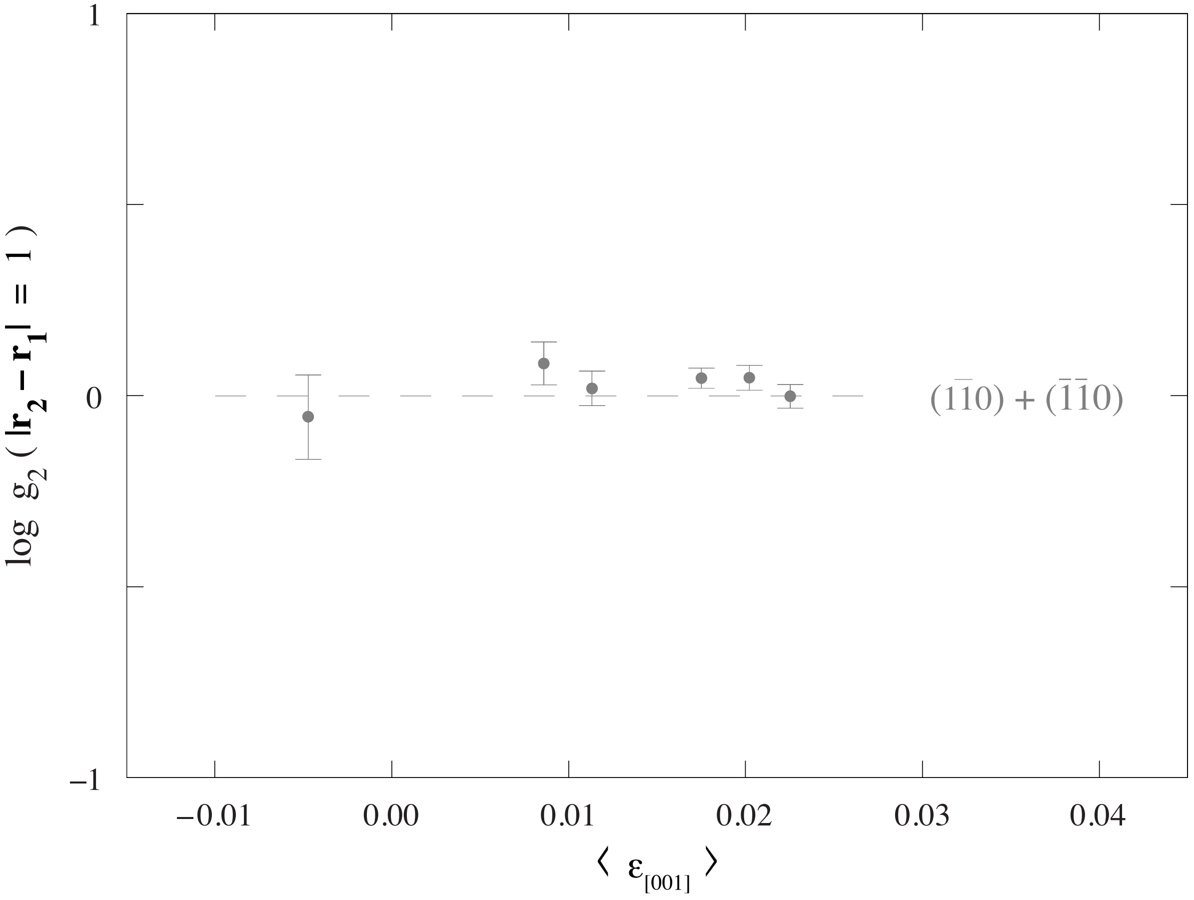


FIGURE 4.36. Logarithm of nearest–neighbor correlations in orthogonal cleavage planes (Fig. 4.35) sum to zero, indicating strain–independent compensation.

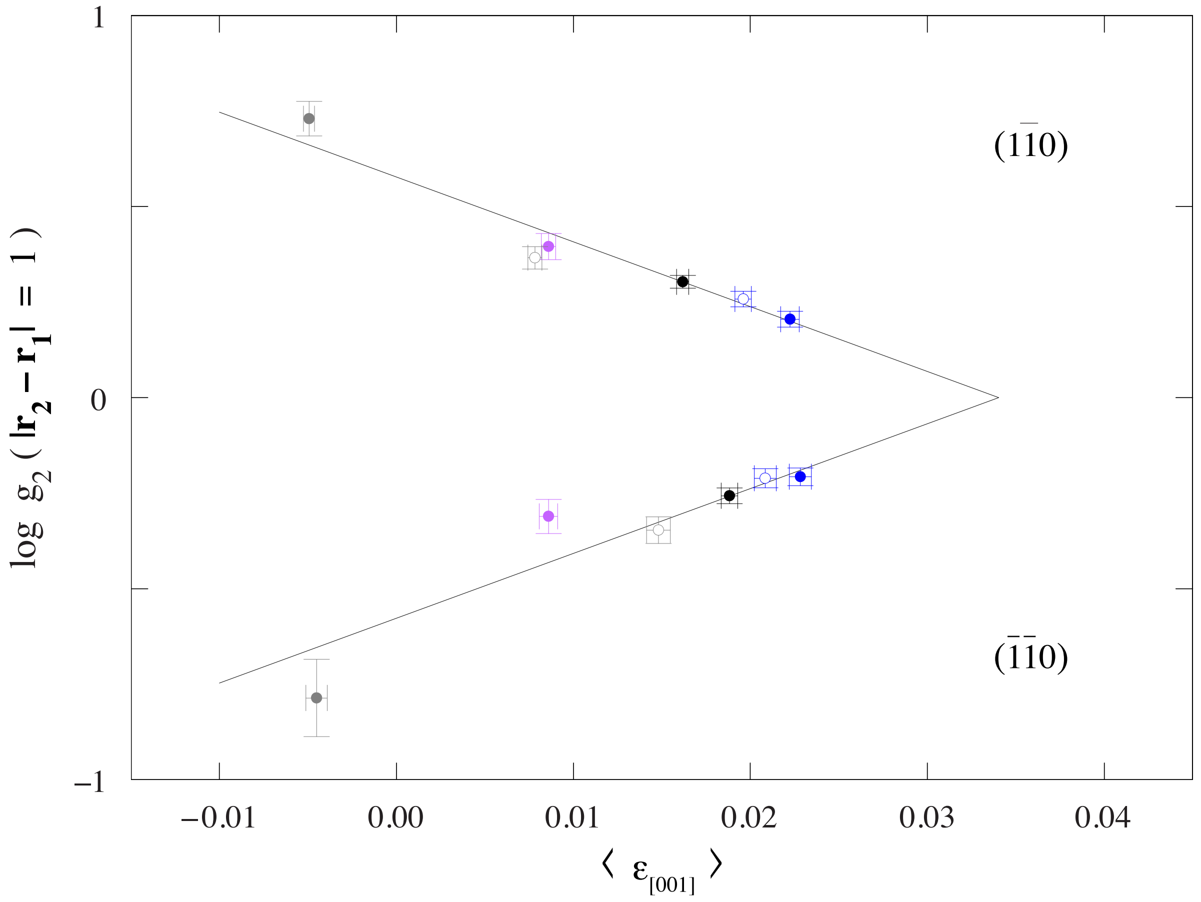


FIGURE 4.37. Nearest–neighbor correlations reproduced from Fig. 4.35. Simultaneous straight line fits to strain–resolved ensembles on complementary cross sections accurately describe the strain ensembles as well as the source–resolved and all–inclusive ensembles. Solid, colored points are strain–resolved, open points are source–resolved, and solid black points are all–inclusive superlattice ensembles. Vertical and horizontal error bars are propagated from counting statistics.

