Multiscale Investigation of Shape Memory Alloy Fatigue

by

William Scott LePage

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Mechanical Engineering)
in The University of Michigan
2018

Doctoral Committee:

Associate Professor Samantha Daly, Co-Chair,
University of California, Santa Barbara
Assistant Professor Neil Dasgupta, Co-Chair
Associate Professor Vikram Gavini
Dr. Alan Pelton, G-RAU, Inc.
Professor John Shaw

William LePage

wlepage@umich.edu

ORCID iD: 0000-0002-6811-3162

DEDICATION

To my family.

ACKNOWLEDGEMENTS

This research was supported by the National Science Foundation (CAREER award 1251891), U.S. Department of Energy (contract DE-SC0003996), the U.S. Department of Defense (NDSEG, 32 CFR 168a), Tau Beta Pi (Anderson Fellowship), and the University of Michigan Rackham Graduate School (Rackham Conference Travel Grants).

I am tremendously grateful to everyone who supported me during my graduate studies. Thank you especially to my advisors, Samantha Daly and John Shaw, for your unwaivering support, thoughtful guidance, and caring patience. I feel so fortunate to have you both as my advisors. As a colleague told me, "You couldn't have picked two better advisors. You really are studying with the best." I completely agree.

Thank you to the many students and post-doctoral researchers from whom I learned so much, especially to Ryan Watkins for coaching me to be a careful experimentalist, to Zhe Chen, Adam Kammers, and Will Lenthe for helping bring my work to the microscale with their new techniques in SEM-DIC, and to Jalil Alidoos, Daniel Biggs, Corin Bowen, Jonathan Estrada, Jason Geathers, Joyce Gong, Michelle Harr, Michael Kimiecik, Marissa Linne, Callan Luetkemeyer, Benjamin Marchi, Benjamin Reedlunn, Avery Samuel, J.C. Stinville, and Jared Tracy for all of your help and thought-provoking conversations.

Thank you to Andy Poli for his continually kind reception of my ideas in their most unpolished state, for assisting me with equipment and experiments, and for sharing his refreshing thoughts.

Thank you to Aslan Ahadi for teaching me new concepts about NiTi, for coaching me in X-ray diffraction of NiTi, and for helping me broaden my work.

Thank you to all of the faculty and staff who shared their suggestions, critiques, equipment, and expertise, including Ellen Arruda, Neil Dasgupta, Allen Hunter, Nakhiah Goulbourne, Bobby Kerns, J. Wayne Jones, Jerry Li, Ying Qi, Michael Thouless, and Chris Torbet.

Thank you to my undergraduate research mentors, John Henshaw and Gordon Purser, for helping me find my passion in science, and for encouraging and aiding me to attend graduate school. Also, thank you to my mentor from Sandia National Laboratories, Aaron Hall, for many formative conversations about excelling in graduate school and beyond.

Thank you to my brother Daniel for encouraging me to pursue undergraduate research opportunities and inspiring me to continue on to graduate school.

Thank you immensely to my parents Bill and Lesley for fostering curiosity and making so many sacrifices for my education and growth.

Finally, I am most grateful to my wonderful wife Becca and our son Peter, for your steadfast love, for giving me countless reasons to smile, and for showing me what is most important in life. Thank you for making my life abundantly joyful.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

ABSTRACT

Shape memory alloys (SMAs) are functional materials with two remarkable properties: superelasticity and the shape memory effect. The reversible, solid-to-solid ("martensitic") phase transformations that enable these useful material responses also complicate their durability predictions. SMAs are used extensively in biomedical devices such as stents and root canal files, and also have promising applications for important frontiers, such as weight-saving actuators, active structures for aerospace systems, and damping components for civil structures. Before SMAs can be widely adopted for these new applications, however, their response to cracking and failure must be understood. Complexities from martensitic phase transformations in SMAs confound many existing predictions of damage, fatigue, and fracture mechanics (including Schmid's Law [1] and linear/elastoplastic fracture mechanics [2]). Furthermore, this work's findings could extend to other emerging materials with phase transformations, such as Heusler alloys for high-efficiency magnetocaloric refrigeration.

To address these knowledge gaps, this work provides new observations and insights on the role of phase transformation during cracking and failure of SMAs. Additionally, new experimental frameworks were developed to characterize cracks in structural materials for unprecedented breadth and precision between the millimeter and nanometer length scales.

The early part of this work optimized full-field deformation measurements to enable the high-precision experiments of the later part of this work. Also, it advanced techniques in the broader experimental mechanics community. First, a new method was developed to enhance measurements from digital image correlation (DIC), a powerful technique for measuring material deformation [3]. Also, clear guidelines were presented for optimizing an important sample preparation step for optical DIC ("speckle patterning" with paint) [4].

The latter part of this work established new insights into SMA cracking by examining the effects of grain size, crystallographic texture, and temperature. There are two types of fatigue in SMAs: functional fatigue (degradation of the phase transformation, especially from reduction in the recovered strain during superelastic cycling), and structural fatigue (the typical fatigue response of metals, with crack initiation and growth). Despite the dramatic enhancement of NiTi's functional fatigue resistance with grain size reduction to the nanoscale, there was no improvement in structural fatigue resistance with respect to grain size [5]. Rather, the largest grain size studied had the slowest crack growth rate, which

was attributed to roughness-induced crack closure (not observed in the other grain sizes). For all grain sizes, the macroscopic fatigue crack growth correlated well with microscopic crack tip observations: the grain sizes with relatively fast macroscopic crack growth rates exhibited large crack displacements at the microscale, and vice versa.

Next, the effects of crystallographic texture on NiTi's functional and structural fatigue were characterized. Unlike the grain size study that was complicated by roughness-induced crack closure differences, the three texture conditions had similar fracture surfaces. Furthermore, there were clear connections between functional and structural fatigue. During cyclic tension, there was about three times as much residual strain accumulation in the sheet's rolling direction (RD) compared to the transverse direction (TD), and the crack growth rates of the RD condition were consistently faster than the 45 and TD conditions.

Finally, fatigue experiments as a function of temperature characterized the role of martensitic phase transformation on cracking. The slowest crack growth rates were measured in the stable martensite and thermal R-phase martensite.

CHAPTER 1

Introduction

1.1 Overview

This dissertation is organized into eight chapters. The first chapter provides an introduction to NiTi shape memory alloy, this work's material of interest, (Section 1.2) and an introduction to digital image correlation, a powerful technique that was used and advanced in this work (Section 1.3).

Chapters 2–4 focus on key experimental techniques that were developed in this work. Chapters 2 and 3 present optimizations for digital image correlation. These two chapters have been published in Refs. [3, 4], and their optimizations were adapted in the experiments throughout Chapters 5–7. The third chapter on experimental methods (Chapter 4) presents experimental measurements of residual stress and crystallographic texture for NiTi tubes. These measurements complement the extensive work from Dr. Benjamin Reedlunn on the multiaxial thermo-mechanical behavior of NiTi tubes [6], the most important product form of the most important shape memory alloy to date.

Chapters 5–7 investigate the fatigue and fracture behavior of NiTi SMA through three parameters that have strong influence on NiTi's mechanical response: grain size (Chapter 5), texture (Chapter 6), and temperature (Chapter 7). The first of these chapters has been published in Ref. [5]. As discussed further in Section 1.2.4, these three parameters can complicate the mechanical response of NiTi and present challenges to the measurement and modeling of NiTi's thermo-mechanical response. However, these three parameters can also be used to extract scientific knowledge about the role of phase transformation in the presence of cracks.

Finally, conclusions and future work are summarized in Chapter 8.

1.2 Introduction to NiTi shape memory alloy

Shape memory alloys (SMAs) exhibit two responses that most metals do not: *superelasticity* and the *shape memory effect*. Both of these responses are driven by martensitic phase transformations. These solid-to-solid transformations involve shifting the material's atomic configurations without diffusion, as described further in Section 1.2.2. Superelasticity and the shape memory effect provide unusually large amounts of shape recovery. With these unique abilities, SMAs have been deployed in biomedical devices, aerospace and automotive systems, civil structures, and beyond.

Many SMAs have been discovered, and the specific phases and transformations of SMAs can vary. Nonetheless, SMAs manifest common features in their thermo-mechanical responses, in addition to superelasticity and the shape memory effect. These common features can include latent heat effects, acute thermomechanical coupling, and localization/propagation of phase transformations (both within macroscopic SMA structures [7] and on the microstructural level [8]). These effects can be useful or problematic, depending on the application. For example, latent heating complicates the thermo-mechanical response of SMAs, but also enables using SMAs as elastocaloric devices to replace vapor compression refrigeration [9]. In general, the unique thermo-mechanical responses of SMAs enable novel applications that conventional metals could not fulfill, but the scientific understanding, measurement, analysis, and simulation of SMAs is also complicated by their unusual features. As described in Section 1.2.3, the fatigue and fracture mechanics of SMAs are further complicated by the unusual responses of SMAs.

NiTi is the most developed, studied, and utilized SMA. The martensitic phase transformation of NiTi was discovered by accident in 1959 by William Buehler at the Naval Ordinance Laboratory. The common name for NiTi, "Nitinol," was coined by combining "Ni" and "Ti" with the first letters of "Naval Ordinance Laboratory." Buehler, a native of Detroit, Michigan, was surveying intermetallics for the high-temperature environment of nose cones for spacecraft reentry vehicles. One of the intermetallics that Buehler produced was NiTi with approximately equal proportions of Ni and Ti atoms. While handling arc-cast bars of new alloys, Buehler would drop the bars as a crude measure of the alloy's damping capacity. Buehler discovered NiTi's phase transformation when he dropped bars of different temperatures: cooler bars made a thud sound, like Pb, while warmer bars rang like steel. Buehler observed that the previously cool bars, when heated a bit, would produce the ringing sound, and upon cooling, would return to making a thud [10].

1.2.1 Applications of NiTi shape memory alloy

Figure 1.1 is a timeline of select commercial applications of NiTi SMA. The first commercial application of NiTi was a hydraulic pipe coupler for the F14 fighter jet, introduced in 1969 [11]. The majority of NiTi applications to date have been biomedical devices, beginning with vena cava filters and root canal files in the late 1970s [12], and continuing with bone compression staples and stents in the 1980s [11]. Today, NiTi's use in health care remains its most common and important technological application. For example, there are about 22 million endodontic (root canal) procedures annually in the United States [13], and about 1 million annual implantations of endovascular cardiac stents during hospital stays [14]. Furthermore, stents represent more than 95 % of all endovascular interventions [15]. Although other materials are used in some cases, the majority of these endodontic files and endovascular devices use NiTi SMA.

Although the first commercial NiTi product was an aerospace component, the majority of NiTi product developments in the 1970s and 1980s were biomedical devices. However, beginning in 1989 with the first NiTi product that leveraged the shape memory effect for linear actuation, there have been a number of NiTi applications in aerospace systems [11]. For example, the solar panels on the Hubble Space Telescope were deployed with NiTi using the shape memory effect. Also, NiTi products have been developed for civil structures for a number of uses, such as earthquake resistance or concrete pre-compression. While most NiTi biomedical devices are produced from tubes or wires, these NiTi products for civil applications are typically cables with bundled wires. [1] Recently, a team at NASA Glenn Research Center developed a superelastic NiTi wheel for Mars rovers to address the fatigue and failure issues of existing planetary rover wheels [20].

[1] For more details about the intricate thermo-mechanical responses of NiTi cables, see Refs. [16, 17, 18] and the theses by B. Reedlunn [6] D. Biggs [19].

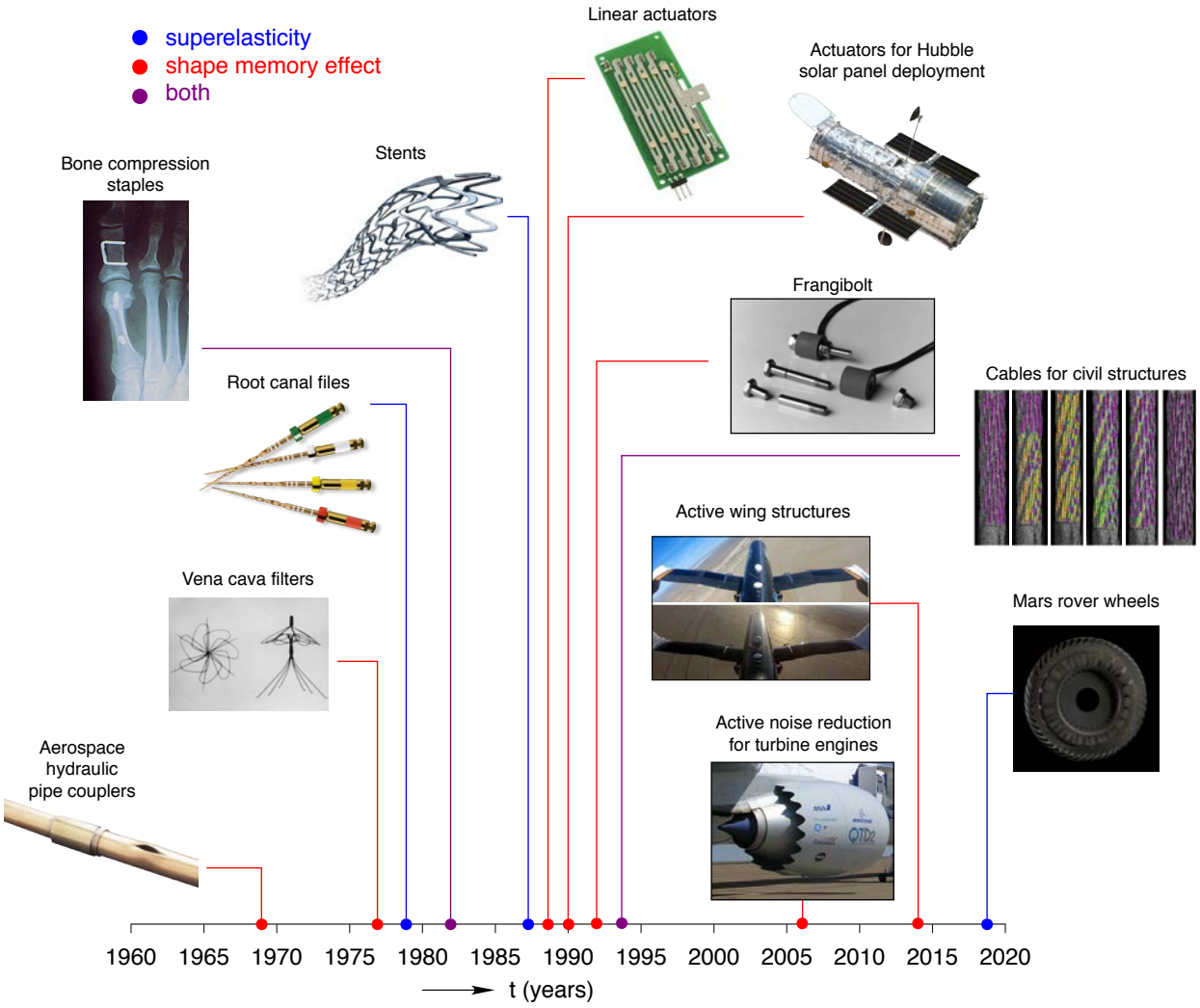


Figure 1.1: Beginning in 1969 with a hydraulic pipe coupler for the F14 fighter, NiTi SMA has been used in many applications for systems and devices in health care, aerospace, and beyond. The lines that connect each application to the timeline indicate the SMA response that is used in that application: blue lines indicate superelasticity, red lines indicate the shape memory effect, and purple lines indicate a mix of both superelasticity and the shape memory effect. Note that superelasticity is shown for stent applications, but early implementations of NiTi stents used the shape memory effect. However, due to superior fatigue performance, stents and most other NiTi biomedical devices utilize superelasticity [21]. Also, some of these technologies use ternary alloys of NiTi, such as NiTiHf in many aerospace applications. The image sources, left to right from the left-most edge of each image: [22], [23], [12], [24], [25] [26], [27], [28], [29], [30], [16], and [20]. All dates prior to 2015 are from Ref. [11] with the application's introduction; the date after 2015 is from Ref. [20].

For the majority of NiTi applications, understanding the material's deformation, damage, and failure is paramount. For biomedical devices with small characteristic length scales, short cracks are important to avoid because length for crack propagation is very short. Thus,

there is a need to understand and mitigate how plasticity accumulates and eventually leads to fatigue crack *initiation* and device failure. Unfortunately, failures have been relatively common for NiTi devices, especially for challenging conditions such as the popliteal artery at the back of the knee (Figure 1.2). As outlined in 2015 in Ref. [15], the prevalence of fractures in stent struts ranges from 2.6 % of stents with severe, "transectional" disruption on blood flow when examined with angiography [31], and up to 29 % of stents with any kind of fractures found at autopsy [32, 33]. As industry experts will emphasize, fatigue remains a major challenge for NiTi biomedical devices. In the words of Dr. Alan Pelton at the 2017 International Conference on Martensitic Transformations, "The three most important things about Nitinol are fatigue, fatigue, and fatigue."

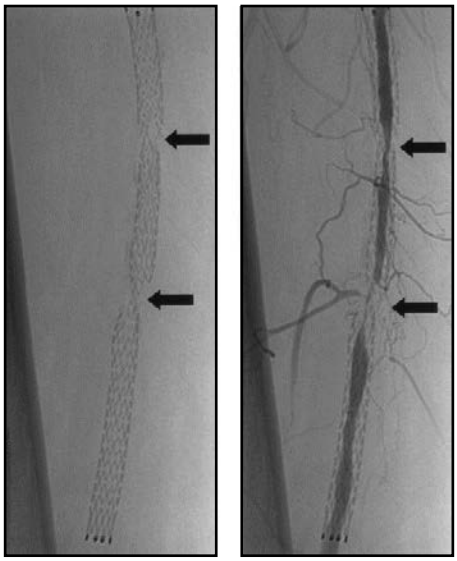


Figure 1.2: A femoropopliteal stent with two strut fractures, shown in the left image in an X-ray radiograph, inhibits blood flow, seen in the right image in an X-ray angiograph. Image source: Ref. [34].

Along with plasticity and fatigue crack *initiation* for NiTi devices with small characteristic length scales, fatigue crack *propagation* and fracture are important for NiTi devices with large characteristic length scales. Larger structures can allow for crack propagation before failure, and in many cases these larges components are still in critical applications, such as aerospace structures. An example of a fracture-relevant NiTi structure in aerospace includes the bulk strips of NiTi that have been used for active noise reduction in turbine engines [35].

In summary, there is both a current and future need to understand the fatigue and fracture response of NiTi, for ensuring the safety of biomedical devices and for enabling future technologies with these materials.

1.2.2 Martensitic phase transformation in NiTi

NiTi's superelasticity and shape memory are driven by transformations between *austenite* (A) and *martensite* (M). A is a cubic phase (B2) with two atoms per unit cell that is stable at higher temperatures. M is a monoclinic phase (B19′) with five atoms per unit cell that is stable at lower temperatures. Below the temperature at which austenite begins to form (A_s), M exhibits *twinning*. In a simplified, two-dimensional view, twinning looks like kinks among alternating rows of atoms. This arrangement in twinned M (also called "thermal-M") is illustrated in Figure 1.3. During deformation, M can undergo *detwinning* as a strain accommodation mechanism. Again in a simplified two-dimensional sense, detwinning is like straightening out the kinks among alternating rows of M to form "oriented-M" (Figure 1.3).

In addition to M, another martensite phase (R-phase, or R) can form in NiTi. R has a rhombohedral B2′ structure with two atoms per unit cell (Figure 1.3). The strain change and stress level for $A \leftrightarrow R$ transformation are relatively small (≈ 0.5 % and < 0.3 GPa) compared with the $A \leftrightarrow M$ transformation (≈ 5 % and ≈ 0.3 to 0.8 GPa) [36, 37]. For this reason, there are very few (if any) applications that actively utilize the R-phase transformation. However, unlike M, R has kinematic compatibility with A, which would enhance the resistance to plasticity for the $A \leftrightarrow R$ transformation (again, however, this transformation is less useful).

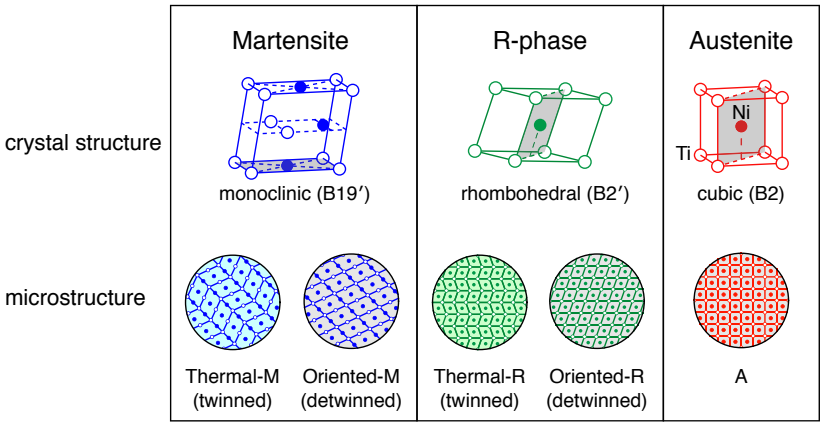


Figure 1.3: The crystal structure unit cells and simplified two-dimensional microstructures of the three phases of NiTi shape memory alloy: B19′ monoclinic martensite, B2′ rhombohedral R-phase, and B2 cubic austenite. Note that martensite has 5 atoms per unit cell, while R-phase and austenite only have two atoms per unit cell. Source: Ref. [36]

Below the austenite start temperature (A_s), the $A \rightarrow M$ phase transformation can be *thermally* induced. In conjunction with applied deformation, the thermally-induced phase transformation can give rise to the shape memory effect, which is illustrated in the low-temperature portion of Figure 1.4. Between ⓪ and ①, mechanical deformation causes

detwinning in the martensite from M to M^+. After relaxing the stress to ② and heating to ③, just below the austenite start temperature, the material remains M^+. However, upon heating through ③ and past ④, the microstructure transforms to A, and the material macroscopically recovers the applied deformation and returns to its original shape. After cooling below A_s, the material returns to the twinned martensite M without changing the macroscopic shape of the material.

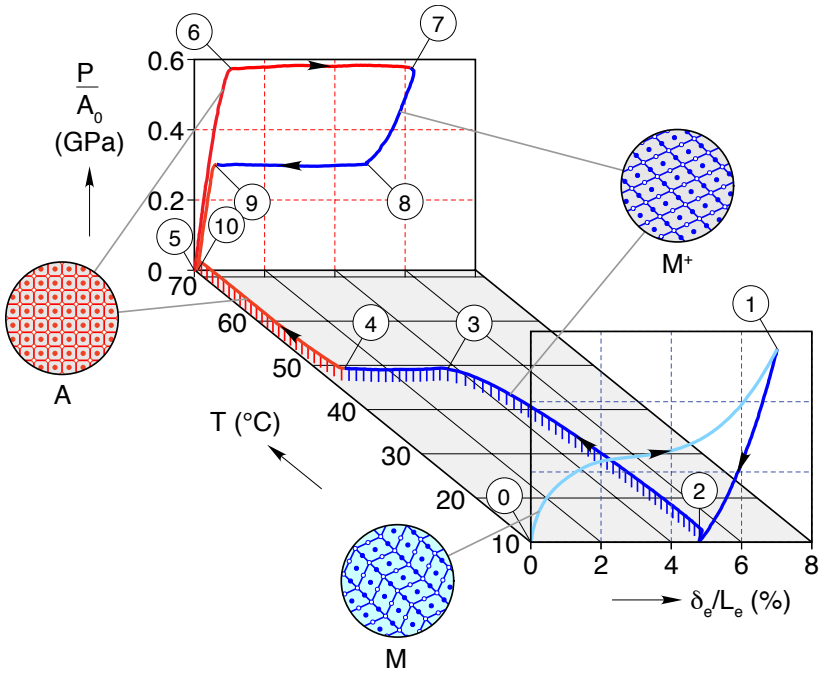


Figure 1.4: From one wire sample in uniaxial tension, both the shape memory effect and superelasticity can be demonstrated. Source: Ref. [36]

A simplified way to envision the mechanism of this shape recovery is the possible configurations of martensite and austenite. Martensite can be detwinned in many different configurations to accommodate deformation in multiple directions (including for more complex deformations than the uniaxial tension example in Figure 1.4), but there is only one configuration for the cubic austenite. This shape recovery mechanism is illustrated in Figure 1.5.

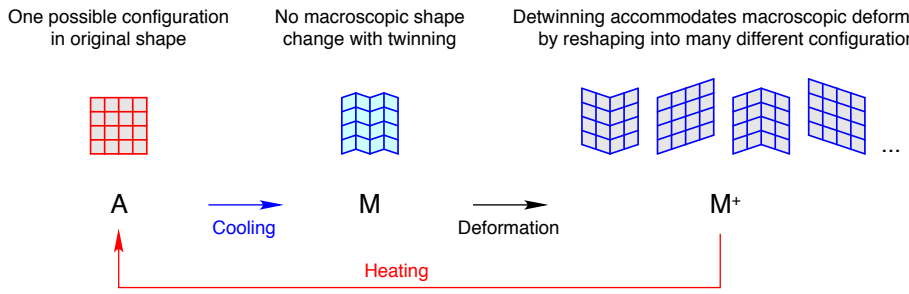


Figure 1.5: The shape memory effect can be envisioned with simple transformations of lines. Note that this diagram oversimplifies the microstructure, since A has two atoms per unit cell, while M has five. In austenite, there is only one possible configuration for the lines to intersect at right angles (analogous to the B2 cubic microstructure of A). Upon cooling to twinned martensite M, there is not a significant macroscopic shape change, because twinning involves relatively small, zigzagged shifts in the atoms. When M is detwinned, however, large macroscopic deformation can be accommodated, and the microstructure can have many possible configurations. Upon heating, however, it must return to the one configuration of A.

Above the austenite finish temperature (A_f), the $A \rightarrow M$ phase transformation can be *mechanically* induced to give rise to superelasticity, which is illustrated in Figure 1.4 with the mechanical response of NiTi uniaxial tension. At ⑤, the material is A. Under elastic deformation to ⑥, the material is macroscopically still austenite (although microscopically phase transformation begins within grains before the macroscopic onset of phase transformation at ⑥ [38]). From ⑥ to ⑦, there is a localization and propagation of $A \rightarrow M^+$ transformation until macroscopically the material is fully M (although on the microscale, there can be regions of untransformed A [8]). Upon unloading from ⑦ to ⑧, the material initially has an elastic recovery of M. From ⑧ to ⑨, there is reverse phase transformation, $M^+ \rightarrow A$. Similar to the forward transformation, the reverse transformation has a structural propagation, although at a lower stress level. Finally, with stress fully relaxed at ⑩, the deformation is mostly recovered, with only small levels of residual martensite and plasticity.

1.2.3 Functional and structural fatigue and fracture in NiTi

Shape memory alloys (SMAs) have promising applications in new frontiers, but fundamental questions remain about their durability because martensitic phase transformations add significant complexity. As outlined in Ref. [39], the complexities for thermo-mechanical behavior of NiTi include:

1. nonlinear, non-monotonic stress-strain response with significant hysteresis and path dependence;

2. acute thermo-mechanical coupling from latent heat effects and enthalpy changes;

3. anisotropy due to microstructural texture and the inherent anisotropy of B19′ martensite [40];

4. tension-compression asymmetry;

5. unstable stress-induced transformation that loads to macroscopic localization and propagation of the phase transformation; and

6. simultaneous slip and phase transformation [38].

With repeated mechanical and/or thermal cycling, shape memory alloys exhibit two types of fatigue: structural fatigue and functional fatigue. *Structural* fatigue is the accumulation of damage that leads to crack initiation and growth. *Functional* fatigue is the decay of properties related to the phase transformation. Structural fatigue is common to other engineering materials, but functional fatigue is typically not (because most structural materials do not exhibit reversible phase transformation). Examples of functional fatigue include "ratcheting", or a gradual length increase during thermal cycling (shape memory effect), and "shakedown," or a reduction in the energy absorbed per during mechanical cycling.

The $A \rightarrow M$ transformation of NiTi has a very small volume contraction (0.023 %), and this near-zero volume change is an important feature for martensitic transformations to exhibit superelasticity and the shape memory effect [41]. Nonetheless, the interfaces between A and M are slightly incompatible, and plasticity is required to accommodate their kinematic mismatch [42]. This plasticity accumulates as residual strain after repeated thermal or mechanical cycling. Functional fatigue includes this irrecoverable strain (from plasticity and residual martensite). More broadly, though, functional fatigue is manifested as changes in phase transformation's characteristic stresses and strains. For superelasticity, functional fatigue involves decreased transformation stresses ("shakedown"), decreased recoverable strain upon cycling, and accumulation of residual strain. For the shape memory effect, functional fatigue involves decreased transformation strains and accumulation of residual strain ("ratcheting").

Functional fatigue remains a significant challenge for the durability and service life of SMAs, and there is a lack of understanding about the fundamental origin of functional fatigue on the microstructural level. For example, irrecoverable strain can be generated with thermal cycling under a bias stress that is much lower than the yield stress [43]. Also, after only a few mechanical cycles, significant residual martensite can remain in grains [8]. Furthermore, after only a few thermal cycles, a high density of dislocations can form in austenite [43].

If enough plasticity accumulates (as part of functional fatigue), then cracking (structural fatigue) will follow. On top of the challenges of understanding functional fatigue in SMAs, structural fatigue adds new complications from fracture mechanics. For example, at relatively low stress intensities (near the fatigue threshold of a material, at which crack initiation occurs), there can be a strong effect from crack closure. Furthermore, the theories of fracture mechanics, especially linear elastic fracture mechanics, can be violated by SMAs [2]. Since existing theories of fatigue and fracture mechanics are limited by these complexities, durability predictions for SMA structures are primarily based on costly and slow empirical measurements. A deeper understanding of SMA fatigue and fracture from new observations and experimental approaches will enable more accurate component life predictions for both existing and future SMA applications.

Although SMAs have dramatically varied mechanical responses as a function of temperature, there are no comprehensive measurements of SMA fracture along a continuum temperature with more than three temperature data points. As described in Section 1.2.2, thermomechanical coupling in NiTi SMA is driven by reversible martensitic phase transformations between B2 cubic austenite and B19′ monoclinic martensite (a third phase, "R-phase", sometimes has a non-negligible role, to further complicate the system). The transformation between austenite and martensite can be either thermally or mechanically activated. Austenite and martensite have different mechanical properties such as elastic modulus, and different volume fractions of austenite and martensite can exist at different temperatures and pressures (analogous to a phase diagram of better-known materials, such as H_2O). Furthermore, SMAs have additional strain accommodation mechanisms beyond elasticity and slip/plasticity. In the martensite phase, SMAs can accommodate additional strain with detwinning. In the austenite phase, SMAs can accommodate additional strain by transforming into martensite (giving rise to superelasticity, or the recovery of large strains on the order of 5 to 10 %).