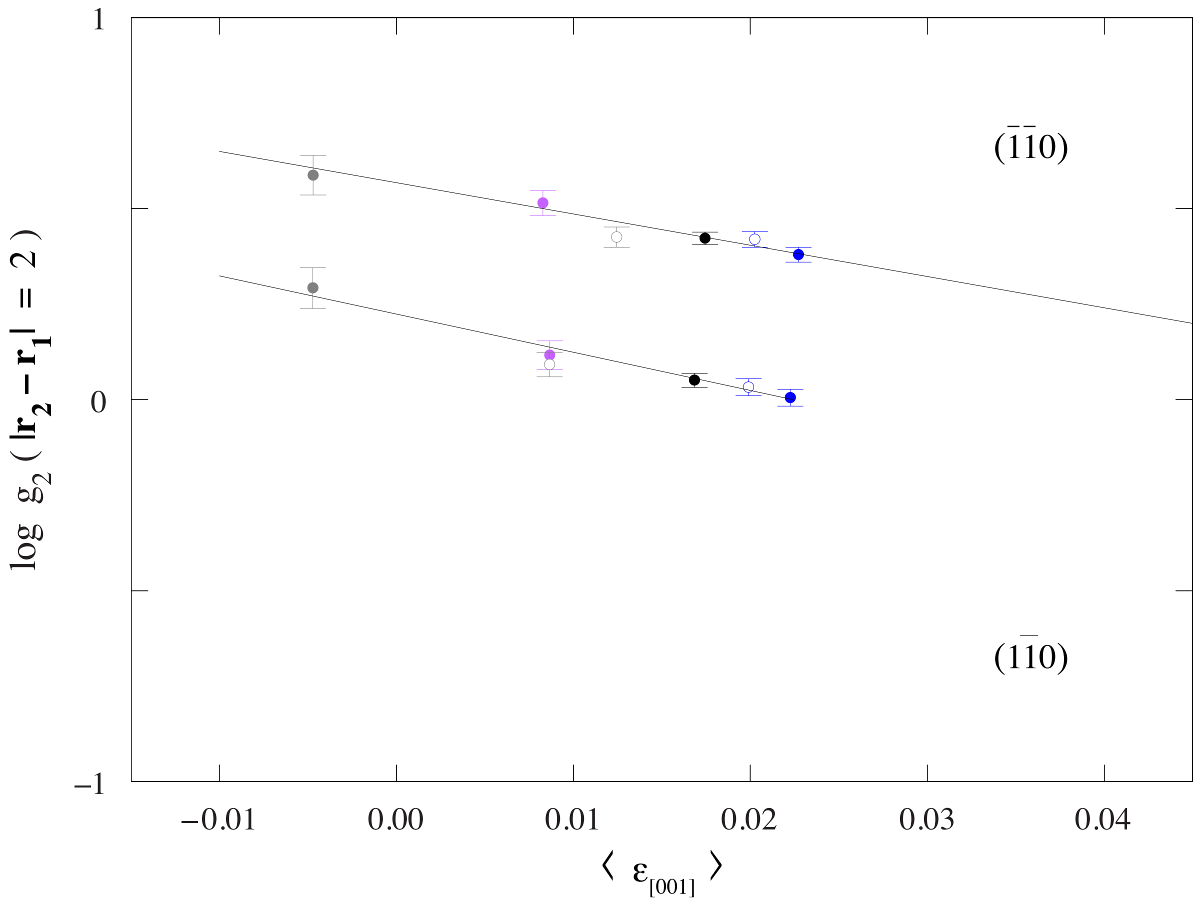


FIGURE 4.38. Next–nearest–neighbor correlations plotted on a log scale. Independent straight line fits to strain–ensembles for each cross section likewise describe source–resolved and all–inclusive ensembles. Solid, colored points are strain–resolved, open points are source–resolved, and solid black points are all–inclusive superlattice ensembles. Vertical error bars are propagated from counting statistics.

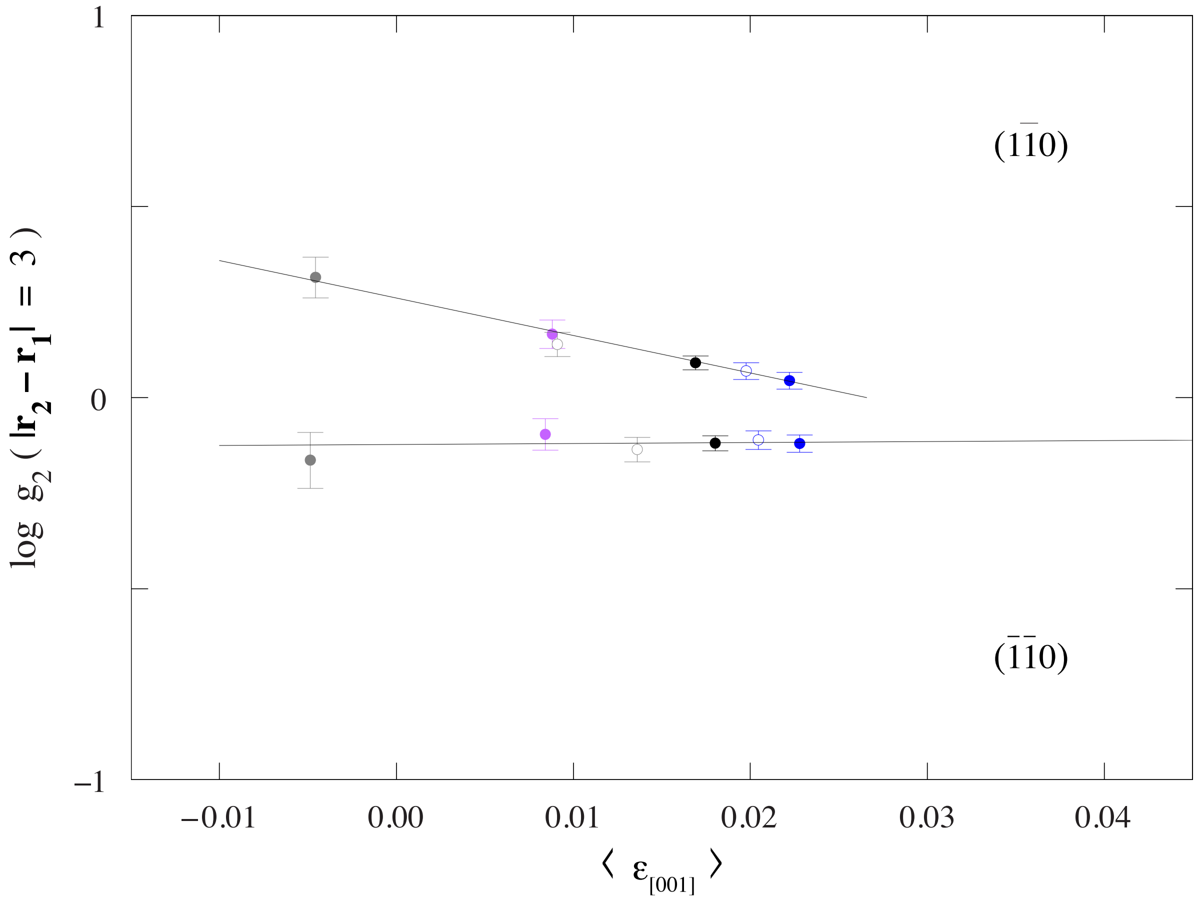


FIGURE 4.39. Correlations for pairs separated by three lattice sites plotted on a log scale. Independent straight line fits to strain–ensembles for each cross section likewise describe source–resolved and all–inclusive ensembles. Solid, colored points are strain–resolved, open points are source–resolved, and solid black points are all–inclusive superlattice ensembles. Vertical error bars are propagated from counting statistics.

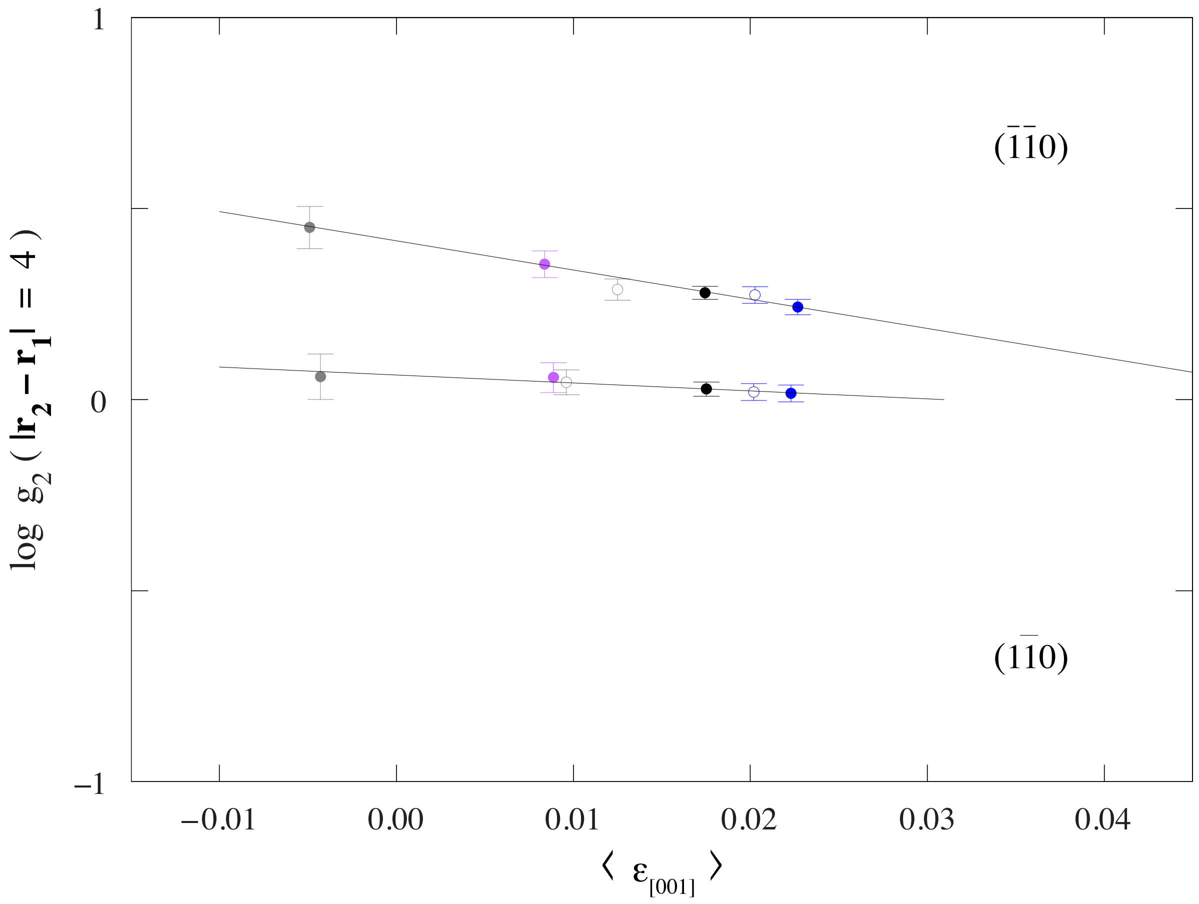


FIGURE 4.40. Correlations for pairs separated by four lattice sites plotted on a log scale. Independent straight line fits to strain–ensembles for each cross section likewise describe source–resolved and all–inclusive ensembles. Solid, colored points are strain–resolved, open points are source–resolved, and solid black points are all–inclusive superlattice ensembles. Vertical error bars are propagated from counting statistics.

show vanishing strain is characterized by comparatively strong pair correlations, bringing the assumption of this parameter as the governing physical variable into question.

It now becomes important to recall the one–to–one correspondence between strain and composition in coherently–strained systems that imposes a physical degeneracy in our choice of independent variable. We previously used the linear correspondence between superlattice composition and strain to define strain–resolved ensembles, but we could have equally well chosen to define composition–resolved ensembles instead. Although the former may, at first, appear more intuitive, our experimental data suggests one must be open to the other.

Converting from ensemble strain to ensemble antimony fraction by replacing ε_i with x_i in (4.9) transforms Fig. 4.37 into Fig. 4.41[10]. The common zero crossing for our combined (–1–10) and (1–10) fit now occurs at $x_{Sb} = 0.325 \pm 0.011$, an instantly recognizable number: the targeted, steady–state antimony fraction for our superlattice. This fraction is directly proportional to the incident antimony flux, but, as we've previously emphasized in connection with Figs. 28, 29, and 30, the pair correlation functions for source on and source off ensembles are nearly indistinguishable, pointing to the surface floating layer, rather than the incoming vapor stream, as the key factor in these nearest–neighbor correlations. In any system evidencing segregation, surface and bulk impurity fractions will be connected through the segregation coefficient R, which may be determined from the compositional grading mapped out with STM [22]; that

---

[10] Horizontal errors seem to be larger when plotting against antimony fraction but the relative errors in Fig. 4.37 and Fig. 4.41 are actually equivalent.