Several works have addressed the role of microstructure in SMA fatigue and fracture, but most microstructurally-sensitive measurements have been application-specific (e.g. with processing conditions to match stents) and have not addressed basic microstructural variations such as bulk anisotropy (such as processing/rolling directions), grain size morphologies, and precipitate distributions. All of these microstructural variations are known to affect the mechanical properties of SMAs, but have not been explored comprehensively in the context of fatigue and fracture mechanics. Additionally, the microstructure and temperature sensitivities of SMAs are intertwined. For example, the connection between transformation temperature and stress for the austenite-martensite phase transformation can typically be modeled with the Clausius-Clapeyron relation, but in NiTi with grain sizes less than about

60 nm, the Clausius-Clapeyron relation no longer holds.

1.2.4 Core: probing the role of phase transformation on SMA fatigue and fracture with advanced techniques in experimental mechanics

In light of these uncertainties from thermomechanical coupling and microstructural sensitivities, this work provides new evidence to inform durability predictions of SMAs. This work builds upon the large body of scientific work about NiTi by using it as a model material for SMAs in order to:

1. address scientific knowledge gaps about martensitic phase transformation in the presence of cracks, and

2. inform both current and future applications of SMAs in areas that are sensitive to fatigue and fracture, such as health care, automobiles, aerospace systems, civil structures, and beyond (as surveyed in Section 1.2.1).

The mechanical response of SMAs is complicated by temperature, crystallographic texture, and grain size (Figure 1.6). Relatively small changes in grain size and temperature can lead to dramatic changes in the stress-strain response within one cycle, as shown in Figure 1.6, as well as among several cycles (with significant differences in functional fatigue for various textures, grain sizes, and temperatures). Although these features of grain size, anisotropy, and temperature dependence still represent challenges for measuring and modeling the thermo-mechanical response, they can also be leveraged to make observations and conclusions about the role of phase transformation on the fatigue and fracture mechanics of SMAs. The undertaking of this work is studying the fatigue and fracture mechanics of SMAs by systematically exploring these features as experimental variables, with grain size in Chapter 5, texture in Chapter 6, and temperature in Chapter 7.

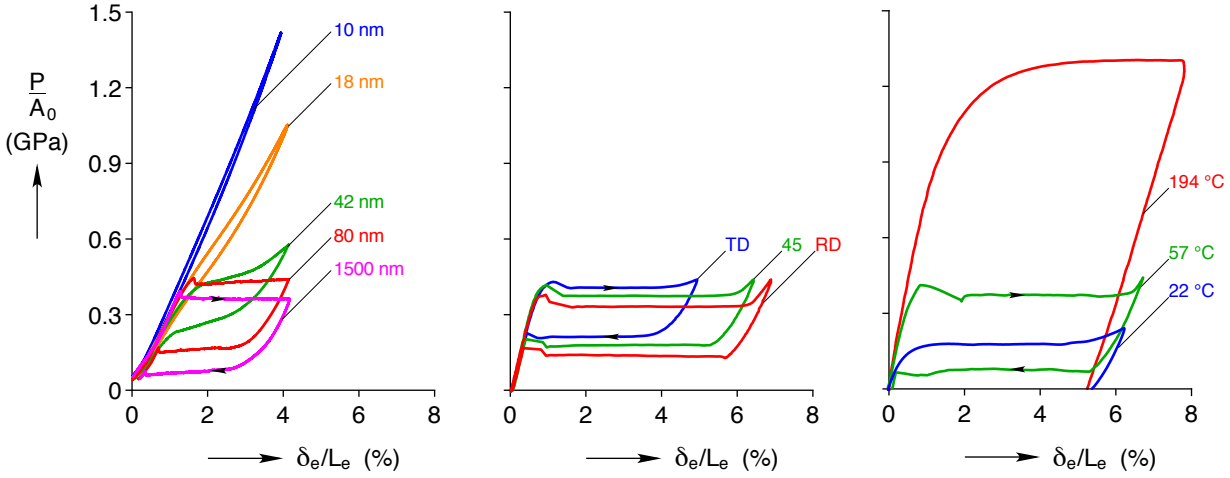


Figure 1.6: The mechanical response of NiTi SMA in one tension cycle is strongly affected by grain size (left; data from Ref. [44]), crystallographic texture (center), and temperature (right). These three features that complicate the mechanical response of NiTi are leveraged as experimental variables in this work to probe the role of phase transformation on the fatigue and fracture mechanics of SMAs.

1.3 Introduction to digital image correlation

1.3.1 DIC fundamentals

Digital image correlation (DIC) is a surface displacement measurement technique that can capture the shape, motion, and deformation of solid objects [45].[2] The basic operation of DIC is tracking a pattern (often called a *speckle pattern*) in a sequence of images. The process of a DIC experiment (illustrated in Figure 1.7) can be divided into three steps:

1. obtain a pattern on the sample for tracking,
2. capture images of the sample during motion/deformation, and
3. analyze the images to compute the sample surface's displacements.

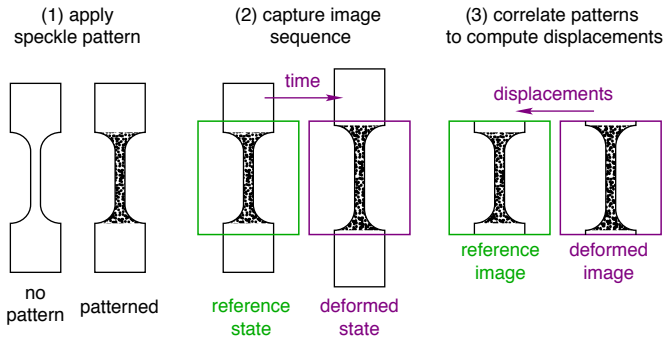


Figure 1.7: The fundamental steps of digital image correlation are: (1) apply a speckle pattern; (2) capture an image sequence; and (3) correlate patterns in the deformed configuration with their positions in the reference configuration.

The first image in the sequence is defined as the *reference* image, orthe baseline to which the other images are compared. DIC matches the pattern between the reference image and a deformed image, and then calculates the pattern's displacements between the reference and deformed images.

A simplification of the DIC analysis is illustrated in Figure 1.8, with the following five steps:

(a) The reference image has a recognizable pattern of dots that will be tracked.

(b) A portion of the pattern, called a *subset*, is selected for tracking. The term *facet* is also used instead of *subset*, but this thesis adopts the term *subset*.

[2]This chapter has been reproduced and adapted from a website by the author, http://digitalimagecorrelation.org/, under the Apache License 2.0 (https://www.apache.org/licenses/LICENSE-2.0).

(c) The center of the subset (the red dot, which is *not* part of the speckle pattern) is the place in the reference image from which the displacement will be calculated.

(d) After the material is deformed from the reference image's initial position, the subset in the deformed image is matched to the subset from the reference image.

(e) Once the subset is matched, DIC calculates the subset center's relative displacement between the reference and deformed images. The displacement here is the (small) difference between the blue and red dots. The next example will show how this basic operation is extended to multiple subsets and DIC points.

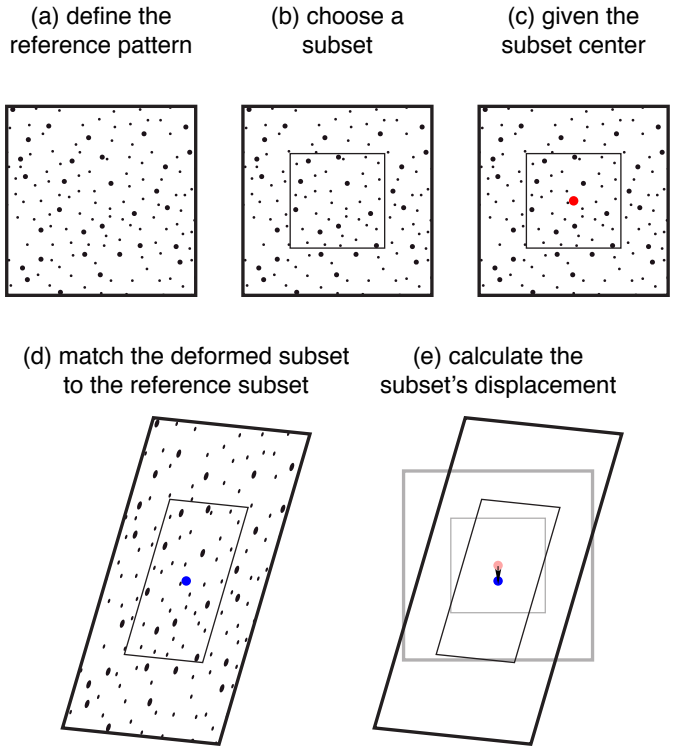


Figure 1.8: For each data point of DIC, there are five steps for a DIC analysis: (a) define a reference pattern with recognizable features; (b) select a subset of pixels within that reference pattern with a certain center position, (c); (d) match the deformed subset to the reference subset; and (e) calculate the subset's displacement.

The previous example computed the displacements from one subset, but DIC computes a *field* of displacements by tracking multiple subsets. The same procedure as before is repeated in Figure 1.9, except this time with four equally-sized subsets in a two-by-two grid. This yields four more points with displacement information, for a total of five data points.

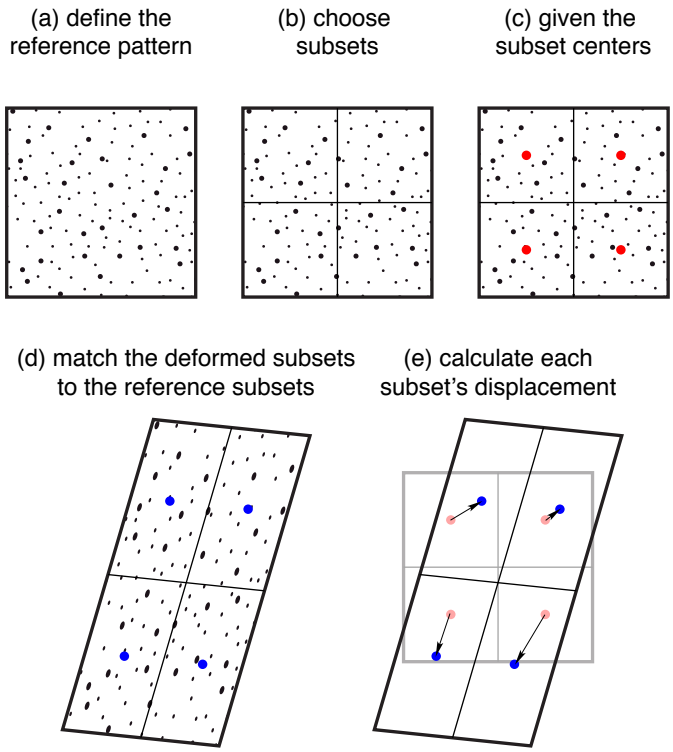


Figure 1.9: The same procedure from Figure 1.8 is repeated for many subsets, with each subset contributing one DIC data point. Shown here are four subsets.

From the five subsets (one from the first example, and four more from the second example), there are five total points for which the displacements have been calculated. Each of these points can be referred to as a *DIC point*.

The displacement at each DIC point is a vector, so the components of the vector can be decomposed. For two dimensions of displacement, the components can be written in a Cartesian coordinate system as the horizontal displacement (u) and vertical displacement (v). Three dimensions of displacements (u, v, and w) can also be measured with a more complicated type of DIC that uses triangulation. More details on the different types of DIC are presented in Section 1.3.2.

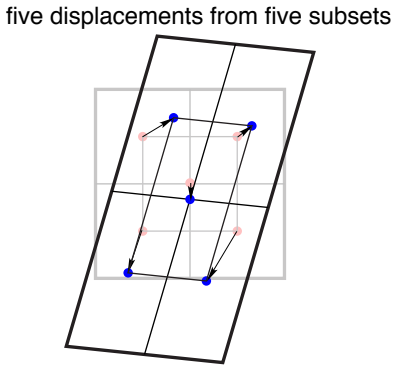


Figure 1.10: The single subset from Figure 1.8 and the four subsets from Figure 1.9 are combined (with overlapping permitted) to form a DIC "field" with five displacement data points.

Macroscopic DIC is commonly utilized to study the mechanical properties of solids. One of the most common experiments for solid materials is a *uniaxial tension* experiment, shown in Figure 1.11. The goal of this experiment is to quantify how much the material deforms when a force is applied. There are many different ways to measure the deformation of the material, including strain gauges, extensometers (mechanical, laser, or optical), and DIC. While strain gauges and extensometers provide a single measurement of strain or displacement in the material, DIC can provide many surface displacement measurements across the material. With displacements across the material, DIC can capture the local deformations that arise from inhomogeneity, cracking, stress concentrations, plastic instabilities, phase transformations, and other localized material phenomena.

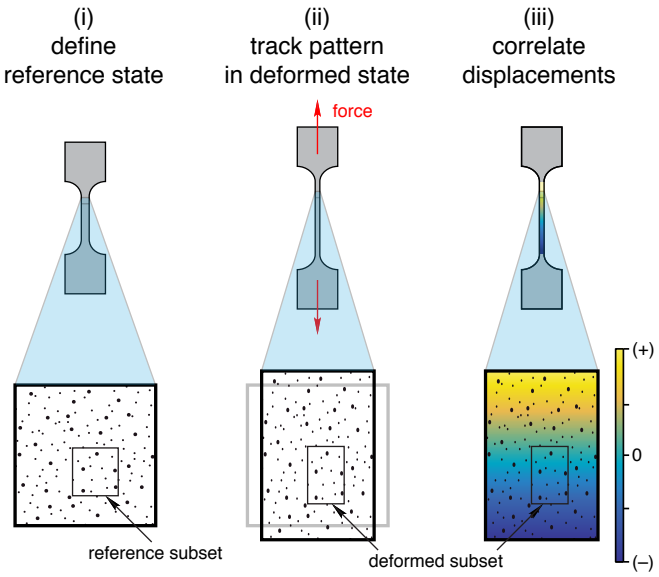


Figure 1.11: In the context of experimental mechanics, DIC is used in conjunction with materials testing to track deformation during a known deformed state (for example, a prescribed global deformation with a measured amount of force).

1.3.1.1 Subset and step sizes

Two important dimensions in a DIC calculation are the *subset size* and the *step size*. The subset size is the width and height of the subset square in the reference image. The step size is the distance between subset centers. Both the subset size and step size are measured in units of pixels. The most important factor for determining subset size is that each subset should contain at least three speckles [45]. A secondary factor for choosing subset size is the competition of better pattern matching for bigger subsets (with more uniqueness for a larger area of features in the image) versus better spatial resolution for smaller subsets (with less spatial smoothing/filtering of the image data). A third factor is that larger subsets require more computation time. The step size has a much stronger effect on spatial resolution than the subset size. Smaller step sizes yield more DIC data points and thus higher spatial resolution. The diagram below shows a range of step sizes for the same field of view and subset sizes.

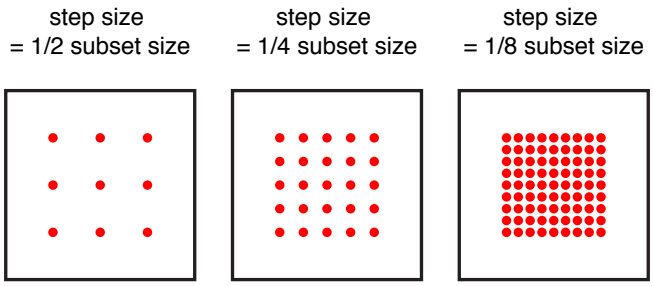


Figure 1.12: Extending from Figure 1.10, subsets can overlap to increase the density of data points in the DIC data field, down to the maximum resolution of neighboring pixels also having neighboring DIC data points (for a "step size" of one pixel, which is the minimum spacing among subsets).

1.3.1.2 Spatial and temporal resolution limits

Two important questions for planning DIC experiments are, (1) What is the smallest displacement that this experiment can reliably measure?, and (2) What is the image exposure time that should be used? Both of these questions are addressed here.

The smallest possible displacement measurement is limited by the quality of the experiment's images. In general, DIC algorithms are capable of detecting sub-pixel displacements on the order of 0.01 px. Sub-pixel displacement resolution is enabled by interpolation (generally bicubic spline interpolation) on the image data. A simple example of this interpolation is demonstrated below on a portion of an image from a DIC experiment. In practice, experimental variables introduce error into the measurements, and the smallest displacement measurements that can be expected from DIC, often called the noise floor, is on the order of 0.10 px.

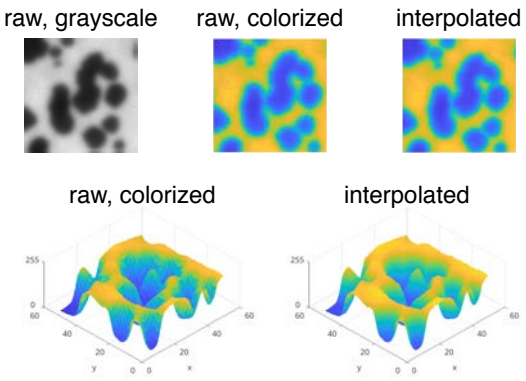


Figure 1.13: Interpolation is performed on the grayscale intensity of DIC images to enable sub-pixel displacement precision.

Since the noise floor of DIC in practice is about 0.10 px, then the images should be captured with an exposure time that limits the sample motion during the exposure time to less than 0.10 px (or even less than 0.05 px to be safe). If the exposure time exceeds the noise floor of the calculation, then the images would have blurring that would deteriorate the DIC displacement accuracy. For example, if the sample displacement is 1 micron/second, and the imaging resolution is 10 microns/px, then the DIC noise floor is 0.10 px (10 microns/px) = 1 micron, so the image exposure time should be less than 1 second (or less than 0.1 second to be safe, by matching the limit of 0.01 px for DIC algorithms).

1.3.1.3 Strain calculation

Computing strains is a common post-processing step with the displacements from DIC. Generally, spatial strains are computed from the displacements with a spatial derivative that has a filtering operator. This filtering in the strain calculation can blur highly-localized deformations with sharp features or discontinuities. With spatial filtering to compute the strains, the sharp features are smoothed out and can hide key phenomena like slip bands [46].

A second precaution for using DIC strain data is that the field of strains calculated with spatial derivatives on DIC displacements make a small strain assumption (from infinitesimal strain theory, in contrast with finite strain theory). In the small strain assumption, the higher order term of the strain calculation is neglected. Broadly, the small strain assumption has less than 10 % error for strains less than 10 %, but is problematic for strains larger than 10 %. An alternative to using spatial strains with DIC data is creating "virtual extensometers" that measure a one-dimensional engineering strain (change in length divided by original length). For a uniaxial tension experiment with DIC, many virtual extensometers can be defined across the gauge length of the sample, and then an average engineering strain (and statistics on the measurement noise) can be computed.

1.3.1.4 Suggestions for when *not* to use DIC

DIC is not a miraculous measurement technique, and there are many experiments in solid mechanics that are better suited to other techniques. Two such cases are:

1. **Measuring very small strains**: The minimum resolvable strain for DIC varies from setup to setup, but a rule of thumb is that the noise floor of DIC in practice is about 0.1 px. If the strain(s) of interest would correspond to displacements of less than 0.1 px displacement, then DIC will not give a reliable measurement. Consider mechanical extensometers, laser extensometers, and strain gauges instead.

2. **Measuring deformations with low pixel-size sensors (e.g. ultra high speed cameras)**: DIC has lower spatial resolution than other full-field measurement techniques (most notably, the grid method). The grid method is a Fourier-based analysis on the deformation of a regular grid pattern. More information and open-source codes on the grid method can be found at http://www.thegridmethod.net/.

1.3.2 Types of DIC algorithms

One way to categorize DIC algorithms is by the dimensions of the calculated displacements. For images collected by just one camera, only two dimensions of displacements can be known. This is called two-dimensional, or *2-D DIC* (also commonly written as 2D-DIC). When images from more than one camera are used, depth can be measured with triangulation. This is called three-dimensional, or *3-D DIC* (also commonly written as 3D-DIC).

2-D DIC assumes that the sample's deformations are constrained to a plane that is parallel to the camera. In practice, out-of-plane motion can be a large source of error for 2-D DIC [47]. Also, images can have distortions that introduce error into DIC measurements. For example, camera lenses and optical microscopes generally have barrel distortions.

For 3-D DIC, out of plane deformations are measured with triangulation. As long as the sample remains in focus, then the out-of-plane deformations do not introduce error in 3-D DIC (unlike 2-D DIC). Furthermore, lens distortions are corrected in 3-D DIC through a calibration procedure. See Section 1.3.5 for more details about calibrating 3-D DIC systems. Another benefit of the 3-D DIC calibration process is that the length scale of the images are accurately connected to the physical length scale of the imaging system. In contrast, the length scale of 2-D DIC is introduced by a simple and less accurate conversion between the pixel size of the images to the physical size of the images (e.g. millimeters).

An important note is that 3-D DIC can only measure displacements on the surface of a material, not within the three-dimensional volume of a material. This extension of DIC from pixels to *voxels* (three-dimensional pixels) is called digital *volume* correlation (DVC) [48]. To measure displacements within a solid, the imaging system must be able to see inside the material, and the algorithms of DIC must be extended to capture displacements through the volume. Two examples of imaging systems that can see inside materials are X-ray tomography and confocal microscopy.

A comparison among 2-D DIC, 3-D DIC, and DVC is illustrated in Figure 1.14.

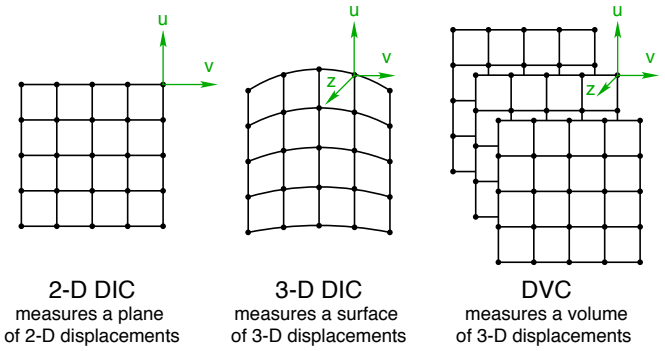


Figure 1.14: Broadly, there are three DIC types with respect to the dimension of the data: 2-D DIC (measures a plane with two dimensions of displacements), 3-D DIC (measures a surface with three dimensions of displacements), and DVC (measures a volume of three dimensions of displacements).

A second way to categorize DIC algorithms is by the pattern matching technique. The pattern can be separated into multiple subsets that are individually matched. This is called *local DIC*. Alternatively, the pattern can be matched in one go using a finite-element based approach. This is called *global DIC*. Local DIC was introduced before global DIC, and local DIC is more popular. Many of the principles in this guide apply to both local and global DIC, but details that only apply to local DIC (such as subsets) are included.

1.3.3 Speckle patterning

To match the reference and deformed images, DIC tracks features on the sample surface that collectively form the *speckle pattern*. Occasionally, a sample's surface will inherently have features that suffice for a *natural* speckle pattern, but typically an *artificial* speckle pattern must be applied to the sample. The quality of DIC results are strongly dependent on the speckle pattern, and optimum speckle patterns meet the following conditions.

1. The pattern covers the sample surface in the area of interest.

2. The pattern moves and deforms with the sample, but does not exert a significant mechanical stress on the sample. In other words, the pattern is fully adhered to the sample, but deforms extremely easily compared to the sample.

3. The features that comprise the pattern (the speckles) are random in position but uniform in size.

4. The speckle size is at least 3 pixels to avoid aliasing [49], but not much more than 7 pixels to achieve a relatively high density of DIC points [50]. If speckles are much

larger than 7 pixels, then there will be relatively few DIC data points possible. Also, note that these speckle sizes are not averages, but are rather the range of the smallest and largest speckles [51].

5. The pattern has good grayscale contrast, which reduces error [45]. One way to visualize this contrast is a histogram: with the number of pixels plotted with respect to grayscale level, the pattern has a mix of dark and bright pixels, indicated by two peaks in the histogram's spectrum, and the separation between the two peaks is broad. Ideally, the two peaks look like a bimodal Gaussian distribution.

6. The edges of the speckles are softened (rather than sharp and distinct with the background). This allows the pixels of the camera sensor to avoid aliasing the speckle edge [52].

7. The pattern is stable in the testing environment. For example, for a high-temperature experiment, the pattern does not decay or darken under heating.

8. The pattern has a speckle density of about 50 % [4]. When the pattern has either too few or too many speckles, then this results in features that are both too big and too small [51]. This concept is illustrated in Figure 1.15.

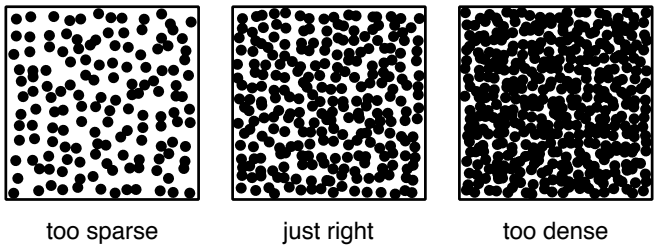


Figure 1.15: The optimum speckle pattern has a speckle density of about 50 % [4]. If there are either too few or too many speckles, then effectively this creates a pattern with features that are both too big and too small. For example, for a pattern that is too sparse (left), the whitespace acts as speckles that are too large, and then comparatively the black speckles are too small. These artificial speckle patterns were generated with the Speckle Generator software from Correlated Solutions, Inc.

In most cases, the sample's natural surface is not the best pattern that could be achieved. There are many ways to apply artificial speckle patterns, and the main techniques are listed below. Also, see Ref. [53] for a summary of speckle patterning for many length scales, and Ref. [54] for a summary of speckle patterning for microscopic length scales.

Paint: Painted speckle patterns are popular because paint is relatively compliant to most engineering materials, and high-quality speckle patterns can be applied quickly with spraying paint. Since paint colors other than black and white will inherently have less contrast, black and white paints are recommended. Using white paint as the background and black as the speckles is favored over the converse order because black paint maintains better contrast over white paint [4], as detailed in Chapter 3.

If the sample will undergo large deformations and/or high strain rates, then plan to perform the experiment within 24 to 48 hours of painting. As the paint dries and hardens, it loses its ability to deform with the sample [55]. A range of speckle sizes can be produced with sprayed paint. Artist grade airbrushes, such as the Iwata CM-B, can produce speckle sizes between 10 and 100 microns by varying the airbrush pressure (more pressure creates smaller speckles). Cans of spray paint can produce larger speckles, in the range of 100 to 1000 microns.

Inks and dyes: For hyperelastic materials (including many elastomers, polymers, and biomaterials), paint is not stretchy enough to track with the sample as a speckle pattern. Inks and dyes that permeate the sample material can be viable speckle pattern options. Stamping, masking, spraying, and stenciling can be deployed to apply the ink or dye. For example, biological soft tissue can be stained with methylene blue and then airbrushed with white paint for speckles [56]. Some DIC practitioners use permanent markers, as well.

Powder particles: For moist or sticky materials, powder particles may adhere better than paint. Graphite powder is popular for dark speckles, and alumina or magnesium oxide can be used for a white basecoat. Another use for powder patterns is achieving smaller speckles than painted patterns can produce. Using a combination of filters and compressed air, powder particle patterns smaller than 10 microns can be deposited on a smooth/polished sample to form a speckle pattern [57].

Nanoparticles: For even smaller speckles than powders (about 20 to 100 nanometer speckle size, for scanning electron microscopy digital image correlation), self-assembled nanoparticles can be utilized [58].

Lithographed patterns: Lithography is another method for achieving small speckle size, with the benefit of a higher degree of control than most other microscale patterning methods [59].

1.3.4 Image capturing

The key step of data collection for DIC is capturing images. In general, good photography or microscopy practices translate to good DIC images, but there are extra considerations for optimizing DIC results.

The first consideration is selecting the appropriate image magnification. The image magnification depends on the length scale of the samples and phenomena that the experiments will investigate. The algorithms for DIC are inherently length scale independent, so the physical length scale conversion arises from the image magnification. For example, DIC has been performed on the length scale of meters to track the collapse of Mount St. Helens [60], all the way down to single atoms with transmission electron microscopy [61]. Most commonly, though, cameras are used to capture DIC images.

Building a successful DIC setup requires making the right equipment choices. For optical DIC systems with cameras, lenses, and lights, there are a few selection criteria that optimize the system:

1. For DIC images, color is superfluous and can be problematic. The best practice is to select black-and-white cameras. Often, cameras that are marketed for machine vision applications are very suitable for DIC, as well.

2. The camera sensor should have low noise, high quantum efficiency, and high dynamic range. Historically, charge-coupled device (CCD) sensors have outperformed complementary metal-oxide-semiconductor (CMOS) sensors, but new advancements in sensor technologies have leveled the playing field between CCD sensors and the next-generation of CMOS sensors (e.g. Sony Pregius).

3. Lenses should have low distortion. The best lenses for DIC are telecentric, which means that the sample's magnification does not vary within the lenses depth or field of view.

4. For lenses with adjustable apertures, use the mid-range apertures, say f-stop 5.6, 8, or 11 for a lens with an f-stop range of 1.8 to 22 (the more extreme apertures introduce more distortions in the imaging).

5. The lenses and cameras should be rigidly mounted on an optics table (ideally) or on a high-quality tripod, and sources of vibration should be minimized. Be sure to clamp, tie, or tape down the camera cables, too.

6. For systems with two or more cameras, much sure the cameras are viewing the same area. Small differences between the height of the epipolar lines can make calibration difficult [62].

7. Good focus is critical. Make sure that the most important region in your area of interest is the best-focused area.

8. Lighting should be evenly distributed along the sample's area of interest. The lighting should be intense enough to achieve sufficient exposure for the images, yet not too intense to introduce saturated pixels. A saturated pixel is the maximum value of the sensor (such as 255 for an 8-bit sensor, $255 = 2^8 - 1$). When a pixel saturates, DIC can no longer perform sub-pixel interpolation at that pixel.

9. The lighting should also avoid introducing too much heat. Halogen lamps are very bright but also very hot, while LEDs are cooler (but high-intensity LEDs can still generate a significant amount of heat).

10. For maximizing optical DIC results, polarizing filters can be placed orthogonally on the lights and lenses (a photography trick called cross polarization). For DIC, cross polarization increases contrast, decreases error, and attenuates saturated pixels that prevent sub-pixel correlation [3], detailed further in Chapter 2.

11. Lastly, using a fan to gently blow air over the DIC setup is pragmatic because the turbulent flow prevents heat waves from distorting the images [63].