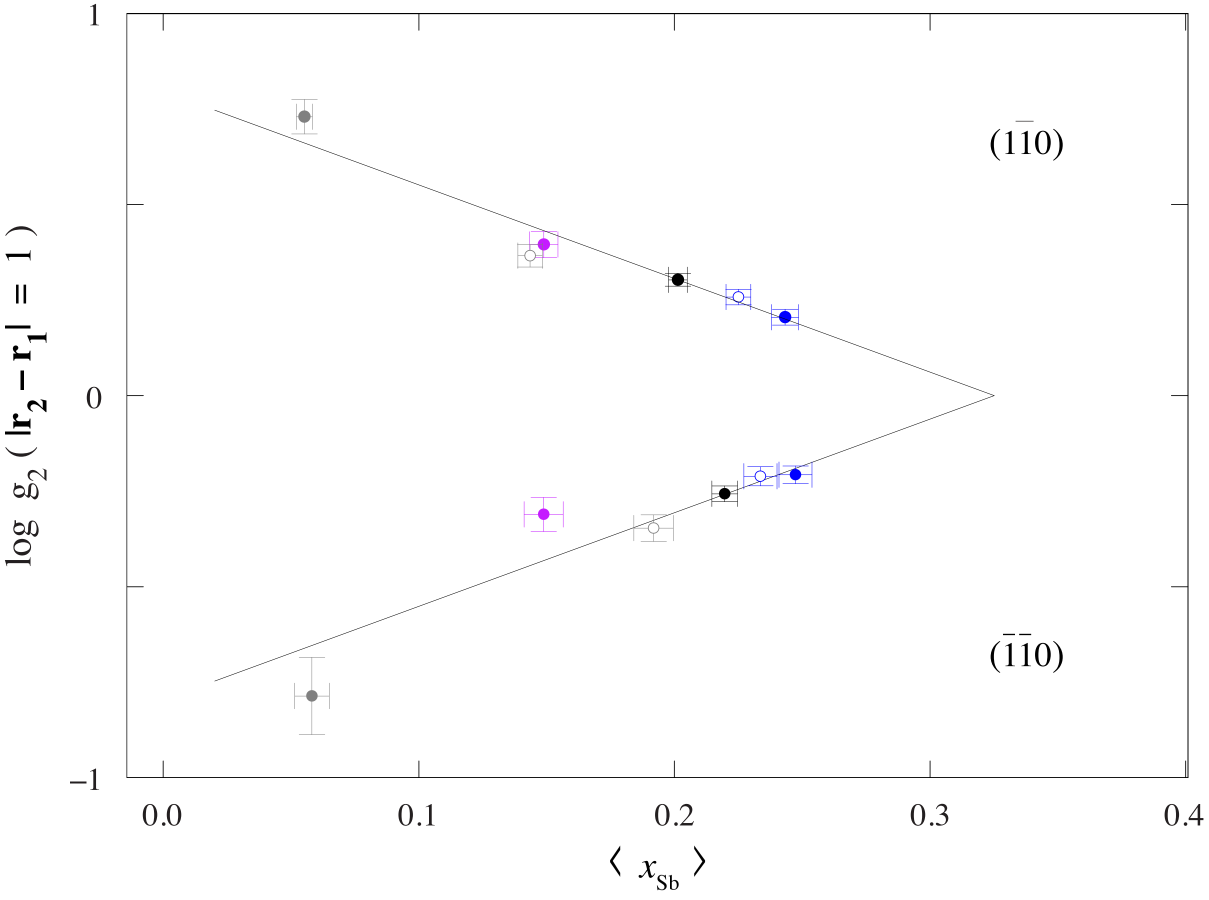


FIGURE 4.41. Nearest–neighbor correlations from Fig. 4.37 plotted against antimony fraction in corresponding ensembles. Solid, colored points are strain–resolved, open points are source–resolved, and solid black points are all–inclusive superlattice ensembles.

number in InAs / InAsSb [20] is very close to 2/3, implying a fixed, monolayer–by–monolayer ratio between surface and bulk–incorporated antimony of 2:1.

The common zero crossing in Fig. 4.37 may therefore be equivalently expressed in terms of its corresponding surface coverage as ~ 2/3 ML, a value whose physical significance is both immediate and independent of source shuttering: it is the maximum amount of antimony that can be accommodated by the antimony–dimer surface reconstructions in Fig. 4.27. The number of [110]–oriented dimers at the growth surface thus appears to be our controlling variable.

Antimony Pair Interactions: Bulk Alloy

Bulk alloy films coherently strained to a virtual substrate offer an independent perspective on the connection between short–range alloy order, bulk strain, and surface antimony fraction, since they enforce a altogether different relationship between strain and antimony fraction than that dictated in the superlattice by the requirement of strain balancing to GaSb. The bulk alloy film made available to us was actually relaxed as opposed to coherently strained[11], and the STM data (taken some 300 nm into the growth) previously summarized in Figs. 4.17 and 4.19.

A relaxed film, by definition, exhibits vanishing (0%) residual strain. As shown in Fig. 4.42, adding this point to Fig. 4.37 makes clear that bulk and superlattice

---

[11] Source temperatures drifted more than usual during growth of the virtual substrate, leading to a template too large for targeted InAsSb alloy composition. Although dislocations necessarily took up that excess strain, substrate and bulk alloy x–ray peaks were of comparable width, suggesting these dislocations were confined to the early stages of bulk alloy growth.

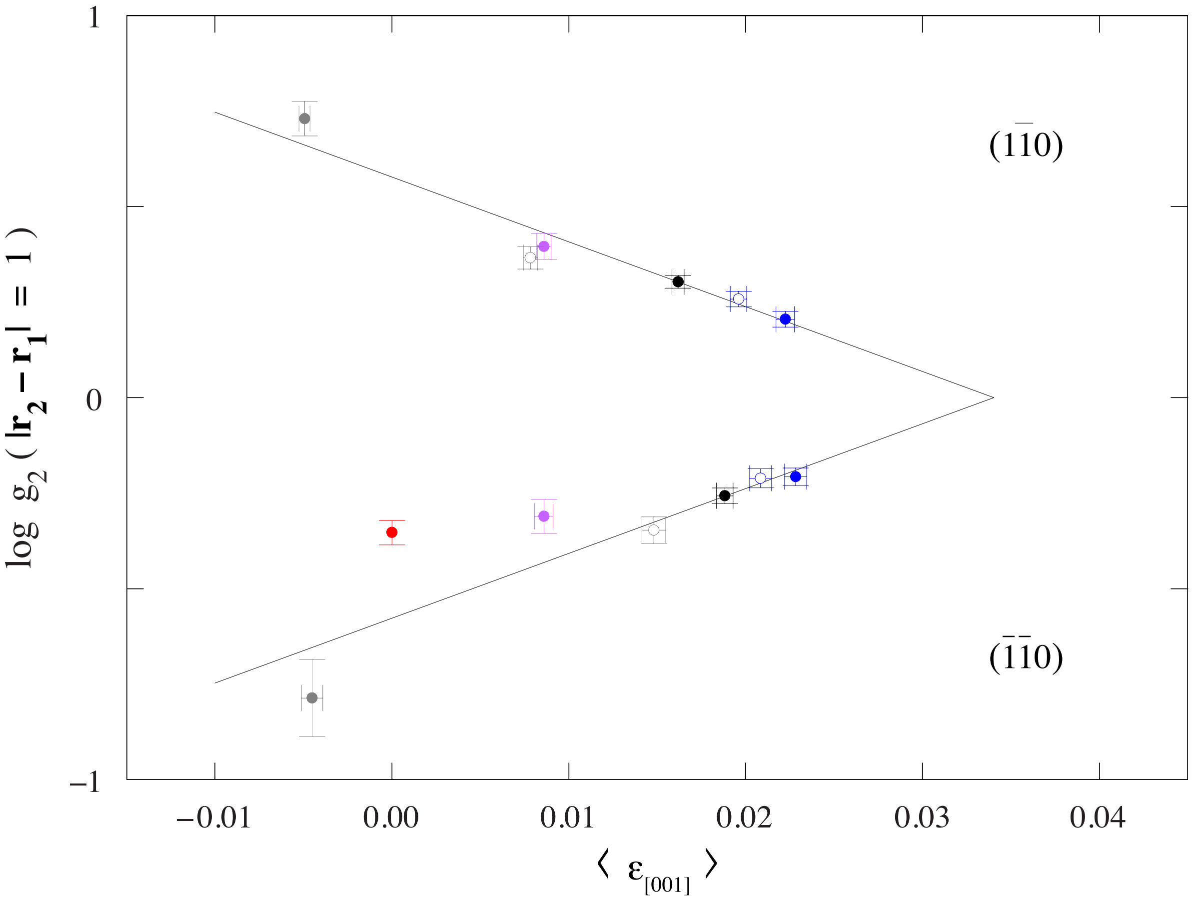


FIGURE 4.42. Logarithm of nearest–neighbor correlations as a function of ensemble strain, and simultaneous fit, reproduced from Fig. 4.37. Nearest–neighbor correlations in the bulk alloy (solid red point) disagree with those in the superlattice. Solid, colored points are strain–resolved, open points are source–resolved, and solid black points are all–inclusive superlattice ensembles.

correlations diverge when expressed in terms of strain. Conversely, as shown in Fig. 4.43, including this data in Fig. 4.41 leaves no doubt bulk and superlattice correlations convincingly agree, provided they are described by their respective bulk antimony fractions[12]. We call attention, however, to the unfortunate absence of (1–10) data for the bulk alloy film from Figs. 4.42 and 4.43, but the previously–drawn parallel between bulk and superlattice (–1–10) correlation functions (Fig. 4.25) suggests the controlling variable here is, again, the number of [110]–oriented dimers at the growth surface[13].