Once a set of fantastic hardware is assembled into a DIC system, care must be taken to ensure its long-term performance. The best practices for cleaning cameras and lenses are summarized below.

1. The lenses and cameras must be cleaned to remove dust. Two easy ways to check for dust on DIC gear:

   (a) Shine a flashlight on the lens or sensor cover and move the flashlight around at different angles. Any specks of dust will be more visible.

   (b) Point the camera at a uniform, diffuse, bright light (e.g. light panel with a diffuser) and increase the exposure just until the image is mostly saturated, then moving the camera and seeing and darker spots in the view that don't move while you're moving the camera.

2. Be especially careful while cleaning cameras, lenses, and sensors (only clean sensors with protective covers or panels), because improperly cleaning the gear can introduce permanent scratches. For a primer on cleaning photography gear, B&H photo has a great guide available at http://www.bhphotovideo.com/explora/photography/tips-and-solutions/how-clean-your-lens-and-filters.

3. To remove dust, do not blow air at the lenses from an air can or from your mouth, which can introduce moisture on the lens that leaves a residue. Use a photography type duster instead.

4. If the duster doesn't get everything, then escalate to lens cleaning tissues. Only use new, clean lens cleaning tissues that have been moistened with lens cleaning solution.

1.3.5 Calibration

For 2-D DIC, the only calibration is a length scale conversion from the pixel space of DIC to the image's magnification, so the calibration needs a line of known length (such as the horizontal field width, or HFW). For 3-D DIC, the cameras must be calibrated with respect to one another in space, so a line is no longer sufficient. Thus, common calibration procedures involve a calibration grid, or a plane of known dimensions. This concept is illustrated in Figure 1.16. The artificial speckle pattern and mock calibration grid were generated with the Speckle Generator and Target Generator softwares from Correlated Solutions, Inc.

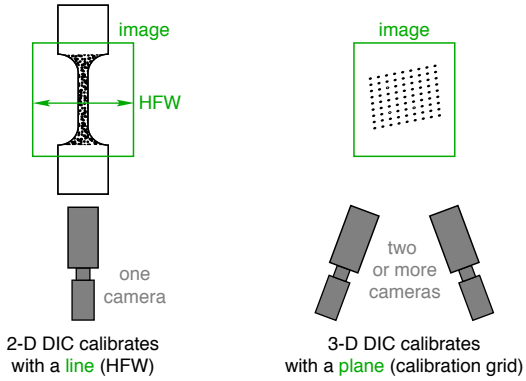


Figure 1.16: DIC calibration takes different approaches for the different types of DIC. For 2-D DIC, only the horizontal field width (HFW) is needed to calibrate the pixel size. 3-D DIC requires a more complicated calibration procedure that involves several images of a calibration grid with a plane of dots.

For the line length measurement to calibrate 2-D DIC, small errors in the line length can create large errors in the resulting displacements, so an accurate reference such as a precision ruler should be used.

For 3-D DIC, the calibration procedure varies among DIC software packages, but general best-practices are listed below.

- For an image horizontal field width (HFW) smaller than about 25 mm, a glass calibration grid with laser-etched marks is recommended due to the high precision at the small length scale. Printed grids on paper suffice for HFWs larger than about 25 mm. If using a printed grid on paper, then firmly affix the printed grid to a flat and rigid substrate, such as a section of PMMA sheet.

- For glass calibration grids, backlight the grid either with a diffuse LED light panel or with indirect backlighting. One indirect backlit option is shining a light on a white poster board with a matte finish.
- With the grid in the cameras, set the light intensity to just before saturation to get good contrast.
- Incrementally move and rotate the grid in the field of view, then take an image pair.
- Take 25-50 calibration images (more can produce even better results, but with diminishing returns).
- To hold the calibration grid, a dial indicator arm mounted on a two-axis stage with a magnetic base is fantastic. A lower-cost option with similar effect is a third-hand. An example of a dial gauge holder with a calibration grid is shown in Figure 1.17.

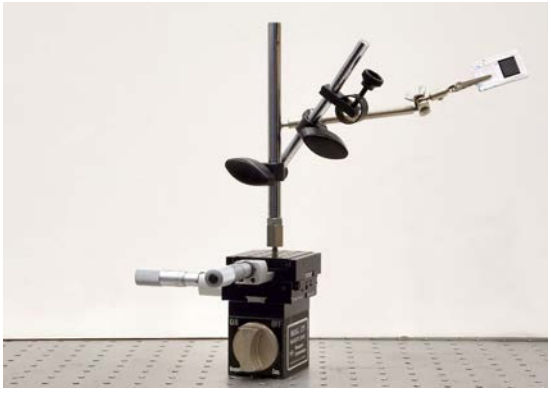


Figure 1.17: An expedient way to hold a calibration grid for 3-D DIC calibration is the magnetic base stand commonly used for dial gauges. Note that the calibration grid is shrouded in tape on the perimeter to cushion a fall in the event of dropping the grid.

1.3.6 Validation and error evaluation

An important question for a DIC practitioner to ask is, "How accurate are these measurements?" Evaluating DIC errors is important because DIC results can vary widely among setups, speckle patterns, correlation parameters, and other conditions. It is difficult to separate the relative contributions of individual error sources in DIC experiments, but the overall error of an experiment can be estimated with appropriate measurements. Two possible validation and error evaluation methods are:

1. Measure the noise floor of the setup and speckle pattern by capturing the "Type A" errors from repeated, static images (no motion or deformation between images) [64].

The mean and distribution of the displacements represent the noise floor (which should be about 0.1 px, and is not zero because of noise and error).

2. Compare the displacements (or strains) from DIC with a second measurement. A direct comparison with displacements can be accomplished with precise rigid body translations on the sample from a trusted source (e.g. Vernier micrometer or precision linear stage). The displacements can also be measured from an LVDT, laser, or other techniques. Strains and deformations can be compared with extensometers (mechanical or laser), as well as strain gauges (although strain gauges are limited to small strains).

CHAPTER 2

Cross polarization for improved optical digital image correlation

2.1 Introduction

Digital image correlation (DIC) is a surface displacement and deformation measurement technique that was pioneered in the 1980s [65, 66] and has become a popular method in experimental mechanics. It provides capabilities beyond the averaged, discrete measurements provided by traditional strain gages and extensometers. DIC is a non-contact, optical method that measures nonuniform displacement fields, providing quantitative measurements of heterogeneous material deformations arising from features such as stress concentrations, sites of damage accumulation, and localized phase transformations. DIC is normally limited to measuring surface displacements, but the principles of DIC have been extended to digital volume correlation (DVC) for measuring internal displacements with through-thickness imaging methods such as x-ray tomography [67] and confocal microscopy [48]. In this work, we will address improvements to macroscopic, optical, surface DIC measurements.

The basic DIC procedure consists of (1) applying a speckle pattern to the surface of a specimen, (2) taking a sequence of grayscale digital images of the specimen during a mechanical experiment, and (3) performing DIC analysis, usually as a post-processing step (since it is computationally intensive) to calculate displacements and strains of the specimen surface. Thus, good measurements require high-quality speckle patterning and high-quality images.

1. The attributes of optimum speckle patterns include a non-repeating pattern of dots (speckles), high grayscale contrast, and speckle diameters of at least 3 px to avoid aliasing [49, 68], but still small enough to maintain desired spatial precision [51]. Good patterning to obtain speckles of a relevant size and random distribution is an art, usually applied for macroscopic testing by airbrush painting (and by techniques such

as self-assembling nanoparticles for microscale testing [58]). While important to the success of DIC, speckle patterning is not the focus of this work. Recommendations can be found in prior work [16, 62].

2. The fundamental data acquisition method of DIC is image capturing. If a single camera is used, only in-plane displacements (u, v) are obtained, so it is termed 2-D DIC. If two synchronized cameras are used, all three displacements (u, v, w) can be obtained, so it is termed 3-D DIC (or stereo DIC)[1].

3. Many DIC codes have been developed, but they use the same methodology to optimize a cross-correlation function of grayscale intensities of the patterned surface to compute point-by-point displacements between a reference image and deformed image(s). The earliest and most common algorithms perform local DIC by discretizing the area of interest into groups of pixels called subsets [66]. Global DIC methods have also been developed [69], with a set of nodes akin to a finite element mesh instead of the subset-based methods of local DIC. We used local DIC in this work, but the same requirements for good speckle patterning and imaging apply to both methods.

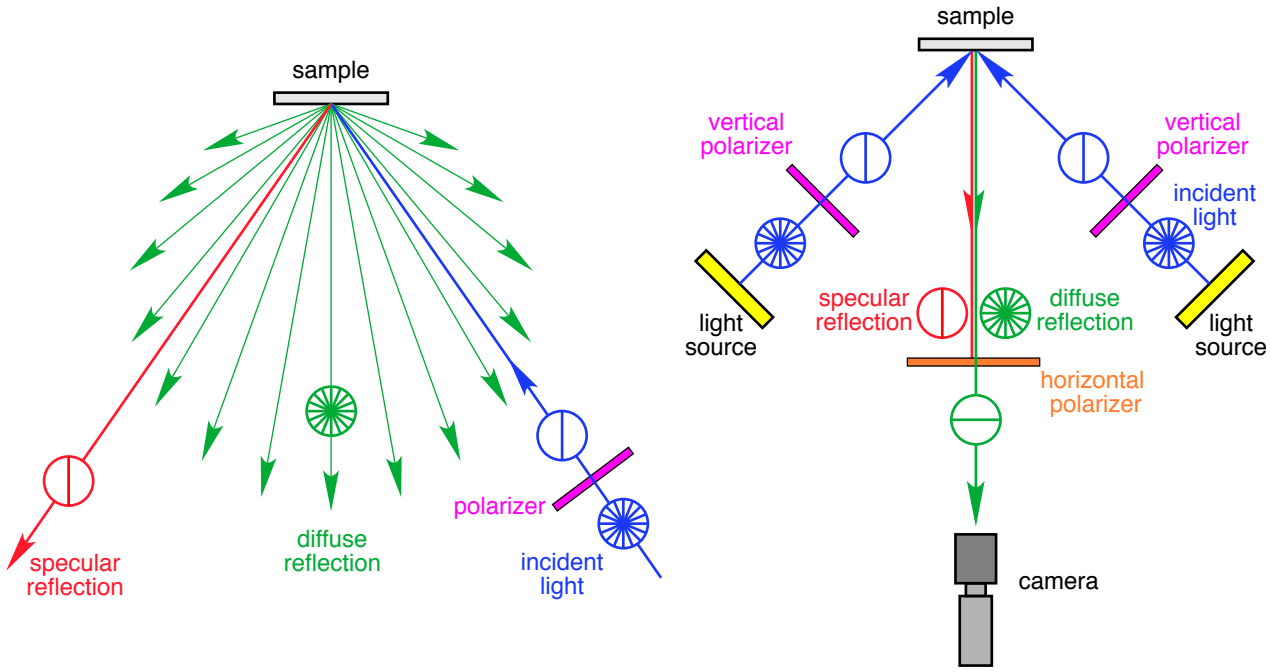
Metrics on the influence of image quality for DIC analysis (step 2 above) and recommendations to achieve optimal results are presented in this work. In short, the usual best practices for good scientific photography apply here. The goals for imaging are sharp focus, good depth of field (where required), high contrast, and uniform and diffuse illumination. Sharp focus may require relatively fast exposure times to minimize noise [70] and motion blur [68], yet this requires intense lighting. The lighting should be bright enough for fast exposure times, yet should not cause saturated pixels, which prohibit sub-pixel displacement accuracy and increase correlation error [68]. Usually, optical DIC is best performed with direct line-of-sight to the specimen surface, to avoid refraction issues and aberrations from intervening optics or other transparent materials (such as environmental chamber windows), since imaging through glass can increase correlation error [71]. Other (perhaps obvious) best practices include the use of rigid camera mounts, the use of moderate lens apertures to avoid accentuated lens distortions at extreme apertures, the removal of dust from lenses and camera sensors, and the avoidance of thermal gradients that also can cause light aberrations.

This work characterizes the technique of cross polarization for optical DIC, and presents two experiments to quantify the effects of cross polarization with common cameras, lights,

[1]We have chosen "3-D DIC" to describe our DIC measurements with three dimensions of displacement. While 3-D DIC does provide all three displacements, it is still a surface measurement that only provides in-plane gradients of displacement (thus, four components of the deformation gradient, or equivalently three strain components and one rotation). It should not be confused with volumetric DIC (DVC), which can provide the full 3-D deformation gradient (nine components, or six strain components and three rotations).

and speckle patterning method. Although cross polarization is an established technique for studio [72], archival [73], forensic [74], dental [75], ophthalmological [76], and dermatological [77] imaging, it is a new practice for DIC. Independently and simultaneously, the group of Cooper, Skaggs, and Reu, and the group of LePage, Shaw, and Daly introduced cross polarization for DIC experiments [78, 79]. When light illuminates a specimen, reflections can be roughly classified into two types, specular reflections[2] (from smooth areas of the surface) that preserve their light polarity and diffuse reflections (from rough areas of the surface) that do not. Specular reflections appear to the camera as high intensity (shiny) spots on the surface and can lead to pixel saturation (glare). From rougher areas of the surface, however, light intensity is diminished and light polarity is strewed when scattered from asperities and sub-surface scattering centers (Figure 2.1a) [80]. With flat samples, specular reflections can be minimized by placing the lights and cameras in positions that avoid direct reflections (which have equal angles for incident and transmitted light rays) from the sample surface. This is not possible with rounded samples such as tubes and cylinders, however. Thus, the benefits of cross polarization described in this work are likely even greater for round samples than for the flat samples presented here. The technique of cross polarization simply consists of placing a linear polarizer between the light source and the specimen and a second polarizer (of a perpendicular polarization axis to the first one) between the specimen and the camera. Thus, the orthogonal polarization between incident and reflected light off the specimen surface selectively attenuates just the specular reflections, resulting in diffuse lighting to the camera (Figure 2.1b).

[2]The terminology is unfortunate in this context. In the optics community, "specular" reflections are defined as those reflections that maintain their polarity. This should not be confused with the "speckle" DIC pattern, which is defined as a random (non-periodic), black-white pattern of dots.



(a) Specular and diffuse reflections

(b) Cross polarization

Figure 2.1: (a) Specular reflections maintain the polarization of incident light, while diffuse reflections do not. (b) The horizontal polarizer on the camera's lens attenuates the vertically polarized specular reflections to avoid saturated pixels.

2.2 Rigid body translation experiment

First, a comparison between unpolarized and cross-polarized imaging for macroscopic optical DIC is presented with 2-D DIC and 3-D DIC setups simultaneously capturing precise rigid body translations of a sample with different speckle patterns.

2.2.1 Speckle patterning

Four different speckle patterns were prepared on a single sample to compare unpolarized versus cross-polarized images. A thin base coat of white paint (Golden High Flow Titanium White, no. 8549-4, lot 366162) was sprayed on a 50 mm × 50 mm area of an unfinished aluminum sheet (80 mm × 80 mm × 0.4 mm thick) with an airbrush (Iwata CM-B with 0.18 mm fluid nozzle and needle, and Iwata Studio Power Jet air compressor) at 210 kPa (30 psi) until all of the aluminum in the painted area was just covered. This was done to avoid glare from the aluminum [55], and complete coverage was verified under a stereo microscope (Olympus BX51M stereo microscope in bright field with 5× magnification MPLN objective lens). Speckle patterns were applied by airbrush with a 1:1 mixture of methanol and black

paint (Golden High Flow Carbon Black, no. 8040-4, lot 345307). Methanol is added to the black paint to decrease viscosity and maintain good paint flow and consistent speckle sizes at low airbrush pressures. To create four patterns of different speckle sizes, a different airbrush pressure was used in each of the four 25 mm × 25 mm quadrants of the white-painted area, while masking the other three quadrants (Figures 2.2 and 2.4). The quadrants are labeled I, II, III, IV, respectively, in order of increasing pressure 70, 140, 175, and 420 kPa (10, 20, 25, 60 psi), corresponding to decreasing speckle size. The best speckle patterns have speckles larger than 3 px (as imaged with the DIC setup) to avoid aliasing [49, 68], but not much larger so that DIC data resolution is maximized [51]. Here, quadrant I represents acceptable speckle size, while patterns III and IV have undersized speckles. The undersized speckle patterns are included to assess the efficacy of cross polarization for imaging sub-optimal speckle patterns.

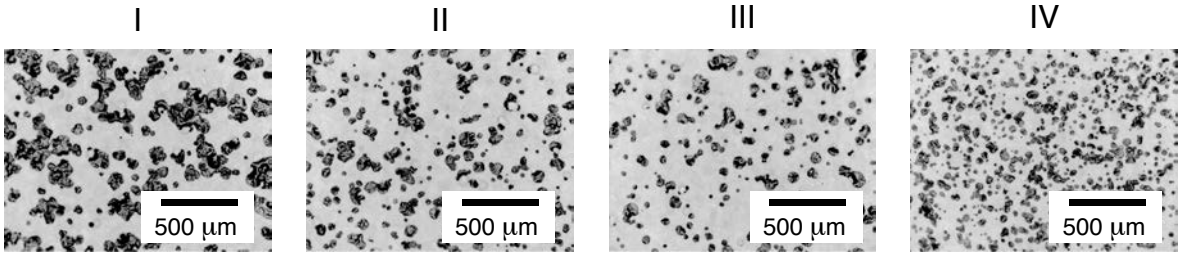
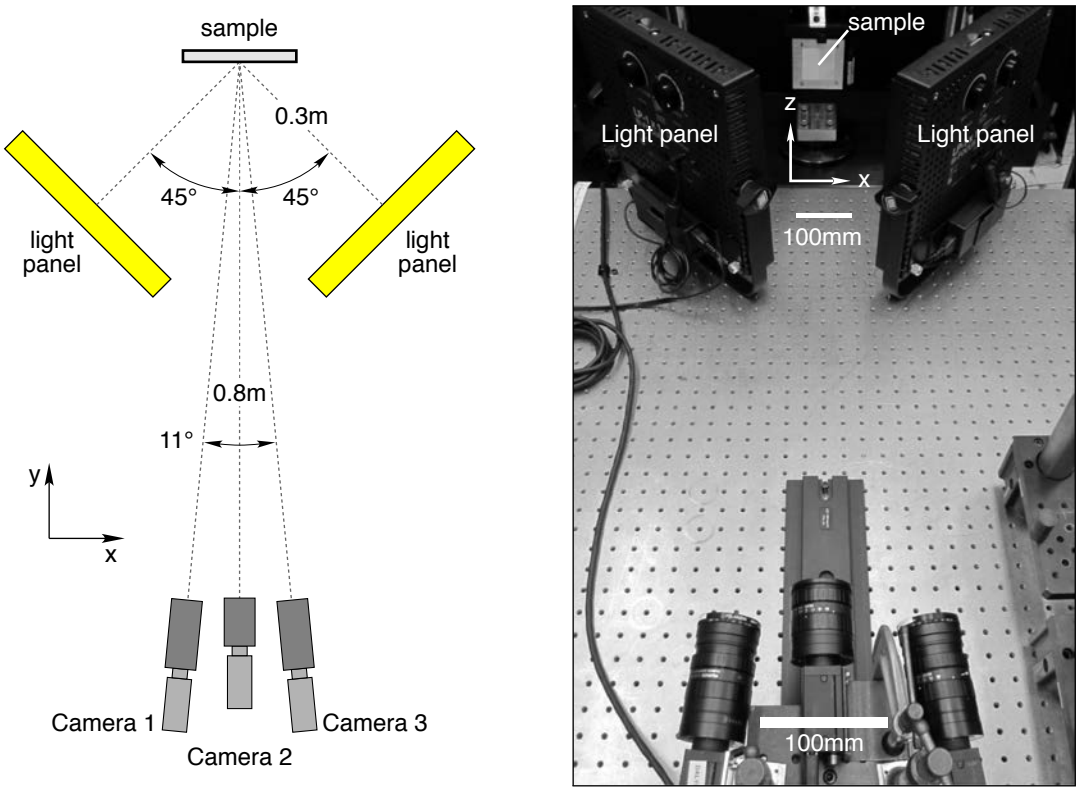


Figure 2.2: Light microscope images showing the four quadrants (I, II, III, and IV) with progressively increasing airbrush pressures (70 kPa, 140 kPa, 175 kPa, and 420 kPa, respectively) and consequent decreasing speckle sizes. Quadrant I has acceptable speckle size when imaged with the DIC cameras, with the exception of some undersized speckles, while the other quadrants, especially III and IV, have undersized speckles to characterize unpolarized versus cross-polarized conditions for sub-optimal speckle patterns.

2.2.2 Imaging setup

The painted aluminum sample was placed in a linear translation stage, driven by a motorized actuator (Newport TRB25CC) with 0.1 μm minimum increment of motion, 1.5 ± 0.2 μm bi-directional repeatability from backlash, and 25 mm range of motion. The actuator speed was kept slow (0.2 mm/s, 10% of maximum actuator speed) for all translations, and images were captured when the actuator was stopped. Three CCD cameras (Point Grey GRAS 50S5M-C) were positioned 0.8 m from the sample, as shown in Figure 2.3. Cameras 1 and 3 were used for 3-D DIC with an included angle of 11° between them, and Camera 2 was positioned between the other two oriented perpendicular to the sample for 2-D DIC. The three cameras simultaneously captured images of the aluminum sample with an aperture of f/8 and 25 ms exposure. Cameras 1 and 3 had 75 mm Fujinon HF75SA-1 lenses, while

Camera 2 had a 50 mm Fujinon HF50SA-1 lens. Two LED light panels (LitePanels no. 903-2076, each with 40W power, or 500W tungsten equivalent, and 1,537 lux at 1.2 m) were placed at 0.3 m from the sample, oriented at $\pm 45^{\circ}$ to provide primary lighting. For the cross-polarized case, linear polarizing PMMA films (American Polarizers, Inc., AP38-030T with 38% transmission) were installed on both light panels in a vertical polarizing direction, and linear polarizing glass filters (Edmund Optics #36-440 with 30% transmission) were installed on all three camera lenses in a horizontal polarizing direction. For the unpolarized case, all polarizing films and filters were removed. Prior to each experiment, the brightness of the light panels was increased just until no saturated pixels were observed in the painted area of the sample. The light panel power knobs were set to about 60% without polarization and 90% with cross polarization to accommodate light transmission losses. Note that the panel power setting changes the light intensity output nonlinearly, so this increase in lighting is about two to three f-stops, or a four-fold to eight-fold increase of the light intensity.



(a) Top view schematic

(b) Photograph of setup

Figure 2.3: Schematic (a) and photograph (b) of the rigid translation experimental setup, with two LED light panels illuminating the sample and three cameras simultaneously capturing images for 2-D DIC and 3-D DIC. The sample was translated upward (z-direction) incrementally to measure known w displacements.

2.2.3 DIC calibration

The 2-D DIC length scale was calibrated with the known distance between corners A and C of the sample's painted area (see Figure 2.4a). The 2-D length scale calibration was measured on the sample with calipers and pinpointed in the images to ±1 px on each endpoint. The paired cameras (1 and 3) were calibrated for 3-D DIC with a laser-printed calibration grid (via Correlated Solutions Calibration Grid Generator software, with 6 mm spacing). The same calibration score (0.059 px residual) was achieved for all 10 trials by selecting among about 25 calibration images.

2.2.4 Procedure

First, the actuator was translated downward to $w = -2.5$ mm, then upward to $w = -1.25$ mm, and upward again to $w = 0$ mm to mitigate any backlash. Reference images were captured with the actuator at $w = 0$, then eight 2.5 mm upward rigid body translations were performed, each followed by an image capture while the actuator was stopped. Five trials were performed for each of the two cases (no polarization and cross polarization), alternating between cases to avoid uneven temporal errors, due to effects such as CCD heating [68]. For each trial, all three lenses were refocused on the center of the sample and the 3-D DIC camera system was recalibrated to the same calibration score. This yielded a total of 80 translated image pairs for 3-D DIC and 80 single images for 2-D DIC.

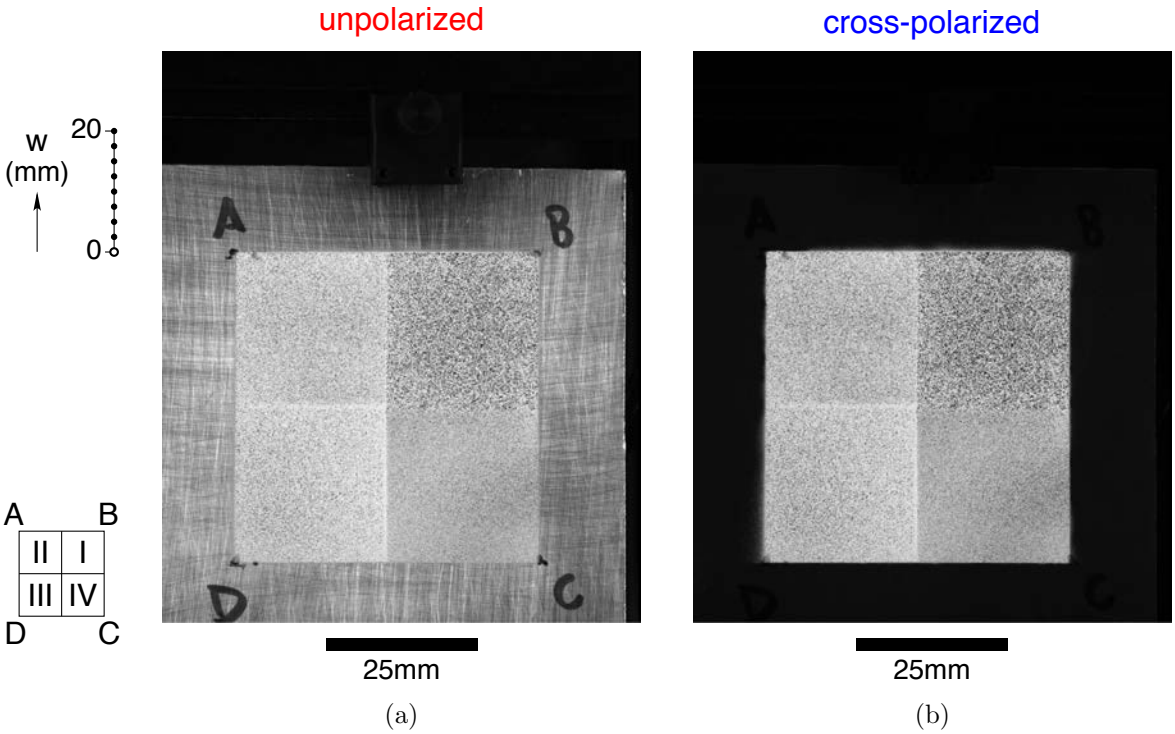


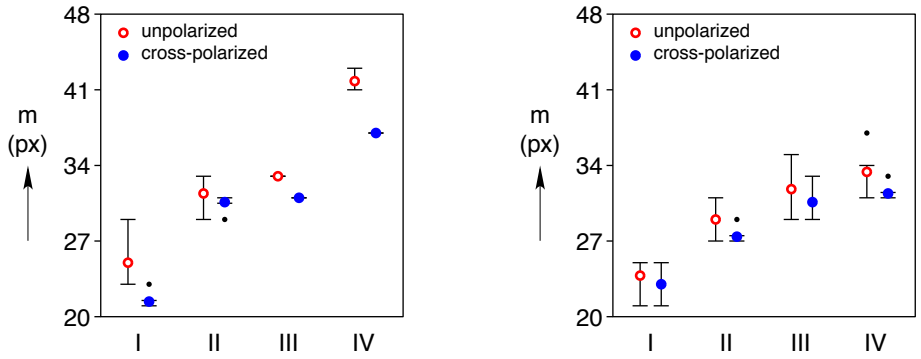
Figure 2.4: A comparison of full images of the sample, (a) without polarization and (b) cross-polarized, shows that specular reflections are attenuated by cross polarization. This is especially evident for the bare metal, but is also true for the speckle painted area (although less obvious to the naked eye). Both images were taken at the reference position $w = 0$ (indicated by the open circle on the vertical axis) by Camera 1 with total field of view of 80 mm wide $\times$ 95 mm tall (2048 $\times$ 2448 px). The sample was translated upward in eight 2.5 mm increments (black dots on the axis) to a final position $w_{\max} = 20$ mm, capturing a sequence of nine images, to assess the effects of distortion as speckle patterns moved away from the center of the field of view.

2.2.5 DIC post-processing

Correlations were performed on the four speckle pattern quadrants, representing different speckle sizes, distributions, and contrasts. The 2-D DIC and 3-D DIC analyses were performed with Vic-2D 2009 and Vic-3D 7, respectively, from Correlated Solutions. The images were first low-pass filtered (with the software's built-in tool) to reduce possible bias errors from aliasing. The images were then correlated using normalized squared differences with Gaussian subset weights and an optimized eight-tap interpolation filter for minimizing bias [68]. Correlation threshold values were 0.02 px consistency threshold, 0.05 px maximum confidence interval, and 0.1 px matchability threshold. Correlation step size (the spacing among DIC data points) was 7 px for all cases, yielding about 2,000 2-D DIC data points and 5,000 3-D DIC data points for each 25 mm $\times$ 25 mm speckle pattern quadrant.

	Optimum subset size (px) for each speckle pattern			
Quadrant	I	II	III	IV
2-D DIC	23	31	33	39
3-D DIC	23	29	31	33

Table 2.1: Fixed DIC subset sizes were chosen to provide consistent comparisons (Figures 2.8–2.10 below) between unpolarized and cross-polarized cases. Each subset size is a selected intermediate value of the optimum values shown in Figure 2.5.



(a) 2-D DIC (b) 3-D DIC

Figure 2.5: Mean values for optimum subset sizes $m \times m$ (px) with cross polarization (blue, closed circles) were consistently smaller than without polarization (red, open circles) in all cases: four speckle patterns I, II, III, IV and using 2-D DIC or 3-D DIC. The error bars (about ± 2.7 standard deviations) also generally show reduced variability of the optimum subset size with cross polarization.

For all four speckle patterns and for both 2-D DIC and 3-D DIC, optimum subset sizes $m \times m$ (px) were consistently smaller for the cross-polarized cases than for their unpolarized counterparts (Figure 2.5). Subset sizes were chosen to optimize the correlation confidence interval (CCI) to 0.01 px for an assumed camera noise level of 8, using the built-in subset size selection tool within the Correlated Solutions software. Average values and spreads for each set of five trials are shown in Figure 2.5. In this (Figure 2.5) and subsequent figures (Figures 2.7–2.9), open (red) and closed (blue) circles denote mean values for unpolarized and cross-polarized cases, respectively, the error bars bound about ± 2.7 standard deviations, and outliers are shown by small black dots. Cross polarization always resulted in a reduction in the average optimum subset size and usually resulted in less variability (all but one case, which was about the same). Smaller subset sizes are generally desirable, since they have less spatial smoothing from their smaller shape functions.