The similarities between our bulk alloy film and superlattice do not end with Fig. 4.25. Exponential fits like those in Fig. 4.25 for the all–inclusive ensembles were repeated for the remaining five superlattice ensembles in Fig 4.43, and the results summarized in Figs. 4.44 and 4.45. The (–1–10) decay length corresponding to the upper branch of each correlation function is found to be independent of antimony fraction, mirroring the composition–independent compensation established in Fig. 4.36[14], whereas the resulting decay–envelope amplitudes are linear in composition. Finally, Fig. 4.46 makes the case that correlation functions for the (continuously antimony exposed) bulk alloy film and (shuttered) source–off superlattice ensemble agree with one another point–by–point, underscoring the irrelevance of an incoming vapor stream, as well as the importance of mean bulk and surface antimony fractions, in scaling these correlations.

---

[12] The antimony fraction was established as 19.5 ± 0.5 % by STM, close to the nominally–targeted alloy composition of 20%.

[13] It's important to point out that the superlattice and alloy film were grown at nearly identical temperatures [23,70], making the presumption of similar prevailing surface reconstructions reasonable.

[14] Recall that strain and composition are interchangeable in any coherently–strained system.

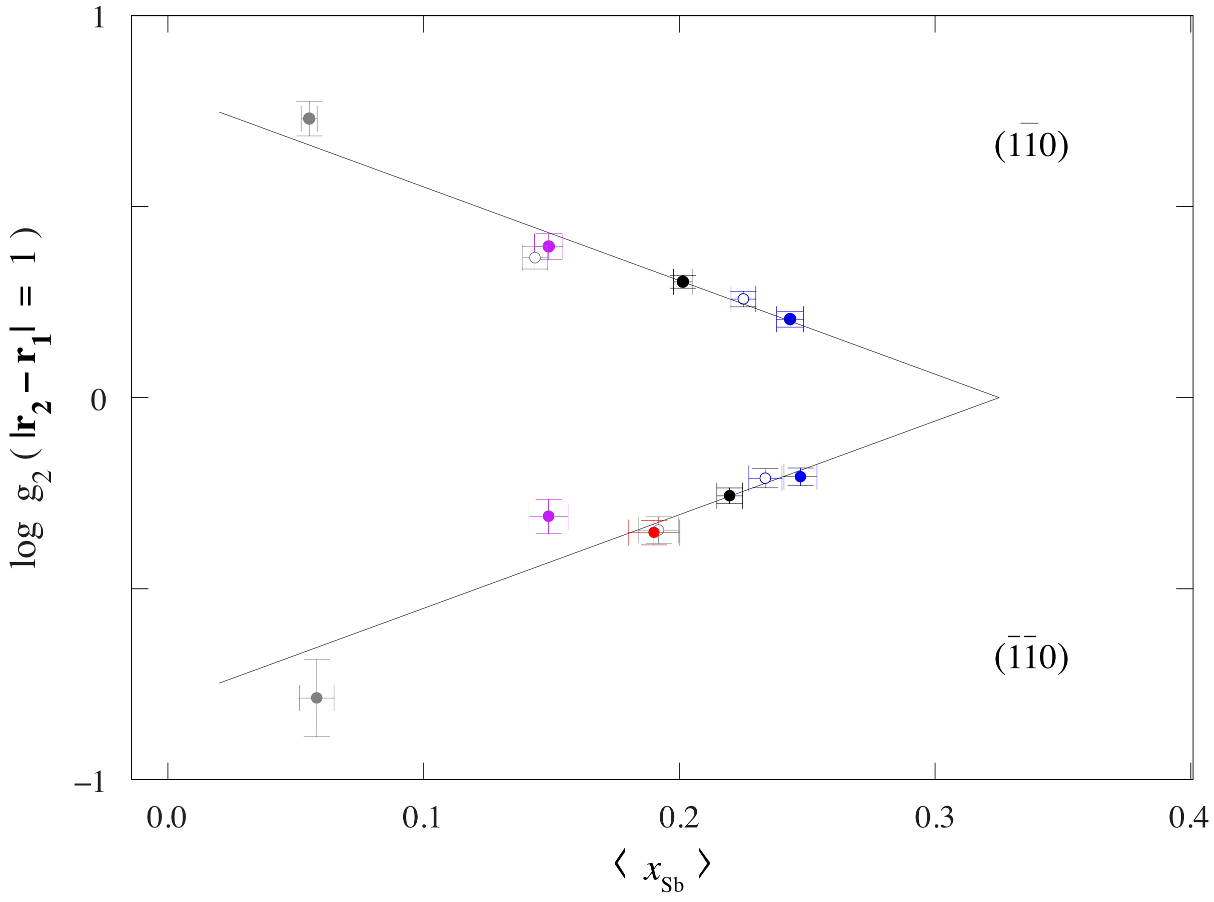


FIGURE 4.43. Logarithm of nearest–neighbor correlations from Fig. 4.42, plotted against antimony fraction in corresponding ensembles. Nearest–neighbor correlations in the bulk alloy (solid red point) now agree with those in the superlattice. Solid, colored points are strain–resolved, open points are source–resolved, and solid black points are all–inclusive superlattice ensembles.