To enable a consistent comparison between unpolarized and cross-polarized results below,

	Total saturated pixel count			
	2-D DIC (Camera 2)		3-D DIC (Camera 1)	
Quadrant	unpolarized	cross-polarized	unpolarized	cross-polarized
I	0	0	36	0
II	31	0	77	0
III	324	0	2965	0
IV	0	0	1	0

Table 2.2: During the rigid body translation experiments, saturated pixels appear in most of the unpolarized trials, despite setting the initial lighting to have no saturated pixels in the reference position ($w = 0$). By contrast, no saturated pixels appear for any of the cross-polarized trials.

however, the same (intermediate) subset size was chosen (Table 2.1) for each speckle pattern and 2-D or 3-D DIC calculation. DIC displacements were computed with 2-D DIC (u, w) and 3-D DIC (u, v, w) and compared pointwise to the actuator translations (w_0) to quantify the magnitude of respective displacement errors, according to

$$\Delta w_{2D} = \left| \sqrt{u^2 + w^2} - w_0 \right|, \tag{2.1a}$$

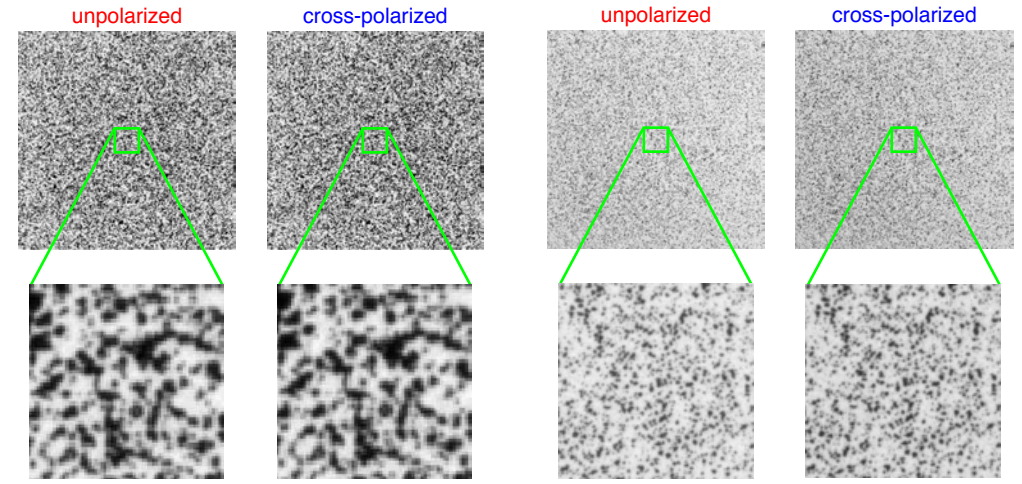
$$\Delta w_{3D} = \left| \sqrt{u^2 + v^2 + w^2} - w_0 \right|. \tag{2.1b}$$

The pointwise errors Δw_{2D} and Δw_{3D} were averaged across each quadrant to obtain four values of $\overline{\Delta w}_{2D}$ and $\overline{\Delta w}_{3D}$ for each image.

2.2.6 Results from rigid body translation experiment

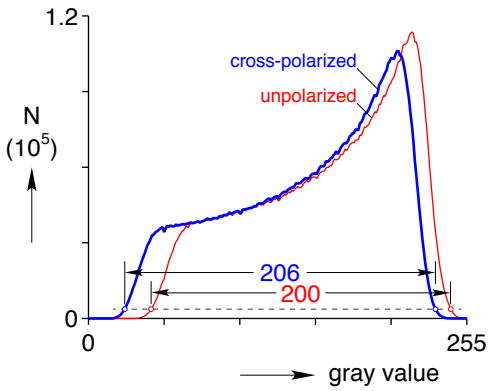
Cross polarization reduces or eliminates saturated pixels, which are undesirable because saturated pixels prohibit sub-pixel accuracy and increase correlation error [68]. The appearance of saturated pixels during optical DIC experiments is difficult to avoid, as lighting conditions effectively change with large translations and deformations, despite fixed light positions and settings. In the experiments presented here, the light panel brightness was increased just until no saturated pixels were observed in the DIC area of interest in a live view of the cameras. After translations of the sample, however, saturated pixels appeared in most of the unpolarized trials, while no pixels were saturated for any cross-polarized cases (Table 2.2). One should recognize that the changes in illumination here are relatively mild for such a simple rigid body translation, and we would expect much larger issues for "real-world" experiments involving out-of-plane displacements and/or curved specimen surfaces (such as column buckling of a cylindrical rod). Thus, the protection that cross polarization

provides against saturated pixels makes DIC measurements significantly more robust.

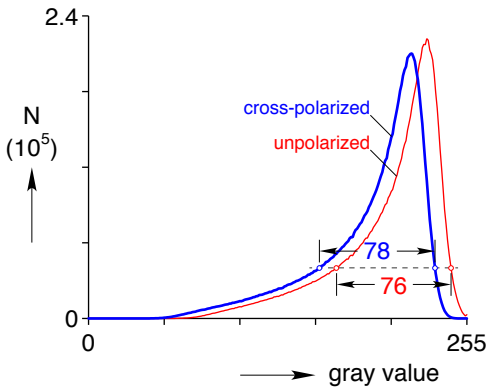


(a) Images of Quadrant I

(b) Images of Quadrant III



(c) Histogram for Quadrant I

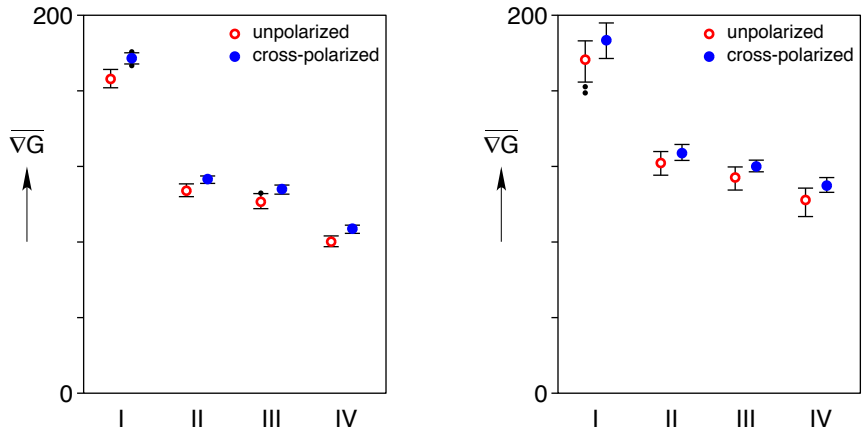


(d) Histogram for Quadrant III

Figure 2.6: A comparison of images and histograms of two speckle patterns (Quadrants I and III) shows that cross polarization results in about a 3% gain in contrast (6 and 2 grayscale counts, respectively) relative to the nominal distribution width. Images (a) and (b) were taken by Camera 1, and their corresponding histograms (c) and (d) give the total pixel count (N) at each of the 256 grayscale values accumulated over all 10 trials.

In addition to minimizing saturated pixels, cross polarization improves image contrast, which reduces DIC correlation error by increasing intensity gradients [70]. Figure 2.6 provides two examples of speckle patterns, Quadrant I (Figure 2.6a) and Quadrant III (Figure 2.6b) both taken by Camera 1, to compare unpolarized and cross-polarized images. The corresponding histograms (Figures 2.6c and 2.6d, respectively) were accumulated across all 10 trials, giving the total pixel count (N) at each grayscale value within the 8-bit (256) dy-

namic range of the CCD sensor. Quadrant I represents a good speckle pattern with a broad distribution of grayscale values, while Quadrant III represents a pattern with undersized speckles as evident from its relatively narrow distribution of grayscale values. Overall, cross polarization darkened both sets of images, which is apparent both to the eye and by the left-shift of the histograms to darker (smaller) grayscale values. The improvement in contrast from cross polarization is less obvious to the eye, but can be seen in the histograms by the slight broadening of each distribution. For Quadrant I, the contrast increased by 6 grayscale counts for 3-D DIC (Figure 2.6c), which might appear to be only a modest improvement, but we consider it an appreciable gain since the unpolarized case already had an optimum speckle pattern and lighting condition. For Quadrant III with finer speckles, cross polarization resulted in a contrast gain of 2 grayscale counts, which is still significant considering the narrow histogram distribution. Relative to the nominal width of each unpolarized distribution (200 and 76, respectively), cross polarization gives about a 3% gain in contrast, even under excellent lighting conditions. We expect larger improvements would be achieved under less optimal lighting, especially over-bright lighting conditions[3]. Furthermore, although not shown here explicitly but still relevant for undersized speckles (like Quadrant III), cross polarization removes any small spots of glare (even black speckles can be smooth and reflective), thereby minimizing the blooming effect[4] of white paint around small spots of black paint (Figure 2.6).



(a) 2-D DIC (b) 3-D DIC

Figure 2.7: Cross polarization increased the mean intensity gradient ($\overline{\nabla G}$) on average by about 7% for 2-D DIC and 6% for 3-D DIC. A large mean intensity gradient indicates high-quality speckle pattern, lighting, and images. Each data set (mean with about ± 2.7 standard deviation error bars) contains 40 data points.

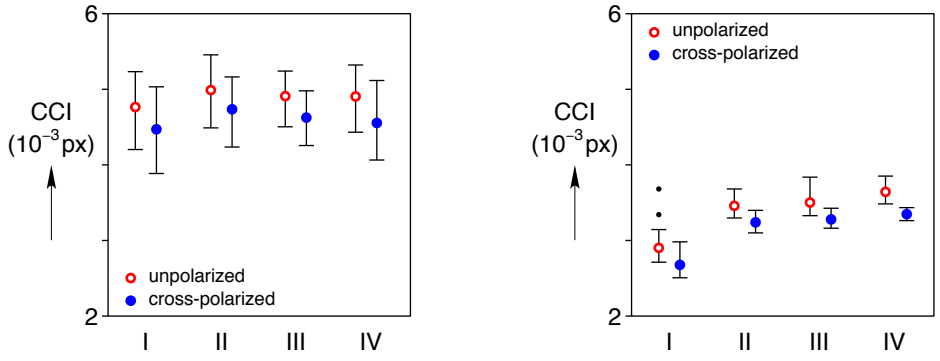
[3]Of course, cross polarization could adversely affect an under-illuminated setup, but that is easily avoided.
[4]Blooming occurs in imaging when fringes extend beyond a bright feature into a dark feature.

Of course, to achieve this increase in contrast, sufficiently bright lights must be used, so that the cross-polarized setup can have the same aperture and exposure time (camera gain) as the unpolarized setup. If the brightness were not adjusted, then light transmission losses from the polarizers would decrease the available light intensity by a factor of about 8. This is why the light source power was increased for the cross-polarized cases, which allowed all images here (whether polarized or not) to be captured at f/8 with a fixed exposure time.

Consistent with the increased contrast, cross polarization also increases the mean intensity gradient (Figure 2.7), which has been proposed as a metric to evaluate DIC speckle pattern quality [81]. The mean intensity gradient ($\overline{\nabla G}$) was computed for each image, (taken by Camera 2 for 2-D DIC and Camera 1 for 3-D DIC) using

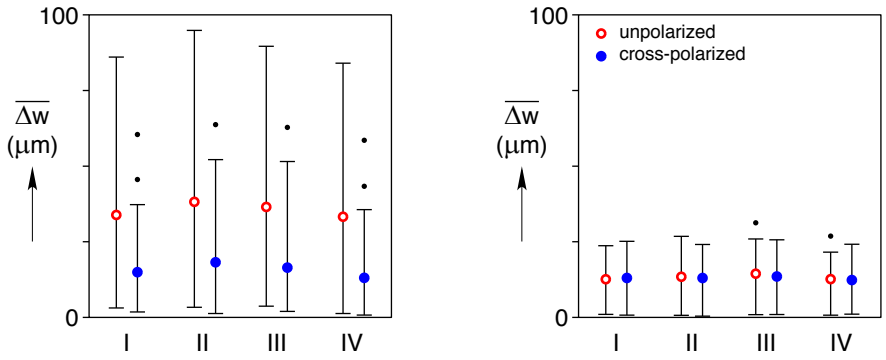
$$\overline{\nabla G} = \frac{\sum_{i=1}^{W}\sum_{j=1}^{H}\left|\nabla f\left(\boldsymbol{x}_{ij}\right)\right|}{WH}, \tag{2.2}$$

where W and H are the DIC area of interest width and height in pixels, $\left|\nabla f\left(\boldsymbol{x}_{ij}\right)\right| = \sqrt{\left(f_{,x}\right)^2 + \left(f_{,y}\right)^2}$ is the magnitude of the local intensity gradient vector, and $f_{,x}$ and $f_{,y}$ are the in-plane derivatives at pixel $\boldsymbol{x}_{ij}$. $\overline{\nabla G}$ was calculated by the Prewitt filter, an approximate differential operator for the intensity gradient [82].



(a) 2-D DIC (b) 3-D DIC

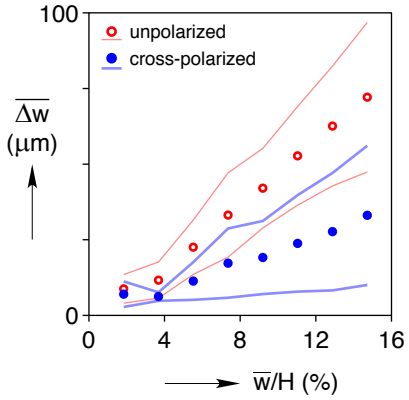
Figure 2.8: Cross polarization consistently provided reductions in the correlation confidence interval (CCI), where a smaller CCI indicates better correlation. Each data set (mean with about ±2.7 standard deviation error bars and dot outliers) contains 40 data points.



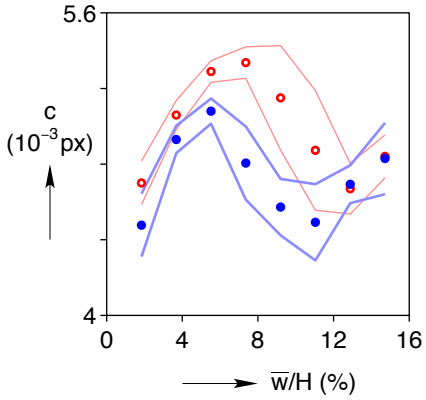
(a) 2-D DIC

(b) 3-D DIC

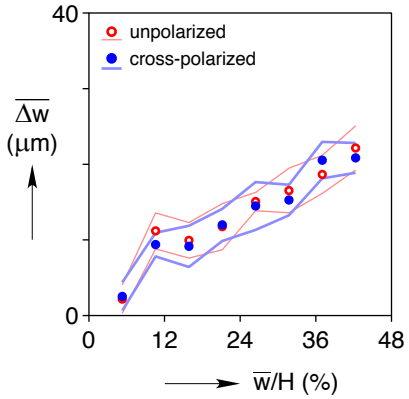
Figure 2.9: For 2-D DIC, significant reductions (56%, 52%, 55%, and 61%) in the mean measurement error were observed for speckle patterns I, II, III, and IV, respectively, by using cross polarization versus no polarization. No significant reduction in displacement measurement error was observed for 3-D DIC by using cross polarization, probably because the unpolarized condition was already optimized for minimum specular reflections by placing the cameras outside of the direct reflections of the light panels. Each data set (mean with about ±2.7 standard deviation error bars) contains 40 data points.



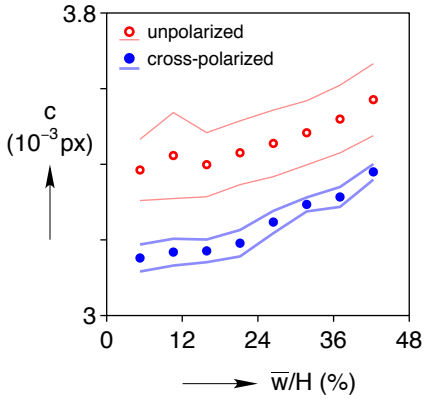
(a) 2-D DIC displacement error



(b) 2-D DIC correlation confidence interval



(c) 3-D DIC displacement error



(d) 3-D DIC correlation confidence interval

Figure 2.10: Measurement errors (Δw) and correlation confidence intervals (CCI) are plotted against the normalized average displacement $\overline{w}/H$ (where H is the field of view height) to show where cross polarization provided improvement. Circles denote mean values for all four speckle quadrants across all five trials of each polarization, and the lines correspond to ± 1 standard deviation from the mean values.

Cross polarization reduced the correlation confidence interval (CCI) by at least 5.2% and 6.3% for 2-D and 3-D DIC, respectively (Figure 2.8). The CCI is a statistical measure of the one standard deviation confidence region for each subset, where a smaller CCI indicates better correlation accuracy. Cross polarization decreased the mean CCI for 3-D DIC, for example, by 7.8%, 6.3%, 6.4%, and 8.1% for speckle quadrants I, II, III, and IV, respectively. A summary of the resulting measurement errors are provided in Figure 2.9. As shown in Figure 2.9a for 2-D DIC, significant reductions in mean measurement error (56%, 52%, 55%, and 61%) were obtained for all speckle patterns (I, II, III, and IV, respectively). Interestingly, no appreciable reductions in displacement measurement error were observed with 3-D DIC

(Figure 2.9b) by using cross polarization, probably since the errors were already quite small.

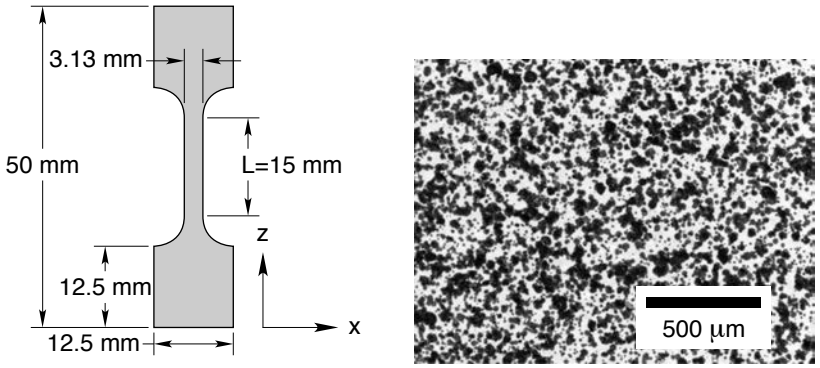
More details of the displacement errors (Δw) and CCIs, plotted against the normalized (average) rigid translation $\overline{w}/H$, are provided in Figure 2.10 to gain a sense of where (and where not) cross polarization provides an improvement. For 2-D DIC (Figure 2.10a), cross polarization on average achieved a 48% lower mean error. Errors were reduced at all displacements, but it made the largest error reductions at the largest displacements. The CCI for 2-D DIC (Figure 2.10b) exhibited a non-monotonic change with displacement, with a local minimum near $\overline{w}/H \approx 12\%$ as the sample translated through the center of the field of view where lens distortions were lowest. For all but one displacement value, cross polarization gave a smaller CCI than without polarization. For 3-D DIC (Figure 2.10c), no appreciable reduction in displacement error was observed, consistent with Figure 2.9b. Cross polarization, however, consistently achieved lower average CCIs across all displacements (Figure 2.10d), indicating better DIC measurements.

2.3 Uniaxial tension experiment

To illustrate the differences between cross-polarized and unpolarized images in a less trivial context with a deforming specimen, a superelastic nickel-titanium (NiTi) specimen was loaded in uniaxial tension. NiTi was chosen for its large recoverable strains, localized phase transformations, and the repeatable pattern of stress-induced martensitic (SIM) phase transformation upon cycling [83]. During tension experiments, these properties result in distinctly nonuniform deformations and allow reasonably congruent comparisons between cycles with different polarization conditions while using the same cameras and lenses.

2.3.1 Uniaxial specimen

The specimen was the same material and dimensions as those from previously published experiments [83, 84]. The test specimen was cut from a flat, 0.281 mm thick sheet of superelastic NiTi, with the rolling direction parallel to the loading z-axis, by wire EDM to the dimensions shown in Figure 2.11a. The front surface of the specimen was patterned with a white base coat at an airbrush pressure of 210 kPa (30 psi) followed by black speckles at 70 kPa (10 psi), using the same equipment reported in Section 2.2.1. A closeup optical micrograph of the speckle pattern is provided in Figure 2.11b.



(a) (b)

Figure 2.11: (a) Schematic of the tension specimen, (b) micrograph of the speckle pattern.

2.3.2 Experimental setup

In this setup all measurements were taken using 3-D DIC. Cameras 1 and 3 were outfitted with 20 mm extension tubes to give an appropriate field of view and were oriented with an included angle between them of 8°, balancing depth of focus while still capturing out-of-plane displacements. All images for the uniaxial tension experiments (about 110 images per load-unload cycle) were captured at an aperture of f/5.6 and a 15.0 ms exposure time. The 3-D DIC system was calibrated with a glass calibration grid (Applied Image p/n 045466, grid A, with 1.780 mm spacing). Between changes to polarization conditions, the lenses were refocused, the sample was removed from the grips, the system was recalibrated to a constant calibration score of 0.036 px residual, and the sample was placed back in the same position in the grips. The same LED light panels as before were used, with brightnesses of about 60% and 90% for unpolarized and cross-polarized imaging, respectively. The specimen was loaded slowly in displacement (δ) control by a TestResources model 830 load frame at constant ramp rates $\dot{\delta} = \pm 0.0024$ mm/s, or $\dot{\delta}/L = \pm 1.92 \times 10^{-4} \mathrm{s}^{-1}$ across the gage section. The axial load resultant P was measured by an MTS model 661.11 (500 N) force transducer. The data presented here for the unpolarized and cross-polarized lighting conditions were captured during loading cycles 4 and 5, respectively, on the same specimen. Figure 2.12 shows the engineering stress-strain response of the specimen for these two cycles, where A_0 and L are the reference cross-sectional area and reference length, respectively, of the specimen's central gage section, and δ/L is strain from a DIC virtual extensometer along the gage length L. DIC results are reported later at two points, ① and ②, on the upper plateau of the mechanical response during stress-induced, phase-transformation front propagation.

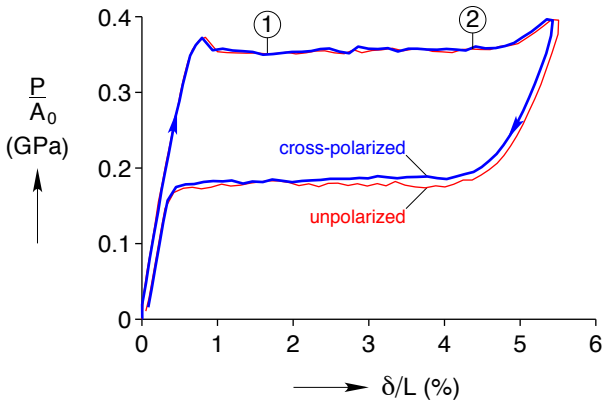
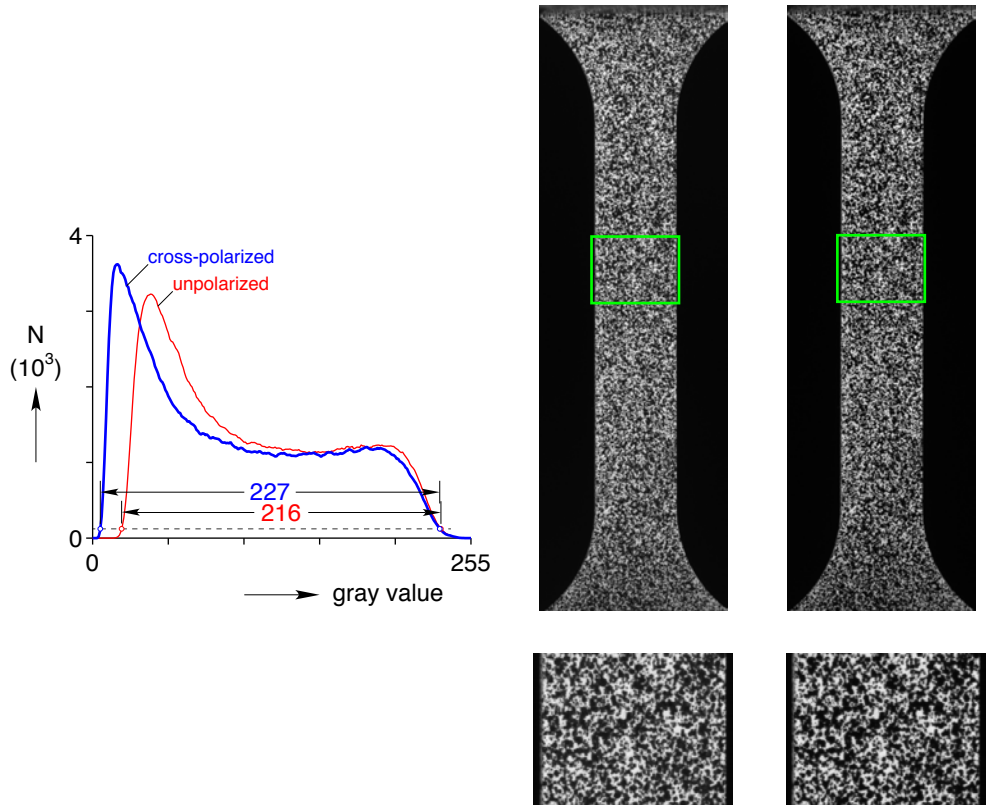


Figure 2.12: Mechanical response of the superelastic NiTi specimen loaded in uniaxial tension and imaged with no polarization and cross polarization during cycles 4 and 5, respectively. DIC results are presented below at ① and ② on the upper plateau of the response.

2.3.3 DIC post-processing

The images were correlated in Vic-3D 7 as before with low-pass filtering, Gaussian subset weights, eight-tap interpolation filter, and normalized squared differences. Correlation step sizes were 3 px to yield about 37,000 DIC data points in the area of interest for subset sizes of 15 px. Correlation threshold values were 0.02 px consistency threshold, 0.05 px maximum confidence interval, and 0.1 px matchability threshold. Using subset size optimization to a CCI of 0.01 px and an assumed camera noise level of 8, the optimum subset sizes for unpolarized and cross-polarized conditions were 27 px and 25 px, respectively. The axial strain and the CCI fields for these optimum subset sizes are presented first in Figure 2.14. Next, the subset size was decreased to 15 px for both lighting conditions to decrease the spatial smoothing at the expense of increased CCI, and the resulting axial strain and CCI fields are shown respectively in Figure 2.15 and Figure 2.16. The strains were computed with a Gaussian filter size of 27 px for the 25 and 27 px optimized subsets and 15 px for the 15 px subsets.



(a) unpolarized (b) cross-polarized

Figure 2.13: Cross polarization increased the contrast by 11 grayscale counts by darkening the black speckles in the image, as seen in the histogram produced from the pixels in the DIC area of interest of the reference image from Camera 1.

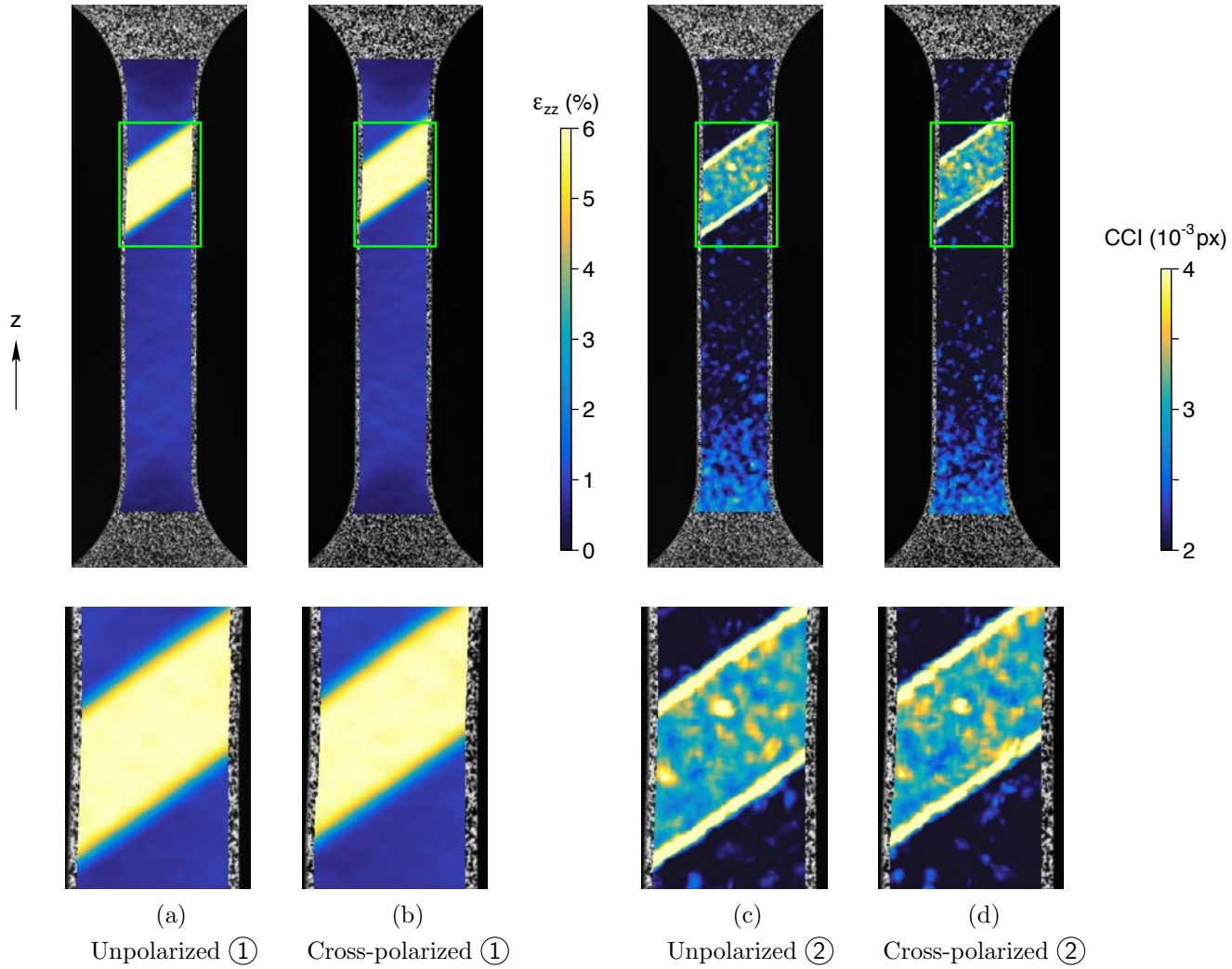
2.3.4 Experimental results

Similar improvements with cross polarization in contrast and mean intensity gradient were observed during the uniaxial experiment as were seen earlier in the rigid translation experiment. As shown in Figure 2.13, cross polarization increased the contrast in the DIC area of interest (across the gage section) by 11 grayscale counts, or 4.3% of the 8-bit sensor's dynamic range. Furthermore, cross polarization increased the mean intensity gradient by 7.4%, from 165.4 to 177.7 grayscale counts. In this experiment, cross polarization again had no saturated pixels, while only two pixels were saturated in the unpolarized images. We expect that glare was less of a problem during this experiment, since maximum displacements were on the order of 1 mm, much smaller than the 20 mm displacements for the previous rigid translation experiment, resulting in more constant lighting conditions.

With the optimum subset sizes (27 px and 25 px for unpolarized and cross-polarized images, respectively), no major differences were observed in the DIC axial strain fields (Figures 2.14a and 2.14b). The same pattern of stress-induced martensitic transformation was

repeated during the two cycles, consistent with the results of Kim and Daly [83]. Cross polarization had a 2 px smaller subset size, but there was no observable difference in the spatial detail of the axial strain fields. However, cross polarization decreased the CCI by 4.6% on average, which can be seen in Figures 2.14c and 2.14d at locations of high strain gradient (across a transformation front).

Smaller subsets increase spatial precision by reducing the size of the DIC data smoothing, but at the expense of reduced confidence in the correlation. Undersized subsets cause points of lost correlation, where correlation thresholds are exceeded. If maximum spatial precision is desired more than complete continuity of data in the DIC area of interest, though, cross polarization was shown to preserve more DIC data points for very small subsets. Although no significant differences were observed with optimum subset sizes, cross polarization decreased CCIs to provide more robust correlations with small subsets. When the subset sizes were reduced to 15 px for increased spatial precision, the unpolarized images lost 1.3% of the DIC data in the area of interest (495 out of 37,459 DIC data points) due to exceeding the correlation threshold values (see the holes in the contour plots of Figures 2.15a and 2.15c). The cross-polarized images Figures 2.15b and 2.15d), however, had no lost correlation points (37,705 DIC data points). Compared to the optimum subset size CCI fields (Figures 2.14c and 2.14d), the decrease in CCI from cross polarization was also more noticeable with the smaller 15 px subsets (Figure 2.16). When the subset size was decreased further to 9 px (not shown), the unpolarized images lost 17.9% of the DIC data points due to exceeding the correlation error thresholds, while the cross-polarized images lost only 2.5% of the DIC data points.



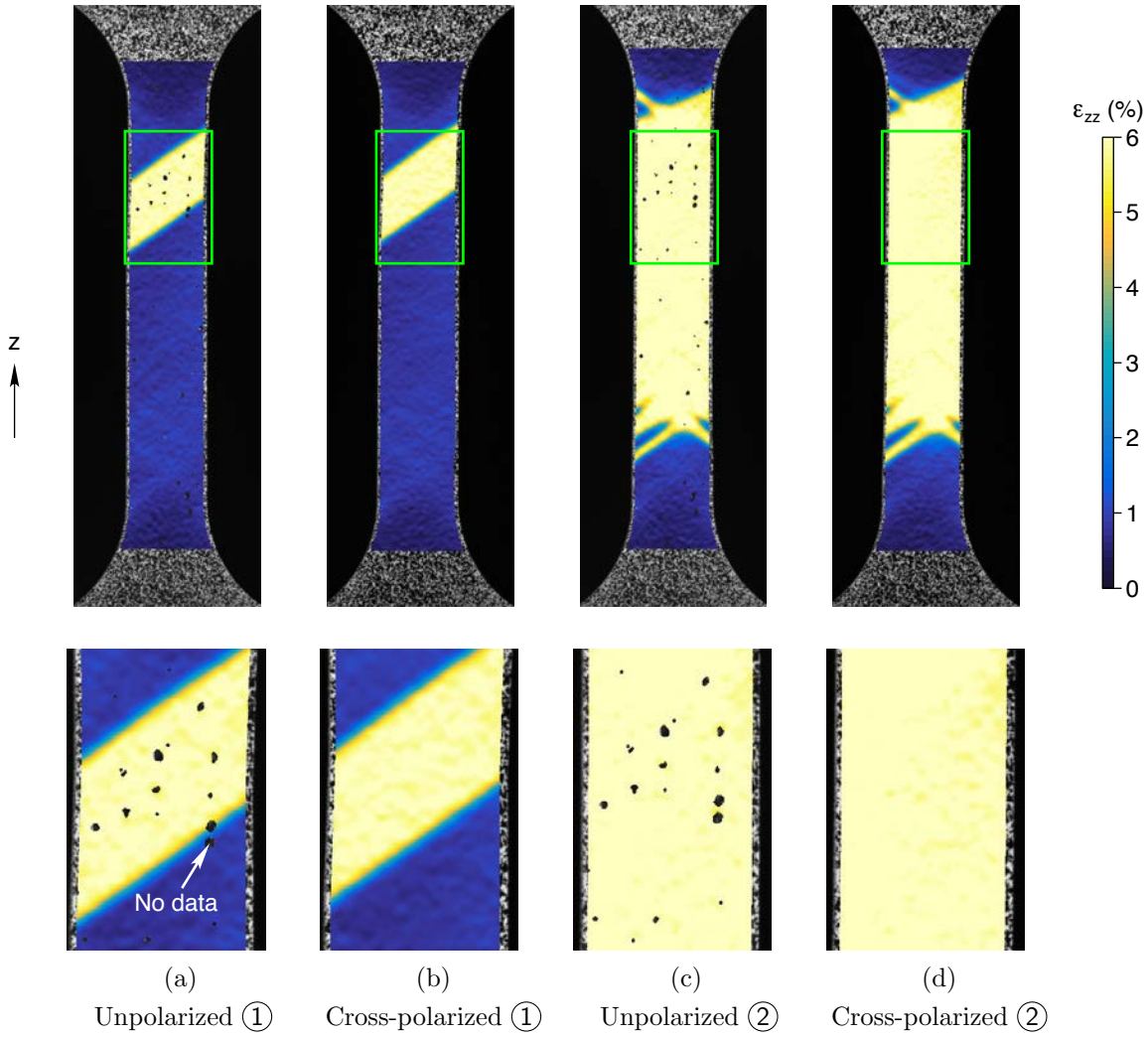
(a) Unpolarized ①

(b) Cross-polarized ①

(c) Unpolarized ②

(d) Cross-polarized ②

Figure 2.14: When using optimum subset sizes (27 and 25 px, respectively), nearly identical axial strain maps were obtained between unpolarized and cross-polarized images, but cross polarization achieved slightly smaller regions of poor (large) correlation confidence intervals.



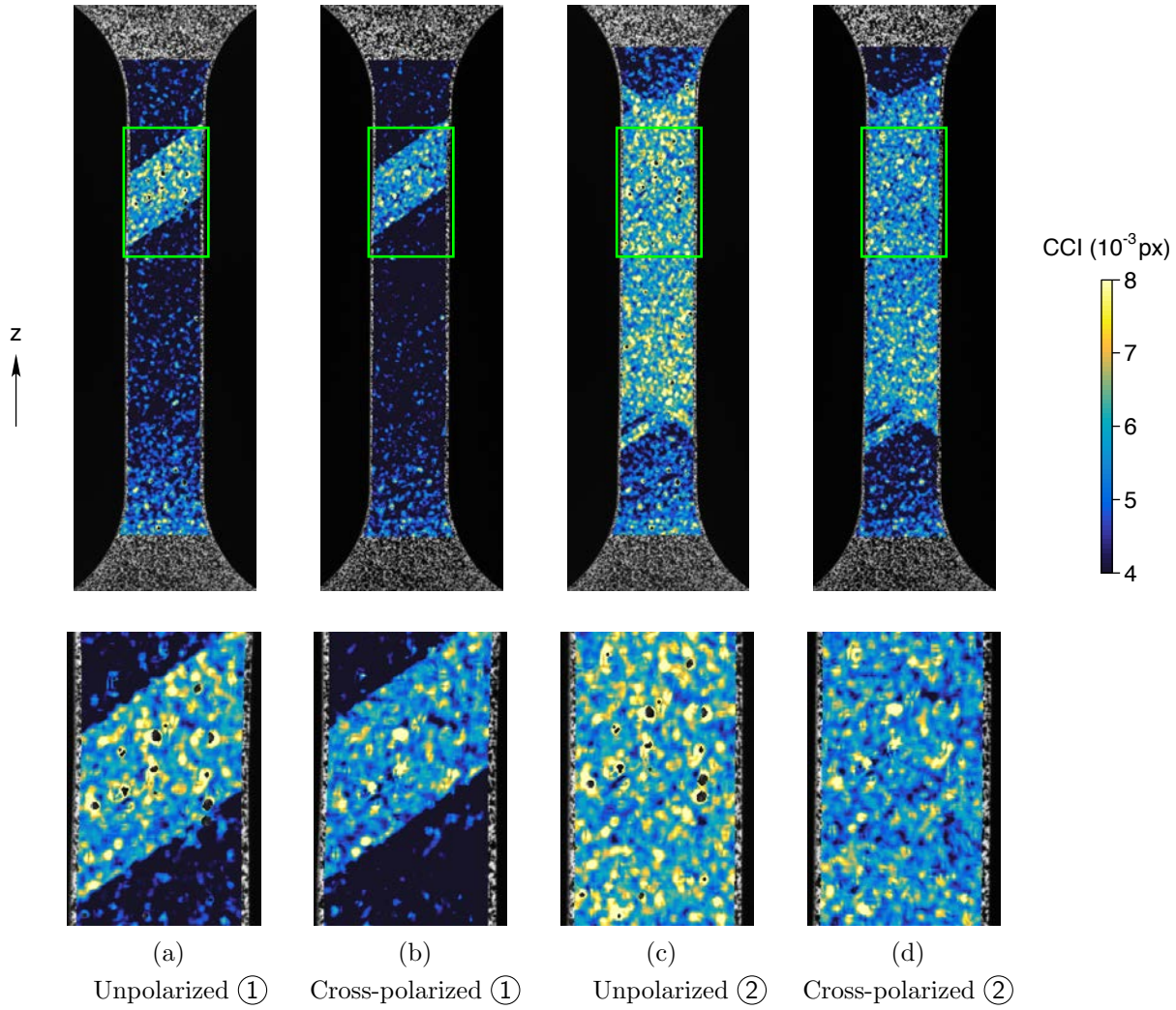
(a) Unpolarized ①

(b) Cross-polarized ①

(c) Unpolarized ②

(d) Cross-polarized ②

Figure 2.15: With a decreased subset size of 15 px, the unpolarized images lost correlation for 1.3% of DIC points, especially in areas of large strains, while cross polarization lost none.



(a) Unpolarized ① (b) Cross-polarized ① (c) Unpolarized ② (d) Cross-polarized ②

Figure 2.16: With a decreased subset size of 15 px, cross polarization maintained better (smaller) correlation confidence intervals (CCI).

2.4 Discussion

As mentioned, cross polarization requires brighter lighting than without polarization because polarizing filters have transmission losses. According to the Law of Malus (Equation (2.3)), who discovered polarized light in 1808 [85], the transmission coefficient Tmax of an ideal polarizer is 1/2, or the average intensity through one period of rotation of the included angle (θ) between the polarizers, calculated as

$$I = I_0 \cos^2 \theta, \tag{2.3}$$

$$T_{\max} = \frac{I}{I_0} = \frac{1}{\pi} \int_{-\pi/2}^{\pi/2} \cos^2 \theta \, \mathrm{d}\theta = \frac{1}{2}, \tag{2.4}$$

where I_0 is the incident light intensity and I is the polarized light intensity. In practice, the transmission coefficient for cross polarizers is less than 1/2. For example, the polarizing film for the light sources in this work had a transmission coefficient of 0.38.

It is not recommended to compensate for light transmission losses by using a larger aperture or a longer exposure time, instead of increasing the light intensity. A larger aperture (lower f-stop) would reduce the depth of field and lead to blurring for out-of-plane displacements. A longer exposure time could also lead to blurring if the specimen displaces too quickly [68] and would increase the effects of sensor noise, a potentially significant source of DIC error [70]. A concern with high-intensity lighting is the potential for heating of the ambient air (causing light distortion) and of temperature-sensitive specimens (such as the superelastic NiTi used here), but this can be avoided with high intensity LED light panels.

Another practical issue is the alignment of the polarizers, which require special care to orient orthogonally. If an imperfection angle θ_0 is introduced between cross polarizers, such that the included angle between the polarized directions is $\pi/2 \pm \theta_0$, the received light intensity $\tilde{I}$ (from imperfectly aligned polarizers) relative to the incident light intensity, extending from Equation (2.3), is

$$\frac{\tilde{I}}{I_0} = \cos^2\left(\frac{\pi}{2} - \theta_0\right). \tag{2.5}$$

This is plotted in Figure 2.17 to show the sensitivity of cross polarization to misalignment. As much as 3% of the light can still pass through the second polarizer with a misalignment of $\theta_0 = \pm 10^{\circ}$. In this work, the angles of the light source polarizers were held equal and constant, and the angles of the camera lens polarizing filters were adjusted for maximum observable glare reduction during live imaging of the specimen.

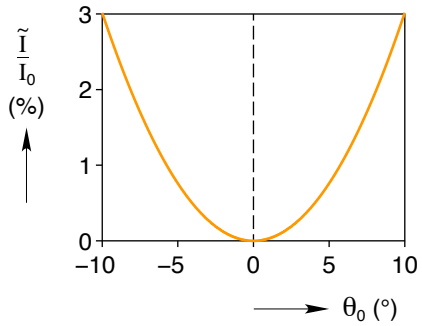


Figure 2.17: Polarizers should be carefully oriented orthogonally to maximize the benefits of cross polarization. The plot shows the relative "leakage" of light with imperfection angle θ_0.

2.5 Summary and Conclusions

Cross polarization is an effective method to prevent saturated pixels, increase image contrast, decrease DIC spatial smoothing, and preserve correlation robustness in optical 2-D and 3-D DIC. Cross polarization can reduce DIC error by attenuating glare that can saturate pixels and prohibit sub-pixel accuracy [68], and by increasing image contrast, which decreases both bias and random errors [70]. Cross polarization can also increase DIC spatial precision while preserving correlation robustness by decreasing the correlation confidence interval and allowing smaller subsets without exceeding correlation error thresholds. The only drawback we see to cross polarization is transmission losses from the polarizing filters, which can easily be counteracted with brighter lighting (although this is more difficult for high-speed imaging setups).

The differences between unpolarized and cross-polarized images were quantitatively evaluated with a rigid body translation experiment and a uniaxial NiTi tension test. For rigid body translations with four representative painted speckle patterns, cross polarization reduced 2-D DIC displacement measurement errors by at least 52% and as much as 61%. No significant changes in displacement measurement errors were observed for 3-D DIC, which already had very low errors for unpolarized images. Also for the rigid body translations, the correlation confidence interval for 3-D DIC was reduced with cross polarization by at least 6.3% and as much as 8.1%, and similar reductions in correlation confidence interval were observed for 2-D DIC.

Cross polarization increased the contrast and mean intensity gradient, two metrics that indicate speckle pattern and image quality. For example, with the uniaxial NiTi specimen, cross polarization increased the contrast by 11 grayscale counts and the mean intensity gradient by 7.4% (from 165.4 to 177.7 units of grayscale count gradient). These increases are significant since the pattern, lights, and cameras were considered optimum before the addition of cross polarization. The increased contrast allowed for very small subsets while maintaining more DIC data points within correlation error thresholds: with 9×9 px subsets, the unpolarized images lost almost ten-fold more DIC data points than the cross-polarized images. The advantages of cross polarization would be even greater for samples with rounded features, such as tubes and rods, since specular reflections cannot be minimized with rounded samples simply by changing the positions of the lights and cameras.

For both the rigid body translation experiment and the uniaxial NiTi experiment, the unpolarized images always had instances of saturated pixels within the DIC area of interest, while none of the cross-polarized images had saturated pixels within the DIC area of interest. Preventing saturated pixels with cross polarization ensures the ability to compute

displacements with sub-pixel accuracy.

Cross polarization requires brighter lighting (about 1.5× higher intensity) than without polarization, since polarizing filters introduce light transmission losses. Although the images would also be brighter with about two to three f-stops wider aperture or longer exposure time, brighter lighting is the best protocol with cross polarization because wider apertures decrease the depth of field and longer exposure times can dramatically increase CCD sensor noise and motion blur. Quality LED light panels are generally suitable because they provide bright, uniform light without specimen and ambient air heating.

CHAPTER 3

Optimum paint sequence for speckle patterns in optical digital image correlation

3.1 Introduction

Optimizing the speckle pattern on a sample is one of the key ways to improve deformation measurements by digital image correlation (DIC). This work shows that, in the case of painted speckle patterns, the order is often important for the pattern's contrasting basecoat and speckles.

DIC is a non-contact surface displacement measurement technique that is popular for characterizing the mechanical properties of solid materials. DIC computes displacements across the surface of a material, so it resolves nonuniform deformations from local features such as material heterogeneities and interfaces. The three fundamental steps of DIC are to (1) apply a pattern of contrasting features, or "speckles," on the sample surface, (2) capture a sequence of images during mechanical deformation of the sample, and (3) compute the sample's surface displacements with a DIC algorithm. The quality of each of these three steps determines the overall accuracy and precision of DIC's displacement measurements.

The ideal speckle pattern has a randomly-distributed arrangement of dots (speckles) with high contrast relative to the background, strong adhesion, and negligible stiffness relative to the sample [68]. Since DIC algorithms compute displacements in pixel space, the length scale of DIC is determined by the imaging system. Reported DIC measurements have spanned orders of magnitude in length, with speckle diameters ranging from individual atoms with high-resolution transmission electron microscopy [61], to nanometers and microns with scanning electron and optical microscopes [86, 87], to centimeters for magnetized plates on bridges [88], and to meters for topological features on volcanoes [60]. At the extreme length scales of nanometers and meters, painted speckle patterns generally do not provide properly sized speckles (about 3 x 3 pixels across) [68, 51]. However, most laboratory DIC exper-

iments use speckles with diameters of about 10 to 1,000 microns. At these typical length scales, speckle patterns of sprayed paint are expedient because of paint's relative compliance on many materials and the ease of applying a random arrangement of acceptably-sized dots. For airbrush painted speckle patterns, guidelines have been established for controlling speckle size and consistency. More consistent speckle sizes were found with diluted paint, increased spraying standoff distance, increased pressure, and decreased airflow. Larger speckles can be created with diluted paint, increased airflow, and decreased pressure [56].

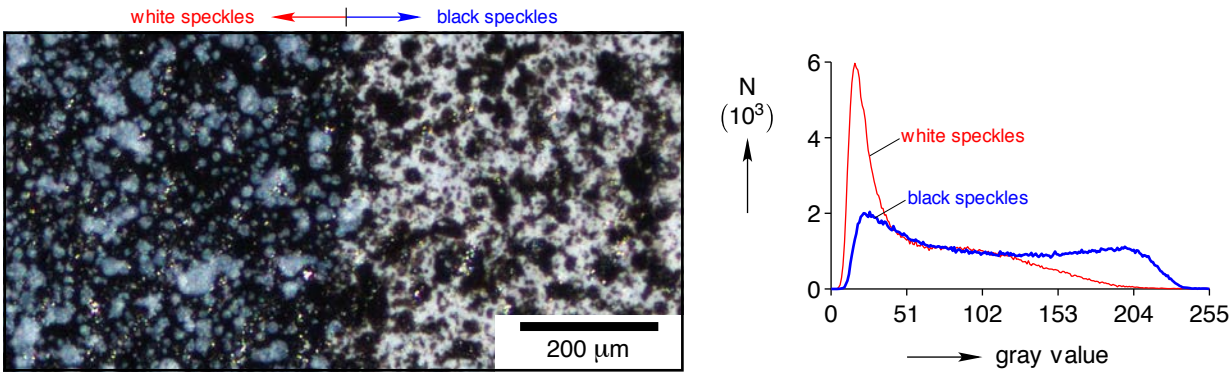
Occasionally, samples already have a uniformly dark or bright surface before painting that serves as the contrasting background, but for most samples, a basecoat is required to maximize contrast and reduce sources of glare such as shiny features on the sample surface. These shiny features can create saturated pixels, which prohibit sub-pixel accuracy and increase correlation error [68]. The basecoat should be thin and uniform to maintain good adhesion and compliance with the sample, and should fully cover the sample surface to avoid the glare sources, particularly with metal samples [70, 55].

For painted DIC speckle patterns, black and white paints are predominantly selected in order to maximize grayscale contrast, because higher contrast lowers correlation error by increasing image gradients [70]. Then the question arises: should the speckle pattern have a black basecoat with white speckles, or a white basecoat with black speckles?

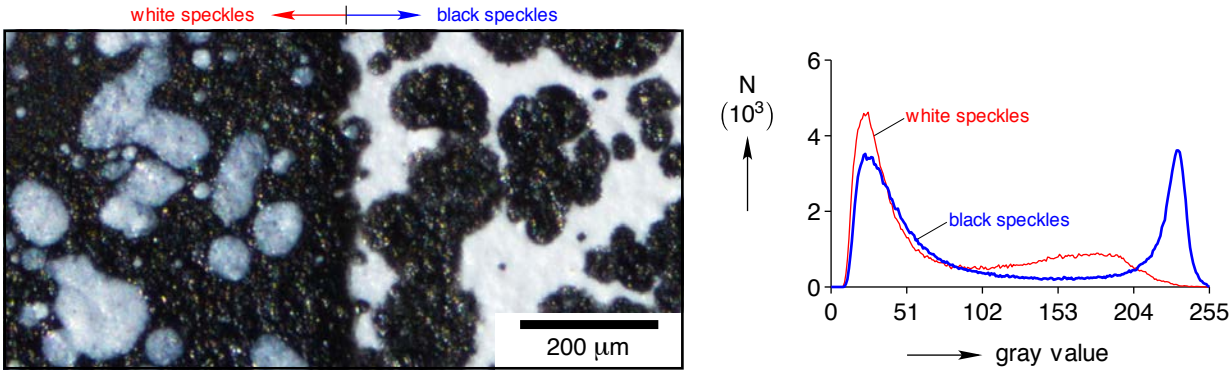
In this work we explore why black paint creates superior speckles over white paint. First, we present observations of paints without performing DIC analyses and discuss the relevant properties of common paints for DIC speckles. Next, we show DIC results of false displacements from static images in order to quantify the noise errors from each paint sequence, and then we conclude with recommendations.

3.2 Observations and properties of paints

On top of the basecoat or sample surface, contrasting dots (speckles) are applied to create a pattern of spatially unique areas of pixels, or "subsets," for correlation. Speckles should consistently be at least 3 x 3 pixels across to avoid aliasing [49], but not much larger than 3 x 3 pixels to maximize spatial precision [51]. Another consideration for painted speckles patterns is the elapsed time between painting and testing. Shorter times are preferred, from hours to a few days, especially for high-strain or high-rate experiments, because the paint becomes more brittle with prolonged drying [68, 55].



(a) airbrush



(b) spray can

Figure 3.1: Black speckles maintain high contrast on white basecoats, while white speckles have less contrast because white paint has lower *hiding power*, or the ability to obscure the underlying basecoat. There is also a slight contrast decrease (when viewed in grayscale images for DIC) from undertone, which is the subtle blue hue in the white speckles from Rayleigh scattering. Optical micrographs (in color) of adjacent speckle patterns were created by an airbrush (Figure 3.1a) and spray can (Figure 3.1b). The optical micrographs shown do not include cross polarization, so there are numerous bright spots due to glare. Cross polarization was used for the DIC images, however, to attenuate the glare [3]. Each image is a single color micrograph (not stitched together), with accurate white balance to show the blue hue from undertone. The grayscale histograms to the right of each image are computed from the micrograph areas that are shown.

Fundamental differences between black and white paints were probed by examining speckle patterns made by two common paint patterning methods, airbrush and spray can. With both samples, paint sequence was alternated in two adjacent regions by selectively masking with low-adhesion tape (ScotchBlue, 3M). This allowed the two paint sequences (black speckles on white basecoat, and white speckles on black basecoat) to be viewed simultaneously for accurate comparison (Figure 3.1).

For the airbrush sample, a 3 mm diameter metallic rod was selectively masked and painted

with thin basecoats of black and white paint (Golden High Flow Carbon Black, no. 8040-4, lot 345307, made with calcined natural iron oxide containing manganese; Golden High Flow Titanium White, no. 8549-4, lot 366162, containing titanium dioxide rutile [89]) using an Iwata CM-B airbrush with 0.18 mm fluid nozzle and needle, with an Iwata Studio Power Jet air compressor at 210 kPa (30 psi). Then speckles were applied with the same conditions using the opposite paint, again with selective masking to create two adjacent patterns with alternate paint sequence. For the spray can sample, a glass slide was speckled in the same selectively-masked method with Rustoleum Flat Protective Enamel black and Rustoleum Flat Protective Enamel white.

Optical micrographs of the samples created by airbrush and spray can methods reveal differences between black speckles and white speckles (Figure 3.1). Similar effects are seen for both patterning methods. Black speckles exhibit good *hiding power*, or the ability to obscure the underlying white basecoat, while white speckles have lower hiding power on the black basecoat. This is particularly true when the white paint is applied in a thin layer (shown later in Figure 3.2). Full opacity to cover the underlying black paint is not achieved because the white paint does not have sufficient hiding power. Later, we show that this incomplete opacity is also manifested by lower speckle densities (Figure 3.4). Black paint is known to have superior hiding power versus white paint, especially because of inherent differences between their origins of hiding power. Black paint creates hiding power by absorbing light, while white paint creates hiding power by scattering light. Light scattering is more difficult to achieve with paints because it requires pigment with a high refractive index and uniformly small particles with diameters slightly less than half the wavelength of light to be scattered (about 200 to 300 nm) [90, 91]. Much work since the 1930s has improved the hiding power of white paints, especially with the development of synthetic titanium dioxide pigments and coatings [92] and an ASTM standard to measure the hiding power of paint [93], but it still lags the hiding power of black paint.

In addition to incomplete opacity from lower hiding power, white speckles exhibit a small but noticeable effect called undertone that reduces contrast slightly. This undertone is evidenced by the bluish hue of the white speckles in Figure 3.1. Undertone arises from Rayleigh scattering, which is the interaction of various wavelengths of light with particles smaller than the wavelength of light. Blue light has a shorter wavelength and higher scattering efficiency, while red light has a longer wavelength and higher absorption, so the reflected light has a bluish hue. Due to a more prominent difference between the scattering of blue wavelengths and absorption of red wavelengths, undertone is more prevalent in white paints than in darker paints [90, 91].

Both the lower hiding power and undertone of the white speckles decrease their contrast

against the black basecoat. No bleeding of the black basecoat into the white speckles is suspected because there was no observable difference in optical micrographs of patterns with a ten minute drying time and patterns with a two day drying time.

3.3 DIC experiments and results

Eighteen speckle patterns were prepared on glass slides with the airbrush method outlined in Section 3.2 with varied speckle density (or coverage). A range of sparse to dense speckles were created by manually airbrushing a gradient of paint onto 9 glass slides for each paint sequence. The speckle patterns are presented in Figure 3.2, with increasing speckle density ρ_s going down the page. The 18 patterned slides were imaged with a 3-D DIC system comprised of two cameras (Point Grey GRAS 50S5MC with Nikon Micro Nikkor 200 mm f/4-30 lenses at f/11) and two LED light panels (LitePanels no. 935-1003), as shown in Figure 3.3. The patterned slides were held stationary, so any measured displacements are false (see Figure 3.5). These repeated measurements access the "type A" errors that define the measurement noise floor [94]. For DIC, type A errors include image noise, contrast, speckle quality, and air turbulence [64]. Cross polarization, or the use of orthogonal linear polarizing filters on the cameras and lights, was deployed to improve DIC quality. Cross polarization increases contrast and attenuates saturated pixels to decrease DIC error and enable smaller subsets [3]. Also, a fan mounted on a tripod separately from the DIC setup maintained turbulent airflow to avoid error from heat waves [95]. The LED light panels operated at full power, and the exposure times texp of the cameras were increased in 0.5 ms intervals just until no saturated pixels were observed with either camera. The exposure times were all relatively short, ranging from 11.5 to 17.5 ms. In general, more white paint in the speckle pattern (whether as the basecoat or as the speckles) reduced the exposure time for which no saturated pixels were observed (Figure 3.4).

step size	7 px
subset size	53 px
consistency threshold	0.02 px
maximum confidence interval	0.05 px
matchability threshold	0.1 px

Table 3.1: Correlation thresholds and settings.

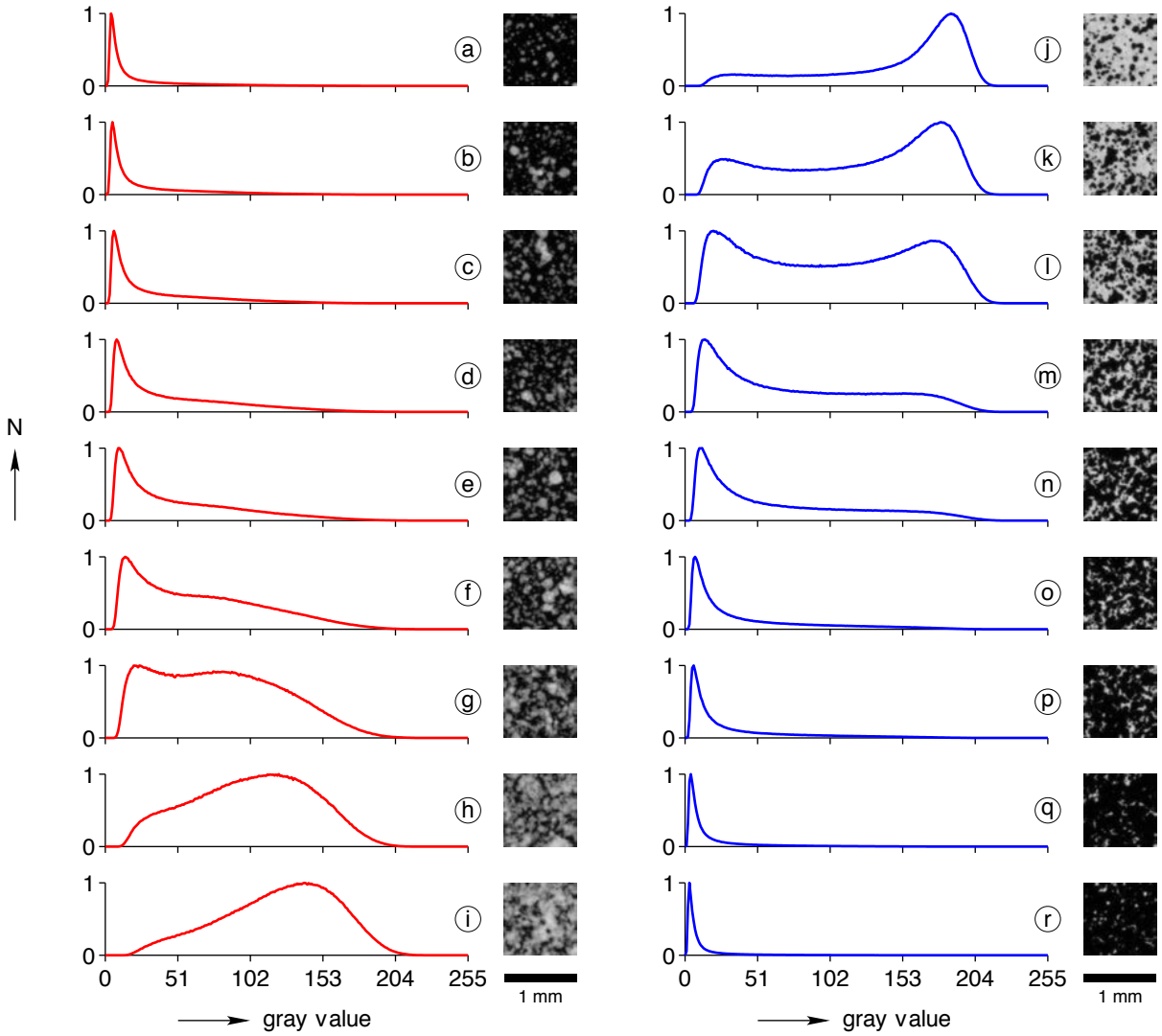


Figure 3.2: Black speckles on white paint (right column, lettered "j" through "r") achieved generally higher contrast, as indicated by wider histograms, than white speckles on black paint (left column, lettered "a" through "i"). The samples are arranged down the page with increasing speckle density (ρ_s), and each image tile shown is the center 1 x 1 mm region of the full 15 x 18 mm image from the left DIC camera. The histograms are computed from the full images. The histogram bins are one grayscale unit, and the frequencies (N) are normalized for each pattern.