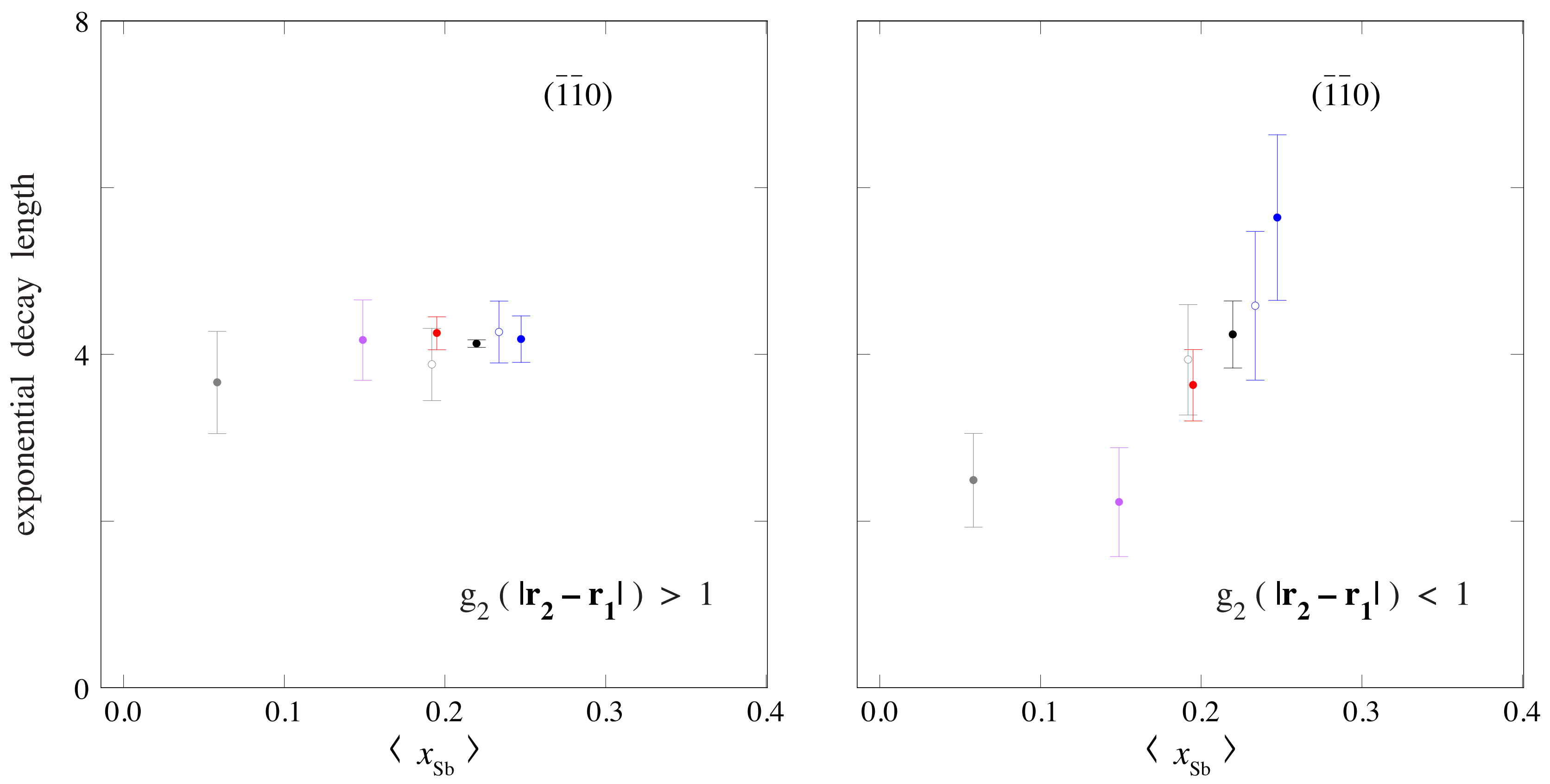


FIGURE 4.44. Exponential decay lengths, in units of the [1–10] lattice spacing (Fig. 4.25), contrasting upper and lower branches (left and right respectively) of bulk and superlattice (–1–10) correlation functions. Upper–branch fall off is the same in all instances, independent of antimony fraction. Symbols are those in Fig. 4.43.

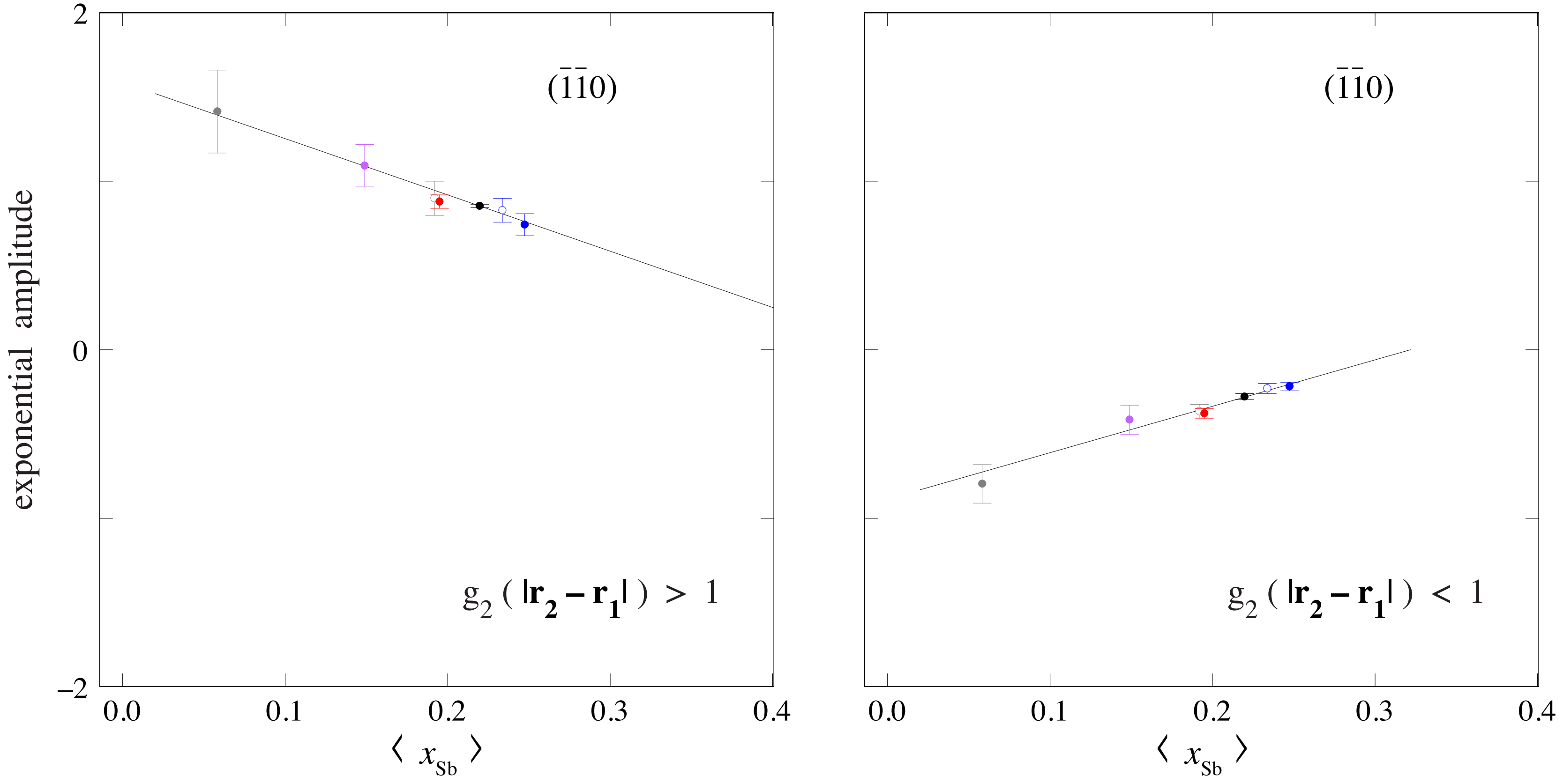


FIGURE 4.45. Amplitudes accompanying the exponential fall offs in Fig. 4.44. Each branch exhibits a linear dependence on antimony fraction. Symbols are those in Fig. 4.43.