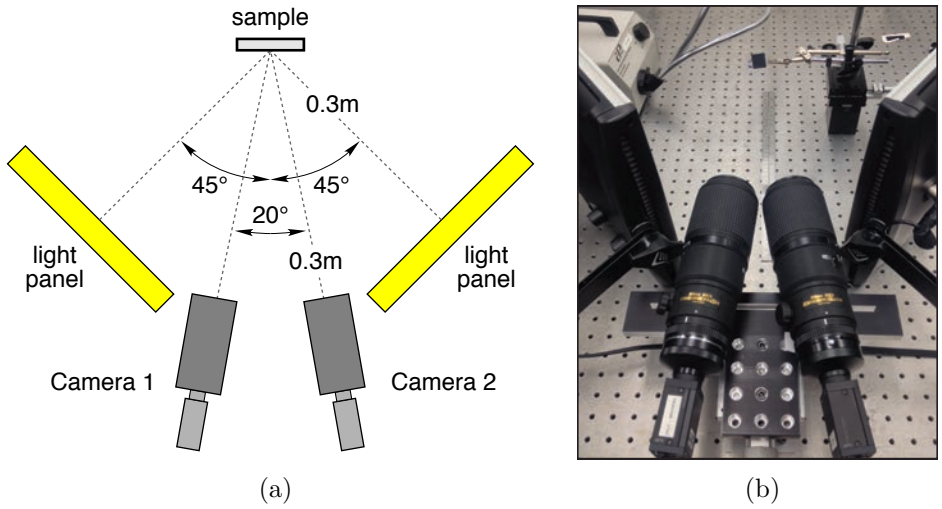


Figure 3.3: 3-D DIC measurements used two cameras at a 20° stereo angle with two LED light panels, shown in the schematic on the left (not to scale), and in the photo on the right. Cross polarization, or the orthogonal placement of linear polarizing filters on the cameras and lights, was deployed to improve DIC quality [3]. Also, a fan (not shown; on a tripod out of view) maintained turbulent air flow across the setup to avoid errors from heat waves [95].

DIC analysis was performed with Vic3D 7 software. The correlation threshold values are listed in Table 3.1. Correlation step size (the spacing among DIC data points) was 7 px for all cases, and the central 25 % of the area (7.5 mm × 9 mm) was chosen as the area of interest. This yielded about 25,000 3-D DIC data points for each speckle pattern. Vic3D software includes a "suggested subset size" feature that computes the subset size that yields a mean correlation confidence interval ($\overline{\text{CCI}}$) closest to 0.01 px. All of the suggested subset sizes (S$sugg$) are relatively large, ranging from 53 to 121 px (Figure 3.4), because most of the speckles are also relatively large (about 100 μm) compared to the pixel size (7 $\mu m/px$). For a fair comparison among speckle patterns of DIC metrics, such as correlation confidence intervals and false displacements, the suggested subset size was not chosen, but rather a subset size of 53 px was used for all of the DIC analyses. The correlation confidence interval (CCI) is a statistical measure of the correlation's accuracy for a subset's displacements from the reference configuration, so a smaller CCI indicates a better correlation. The smallest CCIs are observed for the patterns with black speckles (Figure 3.4).

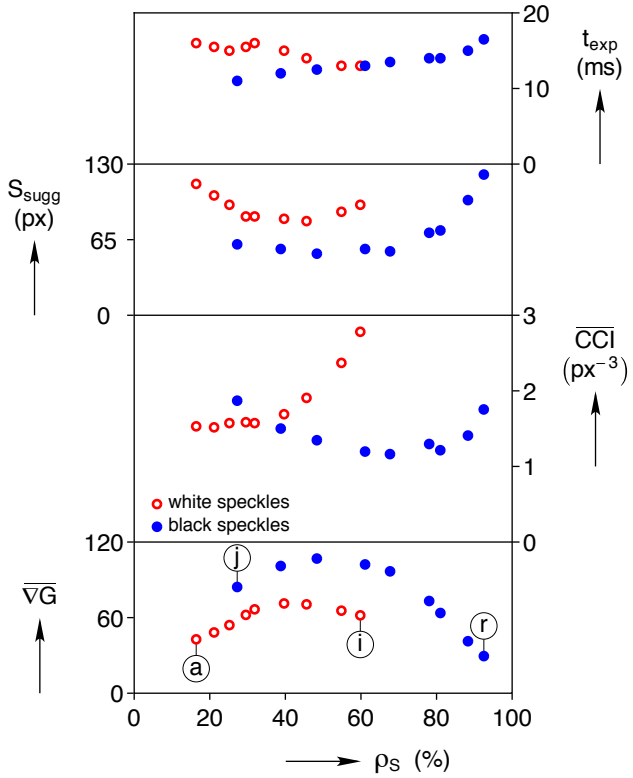


Figure 3.4: Exposure times (t_{exp}) were relatively short, and more white paint in the speckle pattern (either as speckles or the basecoat) generally decreased the exposure time. The patterns with black speckles were superior in three different metrics: smaller suggested subset sizes (S_{sugg}), smaller mean correlation confidence intervals ($\overline{CCI}$), and higher mean intensity gradients ($\overline{\nabla G}$). The patterns with the highest mean intensity gradients also had the broadest histograms (Figure 3.2). The increasing speckle density (ρs) in the positive horizontal direction of this figure corresponds to the patterns progressing down the page in Figure 3.2.

Speckle patterns with the highest (best) mean intensity gradients (Figure 3.4) had the broadest grayscale intensity distributions (Figure 3.2). The mean intensity gradient (denoted here as $\overline{\nabla G}$, but also commonly written as MIG) has been suggested as a good metric to evaluate the quality of DIC speckle patterns and images, because higher values of $\overline{\nabla G}$ indicate higher contrast spatially [81]. The mean intensity gradient is an image-wide parameter based on the local parameter of sum of square of subset intensity gradients (SSSIG). Recently a speckle pattern quality metric was developed based on SSSIG and the secondary auto-correlation peak height. SSSIG was demonstrated to be the best measure of speckle pattern quality, and was the primary component in the new speckle pattern quality metric. The new quality metric also included the secondary auto-correlation peak height because SSSIG alone provides high (favorable) values for repeated patterns such as checkerboards [59]. Repeated patterns can be problematic because spatially unique features are required for

subset uniqueness. For random patterns such as the painted patterns assessed here, though, the contrast measurements locally from SSSIG and thus globally from $\overline{\nabla G}$ are sufficient. For each of the 18 speckle patterns, the mean intensity gradient ($\overline{\nabla G}$) was computed with the Prewitt numerical differentiation filter [82], following Pan, Lu and Xie [81]. The patterns with the highest mean intensity gradients also had speckle densities ρ_s ranging between about 30 % and 70 %. The speckle densities ρ_s were computed with a grayscale thresholding filter, following Reu [51] and Lecompte, et al. [96]. An important consideration is that speckle density is coupled with speckle size: patterns with either too low or too high ρ_s will include both undersized and oversized speckles [50].

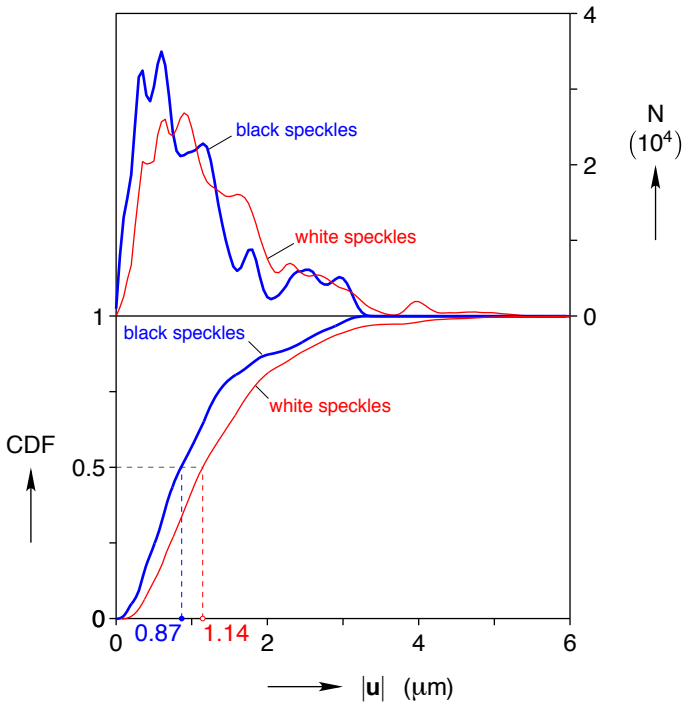


Figure 3.5: The patterns with black speckles had a 24 % lower median normalized false displacement (0.87 μm) than the patterns with white speckles (1.14 μm). False displacement magnitudes ($|\boldsymbol{u}|$) were binned in 0.05 μm intervals from the four patterns from each paint sequence with the highest mean intensity gradients (patterns e, f, g, and h with white speckles, and patterns k, l, m, and n for black speckles). The same false displacement data are replotted as an empirical cumulative distribution function (CDF), also indicating the lower false displacements for black speckles.

The median normalized false displacement for the four black-speckled patterns with the highest $\overline{\nabla G}$ (patterns e, f, g, and h) was 0.87 μm, 24 % lower than the 1.14 μm median normalized false displacement for the four white-speckled patterns with the highest $\overline{\nabla G}$ (patterns k, l, m, and n), shown in the top half of Section 3.3. The same data can also be visualized as an empirical cumulative distribution function, again indicating the lower false

displacements for the patterns with black speckles, shown in the bottom half of Section 3.3. The false displacement magnitudes are $|\boldsymbol{u}| = \sqrt{u^2 + v^2 + w^2}$, where u, v, and w are the horizontal, vertical, and out-of-plane components of displacement. Each speckle pattern was fixed in place while a series of 20 images were captured in one second intervals. Since each sample was stationary, any nonzero displacements are false and indicate the effective noise floor.

3.4 Conclusions

The optimum paint sequence for painted DIC speckle patterns is a basecoat of white paint covered by speckles of black paint. Fundamental differences between black and white paints, especially the higher hiding power of black paint and the undertone of white paint from Rayleigh scattering, enables black paint to maintain higher contrast in a thin layer as speckles. The improved contrast of patterns with black speckles is manifested as smaller suggested subset sizes, smaller correlation confidence intervals, higher mean intensity gradients, and ultimately more accurate displacements (24 % lower median normalized false displacement). In summary, the ideal painted speckle pattern has:

1. a thin basecoat of white paint, applied just thick enough to make complete coverage over the sample surface (especially for reflective samples such as shiny metals),

2. approximately equal amounts of coverage for the basecoat and speckles (speckle density $\rho_s \approx 50$ %), and

3. speckles of black paint with uniform size of at least 3 x 3 pixels to avoid aliasing, and random spatial positions for subset uniqueness.

CHAPTER 4

Residual stress and texture measurement of NiTi tube

4.1 Introduction

4.1.1 Texture measurements

Polycrystalline materials can exhibit *texture*, or preferred crystallographic orientation. Texture can impart anisotropy in the mechanical properties of the material [97]. Experimentally, texture is typically measured with diffraction techniques, although other methods can also measure texture, such as reflected light microscopes [98]. Diffraction measures crystal orientations, and three parameters define the orientation of a crystal with respect to the sample of material. Some diffraction techniques can measure all three parameters, such as electron backscatter diffraction (EBSD) [98]. EBSD was attempted on the NiTi material in this work, but the grains were likely too small ($\ll$ 1 µm) for reliable EBSD patterns. An alternative to EBSD is transmission Kikuchi diffraction (TKD), which can measure crystal orientations for smaller grain sizes than EBSD. However, previous studies with TKD on NiTi had relatively low confidence index values [99].

To overcome the limitation of poor EBSD patterns, X-ray diffraction (XRD) with a pole-figure goniometer can be used. XRD is not limited by very small grain sizes, such as those for the tube in this work. XRD with a pole-figure goniometer does not resolve crystal orientations with spatial positions like EBSD can, but rather averages over the sample surface because the X-ray illuminates a large number of grains simultaneously. Also, XRD cannot measure all three parameters that are necessary to define a crystal orientation. From XRD pole figures, an orientation distribution function, or the probability distribution function for crystal orientations in the material, can be computed [98].

4.1.2 Notation conventions

Miller indices indicate directions and planes in crystal systems. Miller indices for directions are enclosed by brackets, e.g. [110]. Miller indices for planes are enclosed by parenthesis, e.g. (110). For a cubic material, the direction $[uvw]$ is parallel with and equivalent to the plane normal with the same indices, (uvw). Thus, for cubic materials such as the B2 cubic austenite NiTi tube in this work, Miller indices could be written without enclosing punctuation, e.g. 110. In this work, we refer to textures based on the lattice plane normal, with enclosing parentheses.

Texture components in metals can be described as crystal orientations with Miller indices $(hkl)[uvw]$. The first three indices are the lattice plane normal (hkl) in the normal direction of sheet (hoop direction of a tube). The second three indices are the lattice direction $[uvw]$ in the rolling direction of a sheet (drawing direction of a tube) [98].

In addition to textures with a specific crystal orientation relative to the sample direction, there can be *fiber* textures, which are textures with radial symmetry about either a sample axis or another crystal coordinate. The fiber textures in this work are denoted as the crystal plane normal and the sample direction parallel to the fiber. For example, (223)//Z means the (223) plane normal with fiber texture symmetry parallel to the Z axis of the tube (axial/drawing direction).

4.1.3 Tube textures

Texture in tubes depends on the relative amount of reductions in tube diameter and wall thickness [100]. For proportional reductions in tube diameter and wall thickness, the tube's texture is typically similar to that of rod or wire. Reducing the wall thickness without significantly changing the diameter often produces texture similar to rolled sheet. Reducing the tube diameter without significantly changing the wall thickness ("sinking") often gives texture tangential to the circumference [100, 101]. Also, tubes can retain the drawing texture after annealing. For example, copper and brass tubes with proportional wall thickness and tube diameter reductions retain (111) and (100) duplex fiber texture after annealing [100]. Annealing can also modify tube texture, such as the shift upon annealing from $[10\bar{1}0]$ to $[11\bar{2}0]$ texture in the drawing axis for cold-formed Zr or Zircaloy tubes [100].

4.1.4 Orientation distribution functions

Texture is expressed as the probability distribution of crystal orientations, or an orientation distribution function (ODF). ODFs have three dimensions and can be projected into two dimensions as pole figures or inverse pole figures. Pole figures (PFs) are two-dimensional pro-

jections of crystal direction relative to sample coordinates, written as a certain normal/pole to a lattice plane, (hkl). The units of pole figures are normalized to express the probability as multiples of uniform distribution (MUD; also called multiples of random distribution, MRD), with MUD = 1 indicating uniform texture. Inverse pole figures (IPFs) represent the probability of finding a sample direction relative to certain crystal directions. IPFs are helpful to identify axisymmetric textures ("fiber textures") for which one sample direction (e.g. radial direction of a tube) is applicable [98].

Various methods have been developed to estimate ODFs from pole figures, since techniques such as XRD with a goniometer can measure pole figures, but not ODFs [98]. This work calculated the tube's ODF with the MTEX open source Matlab toolbox [102]. In MTEX, the ODF is estimated by minimizing radially symmetric functions [103].

In addition to PFs and IPFs, ODFs are commonly visualized as two-dimensional sections within a space of three Euler angles that relate the two orthogonal, right-handed coordinate systems of the sample and the crystal.

There are different conventions for Euler angles, and the Bunge convention is most common in the metals community. However, Cartesian Euler space is distorted and contains three times as much information as needed, so ODFs represented as Euler angles need to be interpreted cautiously [104, 105]. Cylindrical representations of Euler angles have been proposed to improve their visualization [106, 107], as well as special sections with minimal distortions [108, 109].

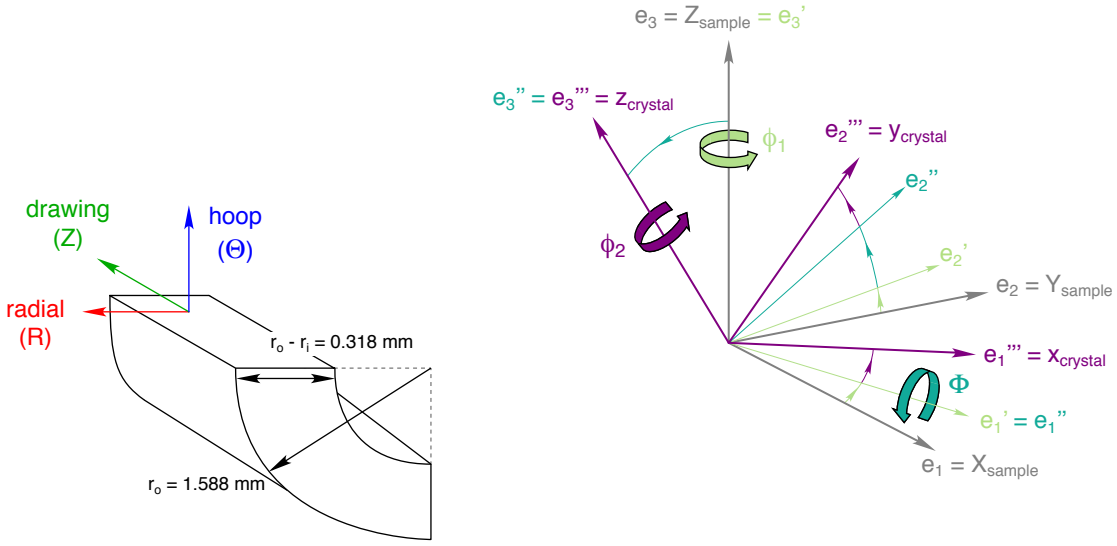
When viewing an ODF as two-dimensional sections (such as the constant ϕ_2 sections in Figure 4.12), only one "zone" should be considered when counting different texture components so that the same texture is not considered more than once [104]. However, texture components can be calculated with the appropriate symmetries and viewed in three dimensions to avoid the issues associated with viewing ODFs in only two-dimensional slices. ODFs can also be viewed in three-dimensional space with modern software such as the MTEX package used in this work (example video available at https://goo.gl/AY6MBE).

In MTEX, the default Euler angle convention is that of Bunge[1]. The Bunge Euler angles are defined by three rotations of the crystal coordinates (x_{crystal}, y_{crystal}, and z_{crystal}) with respect to three sample coordinates (X_{sample}, Y_{sample}, Z_{sample}), as shown in Figure 4.1b. The three Bungle Euler angle rotations in MTEX's default "ZXZ" convention are:

1. ϕ_1: rotate crystal about $e_3 = Z_{\text{sample}}$,
2. Φ: rotate crystal about e_1' (the rotated X_{sample}), and
3. ϕ_2: rotate crystal about e_3'' (the rotated Z_{sample}).

[1] https://mtex-toolbox.github.io/files/doc/Rotations.html

The crystal coordinates correspond to the unit directions in Miller indices (x_{crystal} is [100], y_{crystal} is [010], and z_{crystal} is [001]). The sample coordinates depend on the material geometry. For the most common sample geometry of rolled sheet, X_{sample} is the rolling direction (RD), Y_{sample} is the transverse direction (TD), and Z_{sample} is the normal direction (ND). For the tube sample coordinates in the present work, however, X_{sample} is the axial/drawing direction (Z_{tube}), Y_{sample} is the radial direction (R_{tube}), and Z_{sample} is the hoop direction (Θ_{tube}). Although this incongruent labeling is confusing, it provides the necessary coordinates for matching common texture components in the Bunge Euler angle convention, such as the γ fiber observed in this work.



(a) (b)

Figure 4.1: (a) The tube coordinate axes. (b) A diagram of the "ZXZ" Bunge Euler angle convention used for the orientation distribution function, adapted from a diagram in the lecture notes by Prof. Anthony Rollett. X_{sample} is the axial/drawing direction (Z_{tube}), Y_{sample} is the radial direction (R_{tube}), and Z_{sample} is the hoop direction (Θ_{tube}).

Without any sample or material symmetry, the possible ranges for Euler angles are $0 \leq \phi_1 < 360^{\text{o}}$, $0 \leq \Phi < 180^{\text{o}}$, and $0 \leq \phi_2 < 360^{\text{o}}$. The Euler angle ranges are reduced by B2 NiTi's cubic symmetry (*m-3m* Laue group, or *432* Hermann–Mauguin/international notation). Also, orthorhombic sample symmetry is assumed for the tube drawing process (*mmm* Laue group, or *222* Hermann–Mauguin/international notation). With these crystal and sample symmetries, the ranges reduce to $0 \leq \phi_1 < 90^{\text{o}}$, $0 \leq \Phi < 90^{\text{o}}$, and $0 \leq \phi_2 < 90^{\text{o}}$.

4.1.5 Texture of NiTi tubes

Polycrystalline NiTi shape memory alloy (SMA) can have strong texture that forms during deformation, including drawing and rolling processes. The development of texture in NiTi was experimentally observed from neutron diffraction measurements with in-situ compression. B2 NiTi exhibited (100) texture that aligned parallel to the compression axis, and (110) texture that aligned perpendicular to the compression axis [110].

Texture has been observed in NiTi with various product forms (e.g. wire, sheet, tube), and the texture varies among product forms [111, 112, 113]. Although tube is currently the most commercially relevant product form of NiTi (for endovascular devices such as stents), relatively few studies have characterized the texture of NiTi tubes. Attempts have been made to produce stent tubes with low texture for quasi-isotropic properties [114], but generally, NiTi tubes have strong texture from manufacturing processes.

The texture of NiTi tube has been measured for a selection of heat treatments [111, 112, 113]. Three heat treatments were studied: (1) 485 °C for 5 minutes, followed by an ice-water quench (a common condition in the Nitinol industry for stress relieving, slightly lowering tensile strength, and slightly raising the austenite finish temperature); (2) 850 °C for 30 minutes in vacuum, followed by a slow cool of −10 °C per minute (a typical industry anneal); and (3) 1100 °C for 12 hours in vacuum, followed by ice-water quench (removes residual stress, recrystallizes the material, and removes the R-phase) [113]. For all three heat treatments, the texture components of the tube were generally the same, although there was some redistribution of the texture. Interestingly, the heat treatments at higher temperatures and longer durations *increased* the amount of texture. In the radial direction, there was a progressive redistribution from (111) to (110) texture. In the drawing direction, there was redistribution from (111) and (110) textures to only (111) texture. For all three heat treatments, the drawing direction's texture was stronger than that of the radial and hoop directions [113].

Synchrotron X-ray microdiffraction experiments reiterated NiTi tube's strong texture, and in-situ tension results revealed the macroscopic mechanical responses and microscopic inhomogeneities with respect to grain orientations [115]. Tensile dogbone samples were laser machined from NiTi tube sections that had been flatted with heat treatment (a two-step process at 500 °C). Dogbones were produced with three different tensile axis directions relative to the tube drawing direction: 0° (longitudinal), 45°, and 90° (circumferential). The grain sizes were grown to 10-20 µm with heat treatment (700 °C for 5 minutes) to allow the X-ray (spot size of 1 µm) to resolve transformation within each grain. The 0° sample had primarily ⟨111⟩ grain orientation. The 90° sample had primarily ⟨110⟩ grain orientation. The 45° sample had primarily ⟨100⟩ grain orientation. Microscopically, the difficult-to-

transform ⟨100⟩ grains often remained as islands of austenite, with transformation occurring around bands of ⟨100⟩ grains. The microscopic observations, especially with respect to the ⟨100⟩ grains that tend to resist transformation, connected with macroscopic mechanical measurements [115].

Severely cold-rolled NiTi sheet (42 % thickness reduction) showed similar (110) texture in the rolling direction after two relatively brief, low-temperature heat treatments (485 °C for 2 minutes and 520 °C for 6 minutes, both followed by water quenching [116].

In-situ EBSD studies have been performed on NiTi tubes. To enable EBSD analysis, the NiTi tubes were annealed (800 °C for 30 minutes, then water quenched) to increase the grain size (reported as 10 to 100 µm). By inspection of the ODF in the $\phi_2 = 45^o$ section (Bunge convention), the main texture components were $(110)[1\bar{1}1]$ and $(112)[\bar{1}\bar{1}1]$. From an IPF in the axial tube direction, there was a strong (111) texture [117].

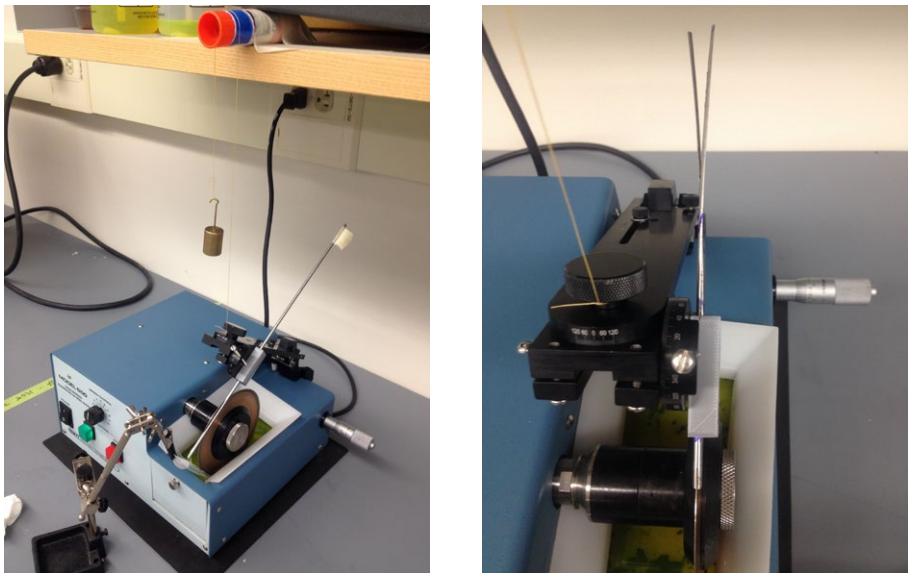
Recently, texture in NiTi tubes was measured with XRD and precession electron diffraction automated crystal orientation mapping (PED-ACOM) performed using a transmission electron microscope (TEM) [118]. Samples for the PED-ACOM were prepared with a focused ion beam scanning electron microscope (FIB-SEM). Samples for XRD were prepared by polishing along the tube axis to produce a surface with a normal in the tube's radial direction (although the surface would also include normal vector components in the hoop direction). From both the XRD and PED-ACOM measurements, the tube's texture was characterized as strong (110) texture in the radial direction. The (110) texture after relatively aggressive heat treatment from [118] is consistent with the progressive redistribution in radial texture from (111) to (110) upon heat treatment shown in [113].

4.2 Methods

4.2.1 Sample preparation

First, an XRD measurement for defocusing correction was performed on B2 NiTi powder (Special Metals, New Hartford, New York; lot 3601, $Ni_{51}Ti_{49}$) using the same XRD conditions that were used for the tube XRD measurements. For a flat sample with XRD in reflection geometry (as opposed to transmission geometry for thin samples), the inclination of the sample with respect to the beam creates defocusing effects, so defocusing corrections are necessary [98].

XRD measurements were performed on NiTi tube (Memry, Bethel, Connecticut; lot 957152, product number 10130W) with outer diameter of 3.176 mm and inner diameter of 2.540 mm. The XRD sample was prepared following the method from [111, 112, 113], with sections from the tube stacked in the radial direction. This stacked configuration increases the sample area



(a) (b)

Figure 4.2: Low-speed diamond sawing. (a) The first axial cut made half sections along the tube axis. (b) During the second axial cut to make quarter sections, bending from residual stress was observed, with the two quarter sections curving enough to overlap.

to improve XRD signal-to-noise ratio, and is similar to the method outlined for wires in [119]. Unfortunately, this stacking enforces orthorhombic symmetry on the sample that assumes there is no texture difference in the radial direction (e.g. no significant texture differences between material points in the middle or near the inner and outer walls of the tube). EBSD measurements on NiTi sheet have shown that there can be through-thickness texture variation for hot-rolled sheet, but not cold-rolled sheet [120]. Hot-rolled NiTi sheet had two primary textures in the "warm inner zone" ($\{111\}\langle 110\rangle$ and $\{110\}\langle 110\rangle$ texture), but random texture near the surface. The inner grains were also larger. The differences between the inner and surface regions was attributed to slip and recrystallization in the inner region, aided by more heat, while the cooler surface had less slip and more dynamic recovery. Unlike the hot-rolled sheet, cold rolling followed by recrystallization did not exhibit texture or grain size differences between the inner and outer regions of the sheet. Cold-rolled and recrystallized sheet had homogeneous grain size and texture through the thickness, random texture in the rolling and transverse direction, and $\{111\}\langle uvw\rangle$ γ-fiber texture in the normal direction [120].

Sections were cut from the tube using a low-speed diamond saw with a cooling oil bath. The tube was held in the diamond saw with a custom fixture and cut twice in the axial direction to produce quarter slices (Figure 4.2a). A counterweight was used to reduce the cutting rate to less than 0.5 cm of the tube length per hour. After the first axial cut, the tube

half sections spontaneously bent towards one another, but only slightly. After the second axial cut, the tube quarter sections bent towards one another markedly (Figure 4.2b).

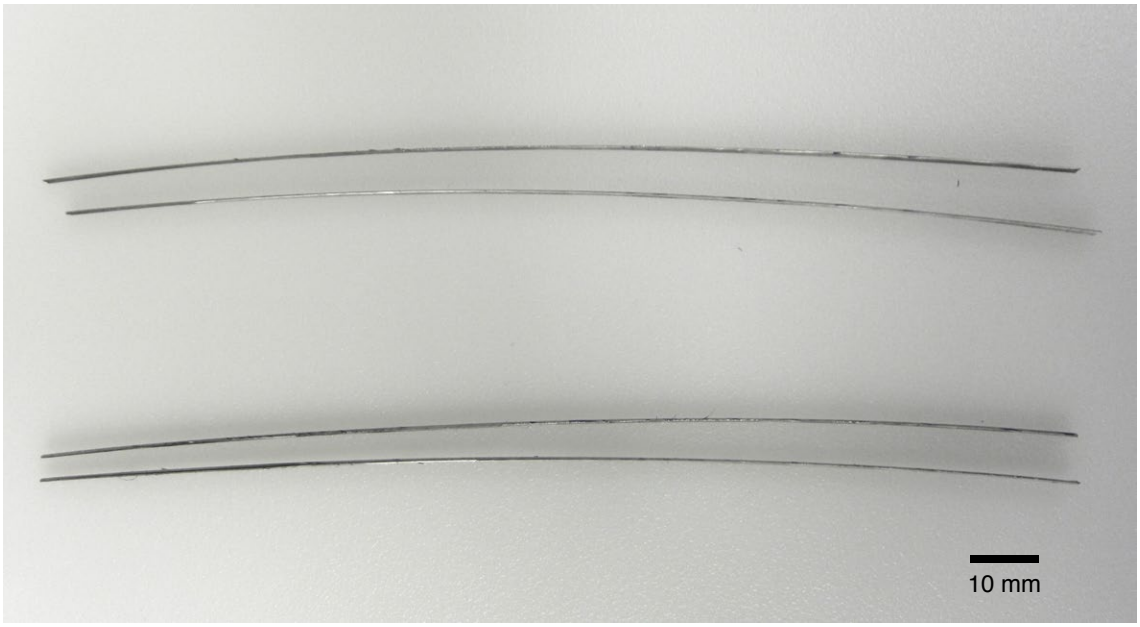
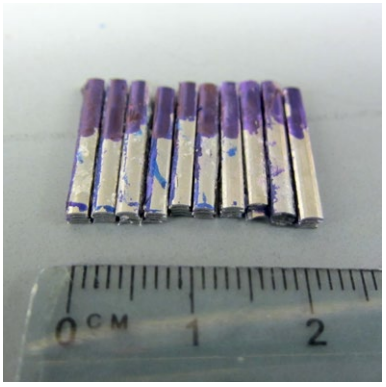


Figure 4.3: Residual stress was observed after tube sectioning.

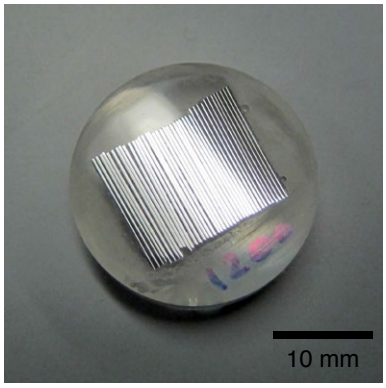
After the axial cuts, the radius of curvature of the quarter sections was 740 ± 100 mm (Figure 4.3). Assuming Euler–Bernoulli theory and radially distributed residual stress, the magnitude of residual stress to produce this bend is -29 MPa at the outer surface and 30 MPa at the inner surface (see residual stress analysis in Section 4.3.2).

The quarter sections of the tube were cut again with the low-speed diamond saw to produce forty 15 mm pieces from the 155 mm quarter slices (Figure 4.4a). These quarter section segments were stacked in the geometry shown in Figure 4.5a and glued with cyanoacrylate "super glue" to remain in this configuration during mounting. The tubes were mounted with an epoxy that cures in about two hours at room temperature (PELCO 2 Hour Epoxy; Ted Pella, Redding, CA). The mounting produced a cylindrical sample 25 mm in diameter (Figure 4.4b). The sample was ground flat with 600 grit silica paper wetted with water, and ground smoother with 800, 1000, and 1200 grit silica wetted papers (CarbiMet S and MicroCut S, Buehler). Grinding was performed until the centerline of the quarter section of tube slices was the exposed surface (Figure 4.5b). The sample was final polished with a 5:1 volume ratio mixture of colloidal silica (Buehler MasterMet 2, 0.02 µm) and 30 % hydrogen peroxide (Fisher H325) on a microfiber cloth (Buehler MicroCut), to produce a samples with a mirror finish on all of the forty quarter sections of tube.



(a) The forty quarter sections of tube (in ten stacks of four) before mounting.

(b) The forty quarter sections of tube stacked radially as shown in Figure 4.5.



(c) The sample after grinding with 1200 grit silica paper.

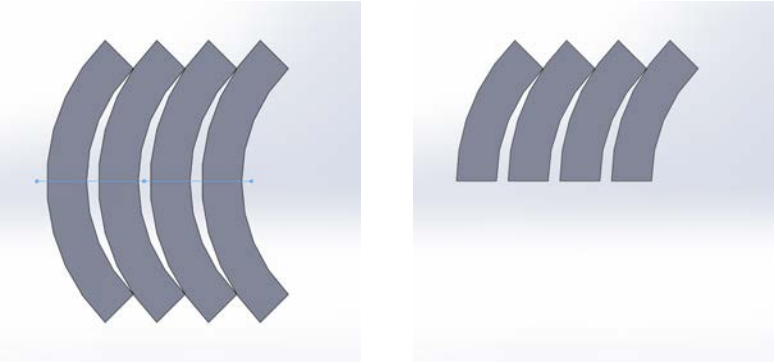
Figure 4.4: Assembly and polishing of tube stack.

4.2.2 X-ray diffraction

X-ray diffraction (XRD) was performed with a Rigaku SmartLab system. The sample was loaded on the "thin film" type stage with the ability to rotate in the plane of the sample surface. The sample was fixed to the stage with tape to avoid any sample motion during the experiment (Figure 4.6). The sample was loaded with the tube drawing direction pointing to $\beta = 0^o$ (towards the instrument's front door, which is the instrument's convention for a point detector). The XRD conditions for all experiments were 40 kV and 44 mA source, parallel beam (PB) collimator slit, 0.5° in-plane parallel slit collimator (PSC), 10 mm length limiting slit, and 0.5° in-plane parallel slit analyzer (PSA). The attenuator was open, the divergence slit was 1 mm, and the Soller slit and receiving slit were 2 mm. The maximum X-ray penetration in NiTi at the experiment conditions was calculated to be 19 µm (Section 4.3.3).

First, a symmetric scan ($2\theta/\theta$ scan) was performed to locate the austenite peaks in the tube sample (Figure 4.7). A Cu-Kβ filter (a thin film of Ni) was not used during the symmetric scan. The scan was performed for 2θ from 30° to 85° with 0.050° and 1 second per step. The symmetric scan showed the expected peaks and intensities for B2 austenite NiTi, with a prominent (110) peak at 42.4°, the (200) peak at 61.6°, and the (211) peak at 77.6° (Figure 4.7).

Next, in-plane pole figure measurement of preferred crystallographic orientation were performed. The in-plane method was chosen over other methods such as the Schulz reflection method and the parallel beam reflection method because the in-plane method can measure the whole pole figure without using transmission geometry that requires a very thin sample [121]. The in-plane method accesses the polar (α) and azimuthal (β) angles of a pole



(a) The quarter sections were stacked radially with their centers aligned.

(b) The sample was ground and polished down to the center line of the stack of quarter sections.

Figure 4.5: CAD of tube stack polishing, with only four of the forty section shown here.

figure by rotating the sample about its plane and sweeping the X-ray source and detector along separate arcs. β is equal to the sample rotation angle, ϕ. α is the angle between the sample surface and the plane defined by the detector position, sample incident position, and X-ray source position (Figure 4.8).

A Cu-Kβ filter (a thin film of Ni) was installed on the detector-side of the instrument for the in-plane pole figure measurements. The same slit settings from the symmetric scan were used. Pole figures were measured for the (110), (200), and (211) peaks with the full β range (azimuthal angle of 0 to 360° with 1° continuous sampling steps and β rotation speed of 60° per minute) and a partial α range (polar angle of 0° to 75° with 2.5° incremental sampling steps). The α range was limited since polar angles greater than about 80° is not reliable [119].

4.3 Results and discussion

4.3.1 Pole figure measurements and orientation distribution function calculation

The three pole figures that were measured with XRD (Figure 4.9a) indicate texture is present in the tube. From the pole figures, a moderate (110) texture was observed in the tube drawing direction (Z). For the tube radial direction (R) in the (110) pole figure, there is a band of relatively uniform intensity with MUD > 1 that indicates less texture in the

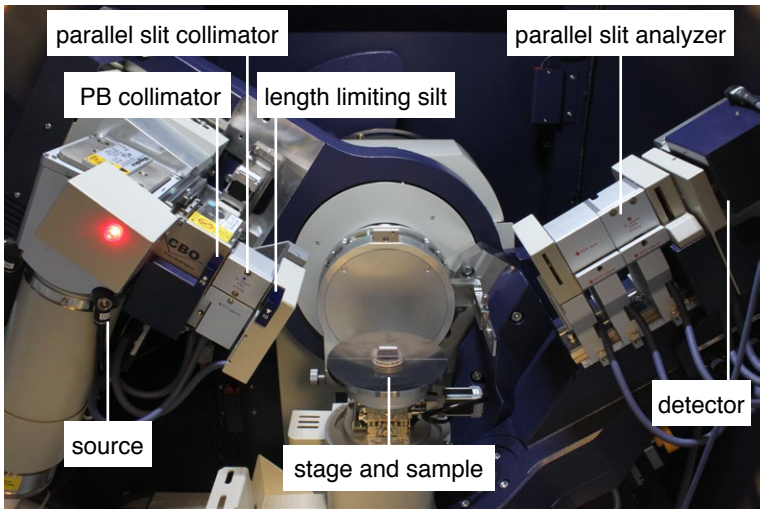


Figure 4.6: Sample in XRD instrument during a symmetric $\theta/2\theta$ experiment.

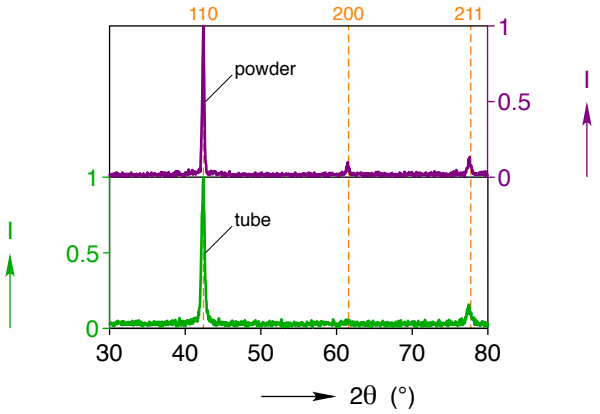


Figure 4.7: XRD profile from $2\theta/\theta$ symmetric scan for the powder sample (top) and tube sample (bottom). The tube sample was positioned with the radial direction parallel to the beam direction.

radial direction than the drawing direction. For the (200) pole figure, the radial direction also has a band of relatively uniform intensity with MUD < 1. From the three measured pole figures, an ODF was calculated using MTEX [102]. Pole figures produced from the calculated ODFs (Figure 4.9b) matched the general shapes and magnitudes of the measured PFs. Generally, the error between the measured and calculated PFs was less than 0.1 |MUD| err, with some spurious signals that had about 0.3 |MUD| err (Figure 4.9c).

From the ODF, inverse pole figures were generated for the tube directions (Figure 4.10). The primary texture was (111) in the drawing direction (Z). Other minor textures included (111) and (122) in the hoop direction (Θ), and (114) in the radial direction (R).

The three poles from the cubic inverse pole figures, (111), (100), and (110), were plotted as pole figures projected in the radial-drawing (R-Z) plane, with the hoop direction coming

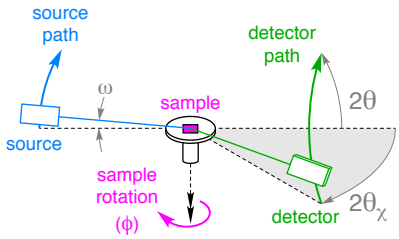


Figure 4.8: The in-plane pole figure measurement method, with the source swept in ω, the detector swept along an arc defined by 2θ and $2\theta_\chi$, and the sample rotated about ϕ. ω and 2θ rotate in the plane of the page, and Φ and $2\theta_\chi$ rotate about the sample normal.

out of the center (Figure 4.11). In this representation, the relative magnitudes of the texture is also apparent, with a strong (111) texture in the drawing direction.

Following the convention for representing textures in BCC metals, the ODF is plotted in Figure 4.12 as ϕ_2 sections (with the Bunge Euler angle convention). From the $\phi_2 = 45^\circ$ section, three texture components were identified (labeled on the plot):

1. (223)//Z fiber,
2. (111)//Z fiber, and
3. (111)//Θ fiber (commonly called the γ fiber for BCC metals).[2]

The (122)//Θ fiber is also labeled on Figure 4.12 to compare with the (111)//Θ fiber. Although the intensity at (122) was slightly greater than the intensity at (111) in the hoop direction (Θ) IPF (Figure 4.10), the intersection of the (122)//Θ fiber with the (223)//Z fiber could increase the apparent intensity of the (122)//Θ fiber. The intersection of the (122)//Θ and (223)//Z fibers is more apparent when the ODF is visualized in three dimensions. A video of the 3D representation of the ODF is available at https://goo.gl/AY6MBE.

The volume fraction η of the fibers was computed as a function of the angle radius around the fiber (Figure 4.13), using the "fibreVolume" function in MTEX. Consistent with the IPF observations (Figure 4.10), the strongest fiber was the (111)//Z. The (223)//Z fiber was the second strongest, followed by the (111)//Θ and (122)//Θ fibers.

[2]The Bunge Euler angles for the γ fiber are $\phi_1, \Phi, \phi_2 = 0\text{-}90^\circ, 55^\circ, 45^\circ$.

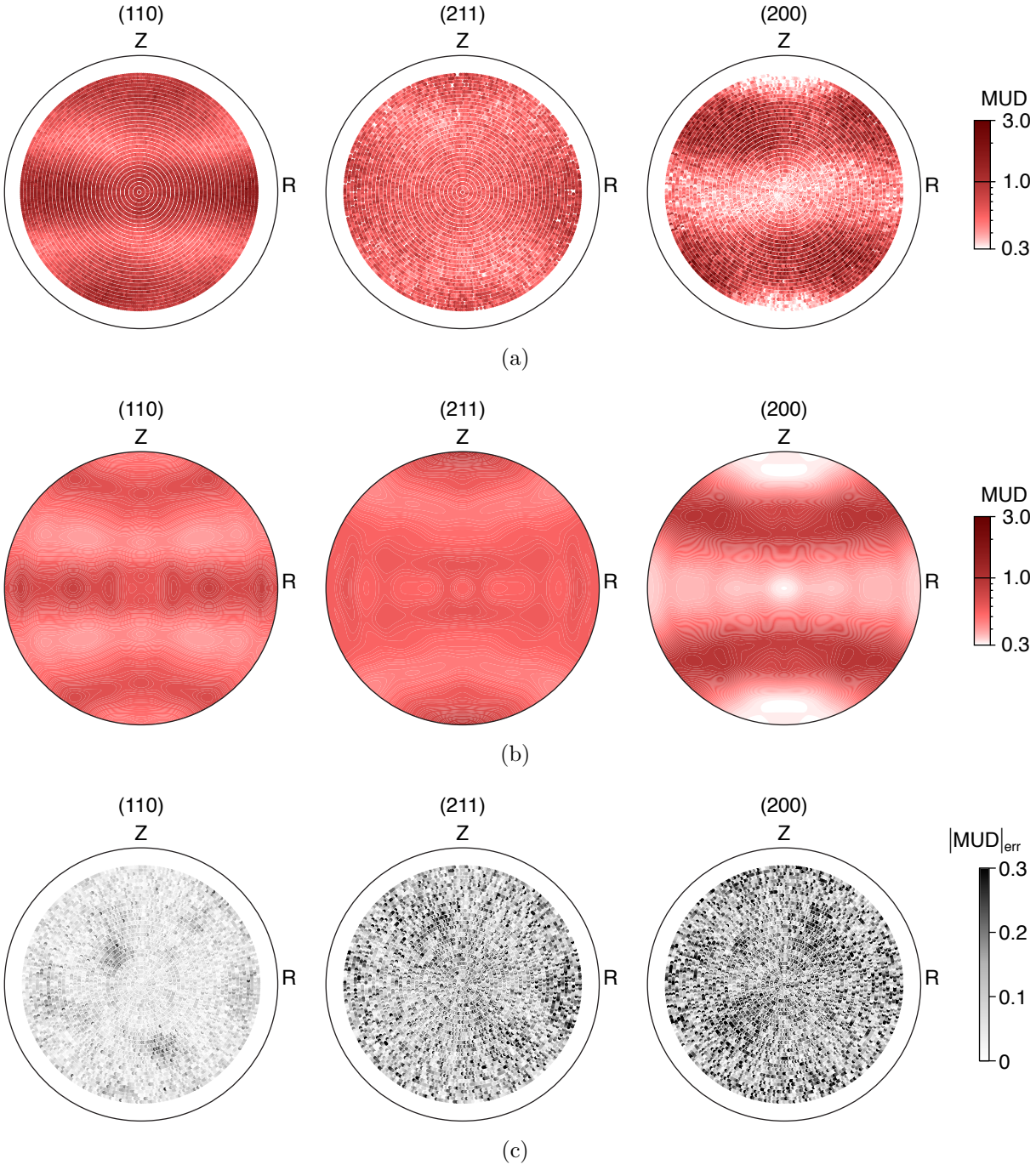


Figure 4.9: Pole figures of (110), (211), and (200) planes, as-measured (a), calculated (b), and the difference between measurement and calculation (c). All units are multiples of random distribution (MRD). The measured values (a) are normalized with respect to the orientation distribution function.

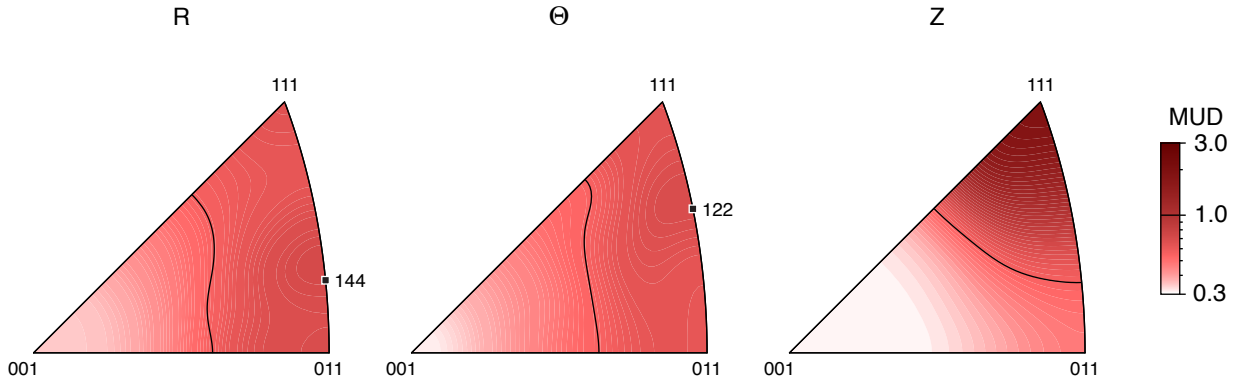


Figure 4.10: Inverse PFs from ODF for R, Θ, and Z sample directions.

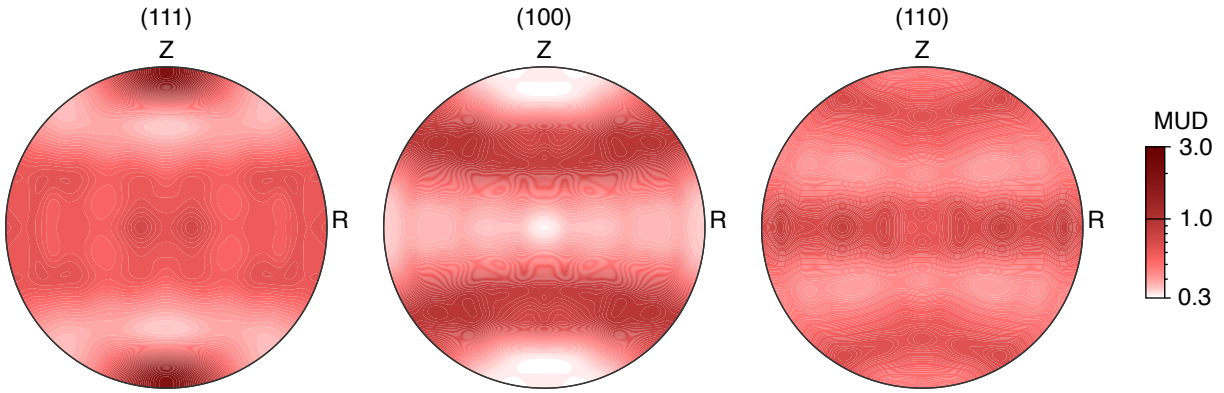


Figure 4.11: Pole figures plotted from the calculated ODF in IPF's three corner poles, (111), (100), and (110).

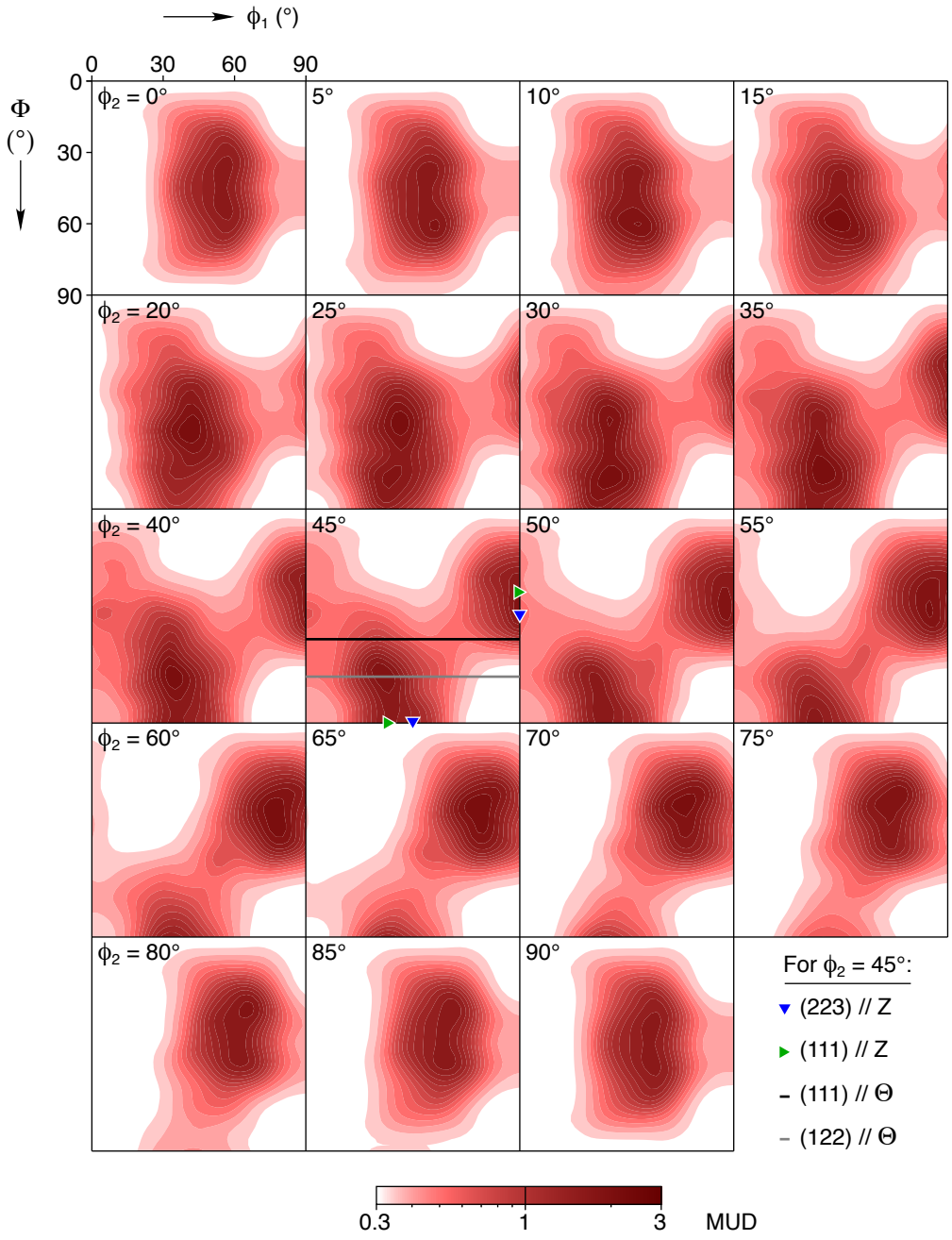


Figure 4.12: Bunge convention Euler angles ϕ_1 and Φ with slices of constant ϕ_2 in 5° increments.

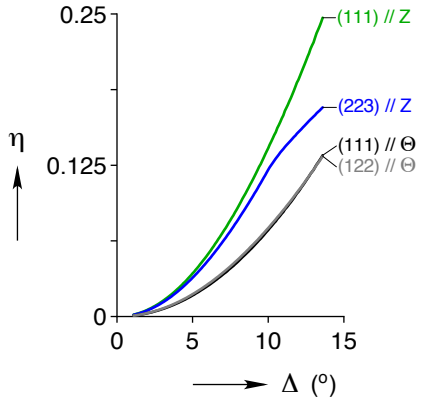


Figure 4.13: Volume fractions (η) of the three primary texture components for radius (Δ) around the fibers. The (111)//Z fiber had the highest volume fraction, followed by the (223)//Z fiber.

4.3.2 Residual stress analysis

Upon cutting the tube with a low-speed diamond saw, bending was observed in the tube quarter sections from residual stress. From this beam theory analysis, residual stress in the tube would result from bending the tube quarter section to a curvature of zero for the as-received, straight tube. The geometry of the tube cross-section yielded the beam's neutral axis (Section 4.3.2.1). From the radius of curvature of the bent tube quarter sections, the residual stress to cause bending, σZZ was estimated with elementary beam bending theory with a stress distribution proportional to distance from the neutral axis (Section 4.3.2.2). Finally, this stress distribution was equated to a stress distribution proportional to distance from a radial axis (Section 4.3.2.3).

4.3.2.1 Tube quarter section centroid

The centroid of the tube quarter section was computed by considering the tube cross-section as a simply connected domain (Figure 4.14a). The distance from the X-axis to the centroid of a quarter section of a circle with inner radius $r\mathrm{i}$ and outer radius $r\mathrm{o}$, with $\hat{r} = (r\mathrm{i} + r\mathrm{o})/2$, is given by

$$\bar{y} = \frac{\int \hat{y}\,\hat{r}\,d\theta}{\int \hat{r}\,d\theta} = \frac{\int (\hat{r}\sin\theta)\,\hat{r}\,d\theta}{\int \hat{r}\,d\theta} = \frac{\int \hat{r}^2 \sin\theta\,d\theta}{\int \hat{r}\,d\theta}. \tag{4.1}$$

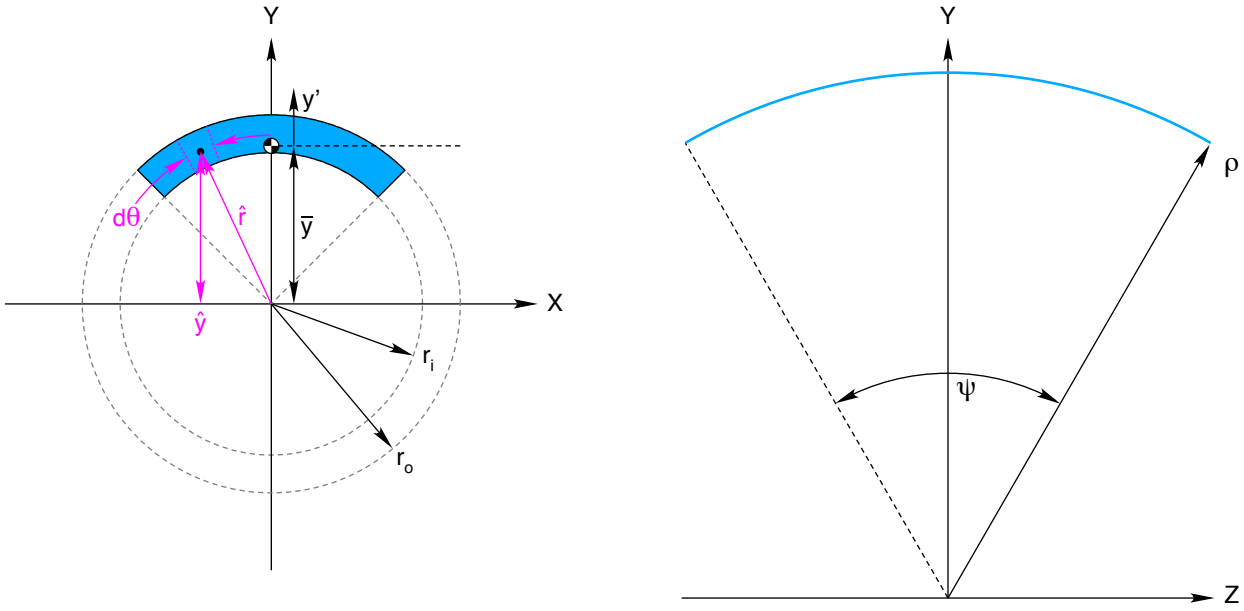
For the limits of integration as θ from $\pi/4 + \theta_{\mathrm{kerf}}/2$ to $3\pi/4 - \theta_{\mathrm{kerf}}/2$, where θ_{kerf} is the angular width removed by the diamond saw blade, this gives

$$\bar{y} = \frac{4\hat{r}\cos\left(\pi/4 + \theta_{\mathrm{kerf}}/2\right)}{\pi - 2\theta_{\mathrm{kerf}}}. \tag{4.2}$$

For the tube in the present work, $r_{\mathrm{o}} = 1.588$ mm, $r_{\mathrm{i}} = 1.270$ mm, and $\theta_{\mathrm{kerf}} = 0.2519$ radians. This yields $\bar{y} = 1.328$ mm.