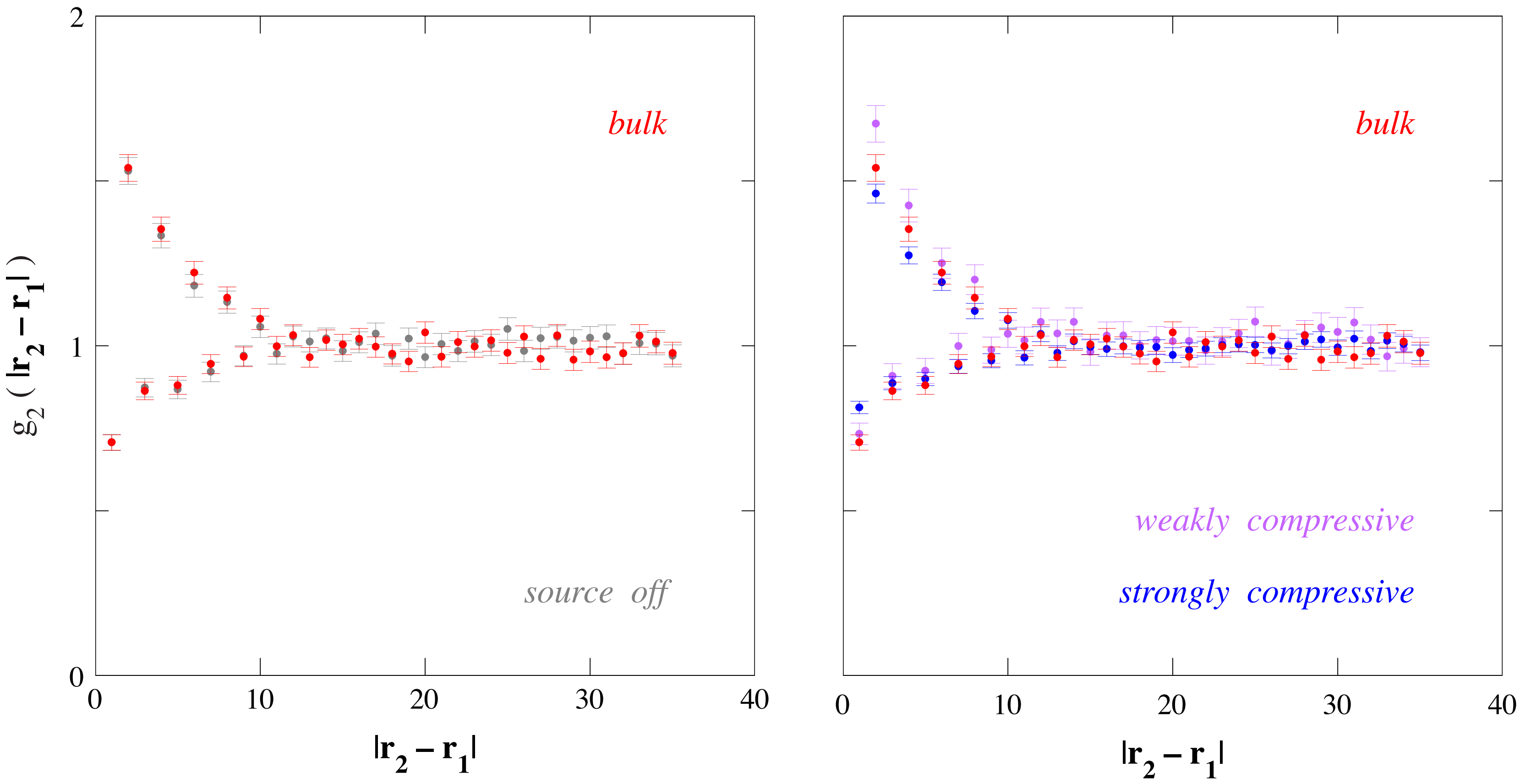


FIGURE 4.46. Point–by–point comparison of bulk and superlattice (–1–10) pair correlation functions. Source–off ensemble most closely resembles the bulk. These two share nearly–identical weighted antimony fractions, $\langle x_{Sb} \rangle$.

Summary

We first showed qualitative evidence in images, thresholded–image autocorrelations, and power spectral densities for atomic order in the InAs / InAsSb superlattice, which is statistically significant on the scale of a micron–long survey. We then described an algorithm that could convert from distorted *x*– and *y*–coordinates into distortion–free [001] columns and <110> rows via lookup images created from filtered DFTs. These columns and rows are central to being able to calculate the separation vector between pairs of antimony atoms in the surface. The pair correlation function was then introduced to calculate the probability above or below random for an arrangement of atoms. We first applied the correlation function to the bulk alloy before particularizing to the superlattice. In the superlattice and the bulk alloy both "attractive" and "repulsive" correlations on the (–1–10) surface were described well by an exponential with fall–off of 4 lattice sites, and the (1–10) surface in the superlattice was nearly random after 1 lattice site implying that antimony is incorporating as dimers aligned with the [110] direction (visible on the (1–10) surface). By splitting the all–inclusive superlattice ensemble into time–resolved sub–ensembles, we observed that the order must have formed on the surface, not in the incoming vapor stream. Comparing strain–resolved ensembles to the bulk ensemble we found the number of [110]–oriented dimers at the growth surface appears to be the controlling variable for short–range alloy order. Finally we found almost perfect agreement between (–1–10) correlations in the bulk alloy and in the source–off ensemble, underscoring the irrelevance of an incoming vapor stream, as well as the importance of mean bulk and surface antimony fractions, in scaling these correlations.

CHAPTER V

CONCLUSIONS

We have used cross–sectional scanning tunneling microscopy (STM) to examine how an *as–grown* InAs/InAsSb superlattice differs from the intended one as regards translational invariance in (001) planes perpendicular to the growth direction. This required atomic–resolution, lateral surveys paralleling the buffer/epilayer interface for up to a micron in orthogonal (–1–10) and (1–10) cross sections, together with repeated lateral surveys at representative vertical locations (i.e., spanned superlattice repeats) within the multilayer stack.

We have shown STM can be used to accurately map the period fluctuations throughout this superlattice. The concept, analogous to Bragg's law in high–resolution x–ray diffraction, relied on an analysis of the [001]–convolved reciprocal–space satellite peaks obtained from discrete Fourier transforms of individual STM images. Properly implemented, the technique enabled local period measurements that reliably discriminated lateral fluctuations localized to within ~ 40 nm along <110> directions in the growth plane. While not as accurate as x–ray, the inherent, single–image measurement error associated with the method may be made as small as 0.1%, which allowed lateral period fluctuations potentially contributing to inhomogeneous energy broadening and carrier localization in these structures to be pinpointed and quantified. The direct visualization of such unexpectedly–large fluctuations on nanometer length scales was tied to a stochastic description of correlated interface roughness.

We also introduced a new technique to automatically tabulate the crystalline coordinates of previously–identified top–layer antimony atoms and construct the antimony pair–correlation functions in orthogonal cross sections. These correlation functions were then analyzed in terms of layer strain as well as antimony fraction, and comparisons drawn between corresponding superlattice and bulk alloy experiments. Nearest–neighbor correlations on opposing cleavage faces were inversely related, with the (–1–10) deficit at nearest–neighbor sites balanced by a compensating (1–10) surplus. The logarithm of this preference scaled inversely with bulk antimony fraction. In more vivid physical terms, this preferential [110]–incorporation of nearest–neighbor antimony atoms in the bulk was traced to the inferred concentration of [110]–oriented antimony dimers at the growth surface.

REFERENCES

[1] R. Caballero, *Physics of the Atmosphere* (IOP Publishing, Philadelphia, 2014).

[2] F. K. Tittel, D. Richter, and A. Fried, in *Solid-State Mid-Infrared Laser Sources*, edited by I. T. Sorokina, and K. L. Vodopyanov (Springer, Berlin, 2003), pp. 458.

[3] C. Gmachl, F. Capasso, D. L. Sivco, and A. Y. Cho, Reports on Progress in Physics **64**, 1533 (2001).

[4] S. Liakat, K. A. Bors, T.-Y. Huang, A. P. M. Michel, E. Zanghi, and C. F. Gmachl, Biomed. Opt. Express **4**, 1083 (2013).

[5] L. Menzel *et al.*, Applied Physics B **72**, 859 (2001).

[6] A. Rogalski, Reports on Progress in Physics **68**, 2267 (2005).

[7] E. F. Schubert, *Light-Emitting Diodes* (Cambridge University Press, Cambridge, UK, 2003).

[8] D. L. Smith and C. Mailhiot, Journal of Applied Physics **62**, 2545 (1987).

[9] G. A. Sai‐Halasz, R. Tsu, and L. Esaki, Applied Physics Letters **30**, 651 (1977).

[10] Y. H. Zhang, Applied Physics Letters **66**, 118 (1995).

[11] B. V. Olson, E. A. Shaner, J. K. Kim, J. F. Klem, S. D. Hawkins, L. M. Murray, J. P. Prineas, M. E. Flatté, and T. F. Boggess, Applied Physics Letters **101**, 092109 (2012).

[12] E. H. Steenbergen *et al.*, Applied Physics Letters **99**, 251110 (2011).

[13] B. C. Connelly, G. D. Metcalfe, H. Shen, and M. Wraback, Applied Physics Letters **97**, 251117 (2010).

[14] D. Donetsky, S. P. Svensson, L. E. Vorobjev, and G. Belenky, Applied Physics Letters **95**, 212104 (2009).

[15] E. H. Aifer *et al.*, Applied Physics Letters **89**, 053519 (2006).

[16] V. Ottesen, (Wikimedia Commons, 2015).

[17] M. H. Hablanian, *High-Vacuum Technology: A Practical Guide* (Marcel Dekker Inc., New York, 1997), 2nd edn.

[18] J. Y. Tsao, *Materials Fundamentals of Molecular Beam Epitaxy* (Academic Press, Cambridge, MA, 1993).

[19] J. van Wingerden, Y. A. Wiechers, P. M. L. O. Scholte, and F. Tuinstra, Surface Science **331**, 473 (1995).

[20] K. M. Kanedy, Ph. D. Thesis, Texas A&M University, College Station, 2016.

[21] J. Harper, M. Weimer, D. Zhang, C. H. Lin, and S. S. Pei, Applied Physics Letters **73**, 2805 (1998).

[22] J. Steinshnider, J. Harper, M. Weimer, C. H. Lin, S. S. Pei, and D. H. Chow, Physical Review Letters **85**, 4562 (2000).

[23] M. R. Wood, K. Kanedy, F. Lopez, M. Weimer, J. F. Klem, S. D. Hawkins, E. A. Shaner, and J. K. Kim, Journal of Crystal Growth **425**, 110 (2015).

[24] A. G. Norman, in *Spontaneous Ordering in Semiconductor Alloys*, edited by A. Mascarenhas (Springer, Boston, MA, 2002), pp. 45.

[25] H. R. Jen, K. Y. Ma, and G. B. Stringfellow, Applied Physics Letters **54**, 1154 (1989).

[26] S. L. Zuo, E. T. Yu, A. A. Allerman, and R. M. Biefeld, Journal of Vacuum Science & Technology B **17**, 1781 (1999).

[27] S.-H. Wei, in *Spontaneous Ordering in Semiconductor Alloys*, edited by A. Mascarenhas (Springer, Boston, MA, 2002), pp. 423.

[28] W. L. Sarney, S. P. Svensson, Y. Lin, D. Donetsky, L. Shterengas, G. Kipshidze, and G. Belenky, Journal of Applied Physics **119**, 215704 (2016).

[29] Z.-Y. Lin, S. Liu, E. H. Steenbergen, and Y.-H. Zhang, Applied Physics Letters **107**, 201107 (2015).

[30] C. A. Hoffman, J. R. Meyer, E. R. Youngdale, F. J. Bartoli, R. H. Miles, and L. R. Ram-Mohan, Solid-State Electronics **37**, 1203 (1994).

[31] S. P. Svensson, W. L. Sarney, K. M. Yu, M. Ting, W. L. Calley, S. V. Novikov, C. T. Foxon, and W. Walukiewicz, Journal of Crystal Growth **425**, 2 (2015).

[32] MBE - Molecular-Beam Epitaxy - k-Space Applications., https://www.k-space.com/applications/mbe/ (Accessed October 23, 2017).

[33] P. F. Fewster, *X-ray scattering from semiconductors* (Imperial College Press, London, 2003), 2nd edn.

[34] K. Matney and M. S. Goorsky, Journal of Crystal Growth **148**, 327 (1995).

[35] E. Luna, F. Ishikawa, B. Satpati, J. B. Rodriguez, E. Tournié, and A. Trampert, Journal of Crystal Growth **311**, 1739 (2009).

[36] D. J. Smith, Micron **43**, 504 (2012).

[37] H. Cerva, Solid-State Electronics **37**, 1045 (1994).

[38] P. M. Petroff, A. C. Gossard, W. Wiegmann, and A. Savage, Journal of Crystal Growth **44**, 5 (1978).

[39] F. Lopez, Ph. D. Thesis, Texas A&M University, College Station, 2015.

[40] C. J. Chen, Physical Review Letters **65**, 448 (1990).

[41] C. J. Chen, *Introduction to Scanning Tunneling Microscopy* (Oxford University Press, New York, 1993), Oxford series in optical and imaging sciences: 4.

[42] J. A. Stroscio and W. J. Kaiser, *Scanning tunneling microscopy* (Academic Press, Cambridge, MA, 1993), Methods of experimental physics: v. 27.

[43] R. M. Feenstra, J. A. Stroscio, J. Tersoff, and A. P. Fein, Physical Review Letters **58**, 1192 (1987).

[44] S. Adachi, P. Capper, S. Kasap, and A. Willoughby, *Properties of Semiconductor Alloys: Group-IV, III-V and II-VI Semiconductors* (Wiley, West Sussex, 2009).

[45] C. B. Duke, Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films **10**, 2032 (1992).

[46] J. H. Davies, P. Offermans, and P. M. Koenraad, Journal of Applied Physics **98**, 053504 (2005).

[47] J. Steinshnider, M. Weimer, R. Kaspi, and G. W. Turner, Physical Review Letters **85**, 2953 (2000).

[48] G. W. Brown, Ph. D. Thesis, Texas A&M University, College Station, 1997.

[49] G. A. Lengel, Ph. D. Thesis, Texas A&M University, College Station, 1995.

[50] J. F. O'Hanlon, *A User's Guide to Vacuum Technology* (John Wiley & Sons, Hoboken, NJ, 2003), 3rd edn.

[51] J. Hansknecht, P. Adderley, M. L. Stutzman, and M. Poelker, AIP Conference Proceedings **1149**, 1143 (2009).

[52] D. R. Lide, *CRC Handbook of Chemistry and Physics* (CRC Press, Cleveland, Ohio, 1990), 71st edn.

[53] D. W. Pohl, Review of Scientific Instruments **58**, 54 (1987).

[54] J. F. Jorgensen, K. Carneiro, L. L. Madsen, and K. Conradsen, Journal of Vacuum Science & Technology B **12**, 1702 (1994).

[55] W. L. Briggs and V. E. Henson, *The DFT: An Owners' Manual for the Discrete Fourier Transform* (Society for Industrial and Applied Mathematics, Philadelphia, 1995).

[56] R. C. Gonzalez and R. E. Woods, *Digital Image Processing* (Prentice Hall, Upper Saddle River, N.J, 2008), 3rd edn.

[57] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes: The Art of Scientific Computing* (Cambridge University Press, Cambridge, UK, 2007), 3rd edn.

[58] C. Kittel, *Introduction to Solid State Physics* (Wiley, New York, 1971), 4th edn.

[59] K. Muraki, S. Fukatsu, Y. Shiraki, and R. Ito, Applied Physics Letters **61**, 557 (1992).

[60] J. R. Taylor, *Introduction To Error Analysis: The Study of Uncertainties in Physical Measurements* (University Science Books, Sausalito, CA, 1997), 2nd edn.

[61] P. R. Bevington and D. K. Robinson, *Data reduction and error analysis for the physical sciences* (McGraw-Hill, Boston, 2003), 3rd edn.

[62] R. L. Ott and M. T. Longnecker, *An Introduction to Statistical Methods and Data Analysis* (Brooks/Cole Cengage Learning, Belmont, CA, 2010), 6th edn.

[63] C. Weisbuch, R. Dingle, A. C. Gossard, and W. Wiegmann, Solid State Communications **38**, 709 (1981).

[64] A. Chomette, B. Deveaud, A. Regreny, and G. Bastard, Physical Review Letters **57**, 1464 (1986).

[65] J. B. Khurgin, Applied Physics Letters **93**, 091104 (2008).

[66] M. B. Priestley, *Spectral Analysis and Time Series: Univariate Series* (Academic Press, Cambridge, MA, 1981), Vol. 1.

[67] G. E. P. Box, *Time Series Analysis forecasting and control* (Holden–Day, San Francisco, 1976).

[68] S. P. Svensson, W. L. Sarney, H. Hier, Y. Lin, D. Wang, D. Donetsky, L. Shterengas, G. Kipshidze, and G. Belenky, Physical Review B **86**, 245205 (2012).

[69] T. L. Hill, *Statistical Mechanics: Principles and Selected Applications* (McGraw-Hill, New York, 1956).

[70] W. L. Sarney, S. P. Svensson, H. Hier, G. Kipshidze, D. Donetsky, D. Wang, L. Shterengas, and G. Belenky, Journal of Vacuum Science & Technology B **30**, 02B105 (2012).