4.3.2.2 Elementary beam bending analysis

First, the beam bending solution was computed for a residual stress distribution that is proportional to y', the vertical distance from the neutral axis $\bar{y}$ (Figure 4.15a). Assuming that the quarter sections of tube obey Euler–Bernoulli beam theory, then the length of the neutral axis for a beam that is bent with radius of curvature ρ and included angle ψ is $\rho\,d\psi$ (Figure 4.14b). The strain at distance y' from the neutral axis along a length of the beam with measured curvature $\kappa_0 = 1/\rho$ and current curvature in the as-received (straight) tube



(a) Differential element to compute centroid

(b) Radius of curvature ρ and included angle ψ

Figure 4.14: Quarter section moment of inertia about X-axis.

$\kappa = 0$ is given by

$$\varepsilon_{ZZ} = (\kappa - \kappa_0)\, y' = -\kappa y'. \tag{4.3}$$

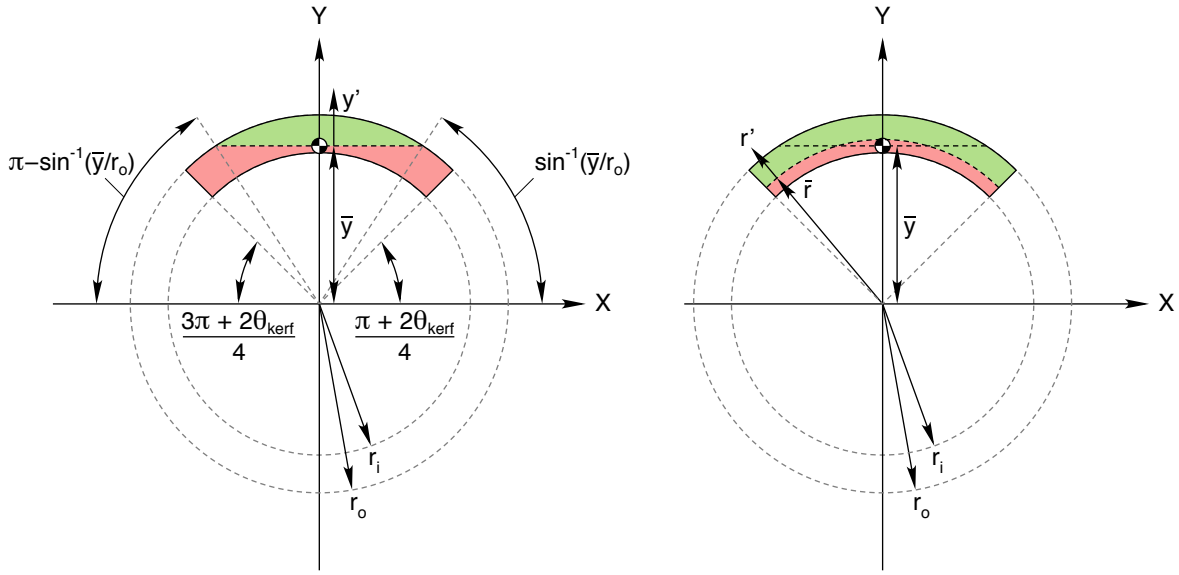
The residual stress in the beam with Young's modulus E_A at a distance y' from the neutral axis is then

$$\sigma_{ZZ}(y) = E_\mathrm{A}\, \varepsilon_{ZZ} = -E_\mathrm{A} \kappa y' = -E_\mathrm{A} \kappa\, (y - \bar{y})\,. \tag{4.4}$$

4.3.2.3 Equivalent radial distribution of residual stress

Next, the equivalent radial distribution of stress was computed for the elementary beam theory's distribution of stress (Equation (4.4)), which is proportional to the vertical distance y' from the neutral axis $\bar{y}$, with

$$y' = y - \bar{y} = r \sin\theta - \bar{y}. \tag{4.5}$$



(a) (b)

Figure 4.15: Coordinates for the tube quarter section with stress distributions proportional to (a) y', the vertical distance from the horizontal neutral axis $\bar{y}$, and (b) r', the radial distance from the radial neutral axis $\bar{r}$.

Similarly, the radial distribution of stress (Figure 4.15b) is proportional to the radial distance r' from a radius with zero stress, $\bar{r}$, with

$$r' = r - \bar{r}. \tag{4.6}$$

The tube was considered as two parts, above and below the neutral axis $\bar{y}$. Above the neutral axis $\bar{y}$, with the stress distribution proportional to y', the stress integrated in the area is

$$\int_{\sin^{-1}(\bar{y}/ro)}^{\pi-\sin^{-1}(\bar{y}/ro)} \int_{\bar{y}/\sin\theta}^{r_o} -\kappa E_A (r\sin\theta - \bar{y})\, r\, dr\, d\theta. \tag{4.7}$$

The stress distribution proportional to y' integrated below the neutral axis $\bar{y}$ is

$$\int_{(\pi+2\theta_{kerf})/4}^{(3\pi+2\theta_{kerf})/4} \int_{r_i}^{\bar{y}/\sin\theta} -\kappa E_A (r\sin\theta - \bar{y})\, r\, dr\, d\theta. \tag{4.8}$$

The radial distribution of stress is assumed to be proportional to radial distance r' from the radius at which the stress is zero ($\bar{r}$), with a fitting constant A to equate the radially

proportional and vertically proportional stress distributions, with

$$\sigma_{ZZ}(r) = -E_{\mathrm{A}}\kappa A\,(r - \bar{r})\,. \tag{4.9}$$

Above the neutral axis $\bar{y}$, the radial distribution of stress integrated in the area is

$$\int_{\sin^{-1}(\bar{y}/r\mathrm{o})}^{\pi-\sin^{-1}(\bar{y}/r\mathrm{o})} \int_{\bar{y}/\sin\theta}^{r_{\mathrm{o}}} -\kappa E_{\mathrm{A}} A(r - \bar{r})\, r\, dr\, d\theta. \tag{4.10}$$

The radial stress distribution integrated below the neutral axis $\bar{y}$ is

$$\int_{(\pi+2\theta_{\mathrm{kerf}})/4}^{(3\pi+2\theta_{\mathrm{kerf}})/4} \int_{r_{\mathrm{i}}}^{\bar{y}/\sin\theta} -\kappa E_{\mathrm{A}} A(r - \bar{r})\, r\, dr\, d\theta. \tag{4.11}$$

To find the radial stress distribution, there are two unknowns ($\bar{r}$ and A), for which two equations were established. For the area above the neutral axis, Equation (4.7) equals Equation (4.10). For the area below the neutral axis, Equation (4.8) equals Equation (4.11). Solving the two equations with the tube dimensions, $\bar{r} = 1.436$ mm, and $A = 2.109$. At the inner radius, $r'\mathrm{i} = r\mathrm{i} - \bar{r} = -0.166$ mm, and at the outer radius, $r'\mathrm{o} = r\mathrm{o} - \bar{r} = 0.152$ mm). Substituting these values into Equation (4.9), along with the Young's modulus of the tube in austenite ($E\mathrm{A} = 66$ GPa) and the radius of curvature in the quarter sections was $\rho = 1/\kappa = 740 \pm 100$ mm (measured from Figure 4.3), yields residual stresses of 31 MPa at the inner radius and -29 MPa at the outer radius.

4.3.3 X-ray penetration depth

From the Beer-Lambert law, for an X-ray beam with intensity I_0 traveling through a homogeneous, isotropic material (assumed) with a linear absorption coefficient μ has a final intensity I after traveling path length L, given by

$$I/I_0 = \exp(-\mu/L), \tag{4.12}$$

for which L can be calculated by

$$L = -\ln\left(I/I_0\right)/\mu. \tag{4.13}$$

For 99% absorption ($I/I_0 = 0.01$), mass attenuation coefficient μ_{NiTi} of NiTi (742 1/cm [122] for Cu-Kα), the path length $L_{99\%}$ is

$$L_{99\%} = -\ln 0.01/\mu_{\mathrm{NiTi}} = 62\ \mu\mathrm{m}. \tag{4.14}$$

The penetration depth, T, for Bragg angle θ is given by

$$T = 0.5L \sin\theta. \tag{4.15}$$

For the maximum Bragg angle in these texture measurements (for the 211 pole at $2\theta = 77.7^{\circ}$), the X-ray penetration depth is 19 µm.

4.4 Summary

The texture of NiTi tube was measured with X-ray diffraction and analyzed as an orientation distribution function. A strong (111) texture was found in the drawing direction of the tube. Additionally, residual stress was observed in the tube during sample preparation, with quarter sections of the tube spontaneously bending with axial curvature. Using an analysis based on Euler–Bernoulli beam theory, the residual stress was estimated as −29 MPa as the outer radius and 30 MPa at the inner radius. Although many NiTi device preparation steps would relax these residual stresses, they would be important for the mechanical measurements of the tubes, as well as the performance of devices without relaxed residual stresses.

CHAPTER 5

Grain size effects on NiTi shape memory alloy fatigue crack growth

5.1 Introduction

Shape memory alloys (SMAs) are functional materials with the ability to recover large deformations via reversible, diffusionless, martensitic phase transformations. These phase transformations can be induced either thermally or mechanically, creating a strong temperature dependence on the mechanical stress-strain response of the material. Thermally activating the phase transformations from below the phase transformation temperature enables the *shape memory effect*, or the recovery of a previous structural shape after seemingly permanent deformation. Mechanically activating the phase transformation at temperatures just above the phase transformation temperature gives rise to *superelasticity*, or the ability to recover unusually large strains of about 5 to 8 % for NiTi SMAs.

Repeated thermomechanical cycling of phase transformations in SMAs can cause two types of fatigue, *functional* fatigue and *structural* fatigue, both of which are relevant in SMA applications. Functional fatigue (or shakedown) is the progressive decay in the phase transformation response during early-life cycling, including the evolution of transformation temperatures, residual strain, latent heat, enthalpy, and hysteresis. Structural fatigue involves material failure with crack initiation, fatigue crack growth (FCG), and eventual fracture. In SMAs, fatigue cracking involves complex thermal and mechanical interactions among plasticity, phase transformation, and other deformation mechanisms such as martensite detwinning.

The focus of this chapter is to characterize the structural fatigue of polycrystalline NiTi SMAs of various grain sizes. The multiscale experiments developed in this chapter for studying grain size effects are extended into Chapters 6 and 7 for studying texture and temperature effects, respectively. A damage-tolerant approach to structural fatigue measurements is

taken, where incremental crack growth rates are measured as a function of stress intensity, as opposed to a total-life fatigue approach. As noted in [123], the majority of structural fatigue studies to date on NiTi have taken total-life fatigue approaches. This is primarily because the biomedical devices that comprise the largest share of NiTi applications have small dimensions and critical crack lengths, such that the time between crack initiation and failure is very short [123]. Total-life fatigue approaches are most applicable to a specific product or device form, while a damage-tolerant approach distinguishes fatigue crack thresholds and fatigue crack growth rates for the material in a more general sense [124].

Damage-tolerant approaches cyclically load a material sample, such as a notched/precracked compact tension specimen, while measuring the crack length (a) during a number of cycles (n). The load is converted to a stress intensity factor (K) that characterizes the elastic stress singularity at the crack tip, and the sample is cyclically loaded between two values, $\mathrm{K_{min}}$ and $\mathrm{K_{max}}$. The stress ratio is defined as $R = \mathrm{K_{min}}/\mathrm{K_{max}}$. The crack growth rate ($\mathrm{d}a/\mathrm{d}n$) is plotted against the rise in stress intensity factor ($\Delta\mathrm{K} = \mathrm{K_{max}} - \mathrm{K_{min}}$). Below a certain threshold value ($\Delta\mathrm{K} < \Delta\mathrm{K_{th}}$), the crack growth rate is essentially zero. At moderate stress intensities above this value, the regime of stable crack growth is often described by a Paris Law [125].

At the microscale, measurements of the local crack opening displacements can capture the behavior near the crack tip. In these measurements, a relevant parameter is the minimum value of the stress intensity necessary to separate the adjacent crack faces ($\mathrm{K_{open}}$). As originally shown experimentally in an Al alloy (with $R = 0$) in [126] and nicely described in [127], a crack may remain closed during a portion of a fatigue cycle, due to residual stresses from plastically deformed material ahead of the crack tip. As the crack advances, material behind the crack unloads, and trailing crack surfaces are driven into contact before the minimum load is reached during the cycle. Thus, a (potentially large) residual compressive stress is generated in the wake of the crack that requires a finite positive load, or $\Delta\mathrm{K_{open}}$, to open the crack. This 'crack closure effect' is the result of elastoplastic behavior of the material, which is not captured by linear elastic fracture mechanics. Crack closure reduces the effective stress intensity factor and generally results in slower crack growth in the material than in a purely elastic material.

In this work, both macroscopic crack growth rates and microscopic crack opening displacements are measured. This multiscale experimental approach is used to study the effects of grain size on fatigue cracking in NiTi, but it could be readily applied to study fatigue cracking in other structural materials.

5.1.1 Nanocrystalline NiTi shape memory alloys

When the average grain size (GS) of the SMA NiTi is refined to the nanocrystalline (NC) regime, several beneficial properties emerge, including a very narrow stress-strain hysteresis, a broad temperature window for superelasticity, improved thermomechanical cyclic stability, reduced latent heat generation, and a reduced sensitivity to deformation rate [128]. These properties give rise to better stability in terms of *functional* fatigue, but the *structural* fatigue of NC NiTi remains poorly characterized.

Interestingly, NC NiTi SMAs do not exhibit shape memory, but still exhibit superelasticity [129]. While the absence of shape memory limits the use of NC NiTi in certain applications, the majority of commercial SMA applications exploit superelasticity, such as biomedical devices [130]. Moreover, compared to commonly available NiTi SMAs, NC NiTi has a wider temperature range for superelasticity and reduced thermomechanical sensitivities. Typical coarse-grained NiTi has a superelastic temperature window of about 50 °C, while 10 nm GS NiTi has a superelastic window of 130 °C [128]. Coarse-grained NiTi has acute thermomechanical couplings that alter the material's stress-strain response with relatively small temperature changes, on the order of tens of degrees Celsius, which leads to hypersensitivities to loading rate and heat transfer characteristics of the ambient medium [131]. These issues are largely avoided by the use of NC NiTi.

The origin of NC NiTi's remarkable properties is a homogeneous phase transformation, unlike the macroscopic nucleation and growth behavior of transformation in coarse-grained NiTi that involves strain localization and propagation phenomena [131, 7, 132, 133]. In typical coarse-grained NiTi (GS $\approx$ 10 µm [123]), martensitic phase transformation localizes as multiple laminates within a grain. With GS below about 70 nm, however, a single laminate of martensite can form. The grain size threshold for this single, continuous martensite laminate in NC NiTi has been measured with X-ray diffraction (XRD) to be 68 nm, under which there is a transition in XRD profiles from multiple peaks to a single, shifting peak [134]. Furthermore, a recent simulation of the interplay between interface and strain energies predicted that multiple laminates of martensite within one grain are more favorable than a single laminate above a grain size diameter of about 70 to 90 nm [135].

Despite the remarkable functional fatigue resistance of NC NiTi, the critical stress intensity factor for crack extension (or mode I fracture toughness, K_{IC}) monotonically decreases with nanocrystalline grain size [136]. For 10 and 18 nm GS NiTi, the crack growth resistance curves (stress intensity factor as a function of crack extension) were nearly flat. SEM fractography revealed cleavage planes with little ductile dimpling in the 10 and 42 nm GS, a mix of cleavage planes and dimpling in the 80 nm GS, and significant ductile dimpling in the 1500 nm GS. Notably, these fracture measurements were conducted on NC NiTi sam-

ples with the same compact tension geometry, cold rolling, and heat treatments used in the present study [136].

There is currently a lack of consensus about NC NiTi's fatigue crack growth resistance. No results exist on damage tolerance with crack growth rates for cold-rolled NC NiTi, though recent works on cold-rolled NC NiTi [137] and equal channel angular pressed (ECAP) NC NiTi [138] suggest that fatigue crack growth resistance may decrease non-monotonically with decreasing grain size. One recent work [137] studied the low-cycle fatigue life of NC NiTi tensile samples without notch defects. Although the NiTi alloy composition and sample geometry were different from this work, the same cold rolling and heat treatment conditions were used. Grain size refinement was shown to have an effect on fatigue life only under certain loading conditions. The 10 nm GS improved low-cycle fatigue life ($N_f < 10^4$ cycles) for a cyclic uniaxial stress of $\sigma_{max} = 450$ MPa with complete unloading (R = 0), but grain size refinement had no effect for the lower cyclic peak stress of $\sigma_{max} = 300$ MPa. Specific mechanisms for the different crack growth responses were not identified. Another recent work [138] showed that ultra-fine grained ECAP NiTi had significantly increased transformation and yield stresses compared to coarse-grained NiTi without negatively affecting the structural fatigue response. The ultra-fine grained NiTi had large fractions with GS < 1 µm, with some large grains up to a few microns, and the coarse-grained NiTi had GS ≈ 100 µm. Additionally, the FCG resistance was mildly enhanced for high mean stresses (R = 0.7) [138]. Although ECAP processing cannot produce grain sizes below 70 nm and does not achieve the exceptional functional fatigue resistance of cold-rolled NC NiTi [139], the fact that severe plastic deformation of NiTi can result in similar or even improved fatigue crack growth resistance is intriguing.

5.1.2 Fatigue crack characterization

On the macroscale, crack growth rates have been measured with indirect techniques, such as electric potential difference (EPD) and the adjusted compliance ratio (ACR, using a back-face strain gage) method. Both EPD and ACR require crack length calibrations, usually using an optical technique. These methods were applied recently to study the fatigue crack growth rates of ECAP NiTi [138], as well as cold-drawn and hot-rolled NiTi [140]. Unfortunately, indirect measurements cannot accurately capture cracks with branching or skewed paths to the tensile loading axis. Even direct optical observation of a crack can be problematic, due to rigid body motion of the sample, requiring position reference marks on the sample [141].

Originally developed by Sutton and co-workers in the 1980s, digital image correlation (DIC) is a powerful optical method for measuring displacement fields on a surface [45].

Measurements can be performed either with a single camera to measure displacements in two dimensions (2-D DIC), or with two or more cameras to measure displacements in three dimensions (3-D DIC). In the context of materials testing, DIC measures displacements on a sample surface from a sequence of images captured during mechanical testing. A DIC algorithm (between reference and current image pairs) is inherently length-scale independent, so the physical length scale is determined by the field of view (FOV) of the images. The technique requires uniformly sized, randomly positioned, and visibly contrasting features on the sample to form a "speckle pattern" for tracking. Unless the sample surface already has suitable contrasting features, an artificial speckle pattern is applied, often by painting for samples at the millimeter scale. The size and distribution of the speckles, combined with the digital camera resolution, determines the spatial resolution of the measurement.

DIC was first applied to problems in fracture mechanics in 1987 to estimate stress intensity factors [142]. The application of DIC to measure crack opening and closing displacements was pioneered in the late 1990s [143], but there have been relatively few studies using DIC to measure crack displacements, as noted by [144]. Optical observations of the crack tip are especially helpful when cracks branch or run along oblique directions to the loading axis. Incremental crack growth rates have been measured with optical DIC (primarily 2-D DIC [145]), but these measurements can exhibit a large degree of scatter due to the limited magnification of optical microscopy. A comparison of DIC with a crack mouth gage and the ACR method (back-face strain gage) was performed in [146], and DIC was accurate for samples that approximated plane stress and do not have significant crack length differences across the sample thickness. Crack length differences here were verified in Section 5.2.3 to be acceptably small.

Microscale optical DIC has been used to examine FCG rates in recent multiscale investigations by Sehitoglu and others [147, 145, 148, 149], at spatial resolutions in the range of 0.4 to 2.0 µm per pixel. The importance of crack closure measurements and the application of DIC for measuring crack opening and closing displacements was discussed by Carroll et al. [147]. Recently these multiscale investigations with optical, microscopic DIC measurements of crack opening have been applied to examine the damage tolerance of advanced alloys, including nanocrystalline Ni-Co alloys [145] and Hastelloy X [148]. Another study on Hastelloy X included sub-grain level measurements of plastic strain accumulation near fatigue cracks, with the microstructure characterized by electron backscatter diffraction and overlaid with optical DIC measurements [149].

DIC within a scanning electron microscope (SEM-DIC) was introduced in the late 2000s and overcomes magnification limitations imposed by the wavelength of visible light [150, 151]. The precision required to measure the microscopic crack opening levels in the present work

was enabled by SEM-DIC. Optical microscopic DIC, with about an order of magnitude coarser resolution from the diffraction limit of light, may not have been able to resolve the sub-500 nm Δv displacement differences in this work between the 1500 nm sample and the other grain sizes. Complicated techniques exist to achieve better resolution than the diffraction limit in optical microscopes (such as interferometric microscopy, 4Pi microscopy, total internal reflectance microscopy, and stimulated emission depletion microscopy), but to the authors' knowledge, these techniques have not yet been applied with DIC.

SEM-DIC requires a sufficiently fine speckle pattern on the sample surface. One option is the application of self-assembled Au nanoparticles, which was successfully demonstrated for SEM-DIC with a spatial resolution of 4 nm per pixel on 99.99 % aluminum, 1100 aluminum, a nickel-chromium superalloy, and silicon carbide samples [86]. A drawback to SEM-DIC, however, is the added complexity from error sources that are not present with optical DIC, including spatial and drift distortions, scanning errors (also called image shifts or jumps), and non-negligible sample stress relaxation [150, 58]. Various techniques, such as time averaging of images and distortion corrections, have been attempted to address these issues. In this work we introduce an improved external scan control technique to address the source of the electron beam scanning errors.

5.2 Methods

This work presents the first measurements of fatigue crack growth rates at both low and high stress intensities in NC NiTi sheet samples with average grain sizes from 1500 nm down to 10 nm (Section 5.2.1). To connect macroscopic crack growth rates to microscopic crack opening displacements, experiments were performed with both optical DIC and SEM-DIC to study the responses across a broad range of length scales. At both the macroscopic and microscopic length scales, spatially-resolved displacements from digital image correlation (DIC) enabled crack tip measurements without the need for crack length calibrations and assumptions required for other techniques. Macroscopic DIC enables accurate crack growth rate measurements and can identify if and when cracks branch or propagate in oblique directions. The issue of scatter was minimized by using an optimized, optical 3-D DIC system (Section 5.2.2) with high-quality lenses, low-noise machine vision cameras, and cross polarization to increase contrast and eliminate saturated pixels [3]. At the microscale, SEM-DIC was used to measure the relative displacements between the crack faces (Δv) close to the crack tip, aiding in the characterization of mechanisms of crack growth resistance, such as crack closure. SEM-DIC was recently applied by others to measure crack opening and closing displacements in an aluminum alloy with a spatial resolution of 70 nm per pixel [144].

In this study, a resolution of 31.25 nm per pixel was achieved with SEM-DIC during loading of the precracked samples. As discussed in Section 5.2.3, a custom, high-precision, external scan controller was used for the first time that largely eliminated scanning distortions and drift in the electron beam, providing improved accuracy in the SEM-DIC measurements. The experimental program consisted of samples with 10, 18, 42, 80, and 1500 nm GS, each subjected to a three-step sequence of experiments: (1) high ΔK fatigue crack growth measurements with optical DIC, then (2) one loading cycle at high ΔK to measure crack opening displacements by SEM-DIC, then (3) low ΔK fatigue crack growth measurements with optical DIC.

5.2.1 Sample preparation

Superelastic NiTi sheets of 1.72 mm thickness with a slightly nickel-rich composition (50.9 at.% Ni, 49.1 at.% Ti) were obtained from Nitinol Devices Company (NDC, alloy SE508). The as-received sheets were annealed at 800 °C for 1 h in a furnace with flowing argon, and then quenched in room temperature water. The sheets were then cold rolled between stainless steel plates to 1.00 mm (42 % thickness reduction). Compact tension samples conforming to ASTM E561 [152] with dimensions shown in Figure 5.1 were produced by wire electrical discharge machining (EDM), with the crack notch direction perpendicular to the rolling direction. The cold-rolled compact tension samples were heat treated at different temperatures and durations (Table 5.1) to produce average grain sizes of 10, 18, 42, 80, and 1500 nm, as measured by transmission electron microscopy [44, 116]. Above 100 nm GS, the mechanical response of NiTi does not change significantly, so the 1500 nm GS sample is a representative baseline for comparison with the nanocrystalline grain sizes [44, 116]. The condition without heat treatment after cold rolling had an average grain size of 10 nm, so only the larger grain sizes had subsequent heat treatments. The sample surfaces were ground and polished with a sequence of 400, 600, 800, and 1500 grit silica papers, then final polished with a 5:1 volume ratio mixture of colloidal silica (Buehler MasterMet 2, 0.02 µm) and 30 % hydrogen peroxide (Fisher H325) on a microfiber cloth (Buehler MicroCut). The hydrogen peroxide accelerated the final polish with surface oxidation and aided the silanization step for self-assembled gold nanoparticle patterning for SEM-DIC, described in Section 5.2.3.

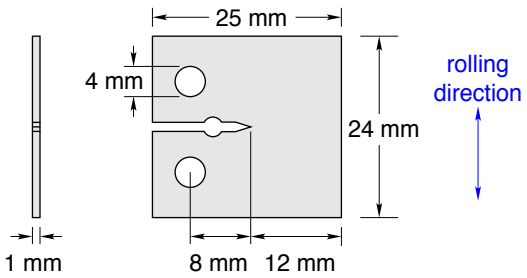


Figure 5.1: Compact tension samples conforming to ASTM E561 were cut with wire EDM from the cold-rolled sheets with the rolling direction perpendicular to the crack notch.

Table 5.1: Heat treatments for grain growth and the grain sizes measured by XRD and TEM. Note that the GS for the as-rolled condition was only measured with XRD. Also, the TEM field of view for the 1500 nm condition contained only a few grains (too few to make a distribution and statistics on grain size).

average GS, XRD (nm)	average GS, TEM (nm)	temperature (^{o}C)	time (min)
10	n/a	n/a	n/a
18	18 ± 12	250	45
42	48 ± 18	520	2
80	80 ± 13	520	6
1500	n/a	600	45

For optical DIC, samples were speckle patterned after polishing using an Iwata CM-B airbrush with a basecoat of white paint (Golden High Flow Titanium White), followed by the application of black speckles (Golden High Flow Carbon Black). This paint sequence, white paint followed by black paint, provides higher contrast and lower DIC errors than the converse [4].

After performing macroscopic, high ΔK fatigue crack growth measurements (Section 5.2.2), the painted speckle pattern was removed with methanol, and the samples were coated with self-assembled gold nanoparticles (AuNPs) to create a suitable speckle pattern for SEM-DIC. The patterning procedure in [86] had three basic steps: (i) promote oxide/hydroxyl groups on the sample surface, (ii) form a thin organosilane layer with silane molecules that covalently bond to the reactive oxide/hydroxyl groups and have pendant functional groups that point away from the oxide/hydroxyl groups, and (iii) soak samples in AuNP colloid and allow AuNPs to self-assemble on the pendant functional groups of the silane molecules. We modified this procedure slightly: step (i), final polishing with the 5:1 mixture of colloidal silica and 30 % hydrogen peroxide was sufficient; step (iii), samples were soaked in AuNP colloid with an average diameter of about 50 nm for an extended period of 7 days, to allow

for the formation of AuNP aggregates for the relatively wide SEM-DIC horizontal field width (HFW) of 200 µm, and on third and fifth days the samples were rinsed and placed in fresh AuNP colloid. After the SEM-DIC crack displacement measurements (Section 5.2.3), the samples were cleaned again with methanol and speckle painted again for the low ΔK fatigue crack growth measurements (Section 5.2.4).

5.2.2 Macroscale fatigue measurements at high stress intensity

The first stage of the fatigue procedure for each sample consisted of macroscale crack growth measurements during cyclic loading at high stress intensity (K). Fatigue cracks were grown in the compact tension samples from the initial EDM notch length of $a_0 = 8.0$ mm to a crack length of $a = 10.0$ mm. This was performed in room temperature air with a TA ElectroForce LM1 system operating at 10 Hz in load control at an initial peak stress intensity factor (K_{max}) of about 6 MPa$\sqrt{m}$ (based on the notch length, a_0) and a stress intensity ratio $R = K_{min}/K_{max} = 0.1$. At this stress intensity level, the fatigue crack growth rate was in a high ΔK regime, with crack growth rates exceeding 10^{-8} meters per cycle.

The stress intensity factor (K) was calculated following ASTM E561 11.3 [152]. For an applied force P, crack length a, sample thickness B, and sample width W,

$$K = \frac{P}{B\sqrt{W}} \frac{\left(2 + \frac{a}{W}\right)}{\left(1 - \frac{a}{W}\right)^{3/2}} f\left(\frac{a}{W}\right), \tag{5.1a}$$

with

$$f\left(\frac{a}{W}\right) = 0.886 + 4.64\left(\frac{a}{W}\right) - 13.32\left(\frac{a}{W}\right)^2 + 14.72\left(\frac{a}{W}\right)^3 - 5.6\left(\frac{a}{W}\right)^4. \tag{5.1b}$$

This formula is based on linear elastic fracture theory, which is not strictly valid for SMA behavior [2]. However, it was adopted here, since it is proportional to the applied load P, provides a simple measure of the stress intensity, and provides a reasonable estimate if the inelastic zone at the crack tip is small.

The crack length was monitored with an optical 3-D DIC system with a precision of about 25 µm minimum detectable crack extension and an accuracy within about 50 µm, as determined from postmortem SEM images. The optical 3-D DIC system (shown in supplemental Figure 5.2) consisted of two cameras (Point Grey Grasshopper GRAS 50S5M-C with Nikon Micro Nikkor 200 mm f/4 lenses) and two LED light panels (LitePanels Astra 935-1003). The field of view was 8.0 mm and 2448 px for a pixel size of 3.3 µm. An image

pair for 3-D DIC was captured every 1,000 fatigue cycles during the minimum and maximum applied loads, and the fatigue cycling was paused for 2 seconds for each image pair. Cross polarization was utilized by placing linear polarizing filters (Tiffen 62POL) on the lenses in an orthogonal position to linear polarizing filters (American Polarizers AP38-030T with 38 % transmission) on the LED lights. Cross polarization increases contrast, lowers optical DIC error, and attenuates saturated pixels to preserve sub-pixel displacement accuracy [3]. Distortions from thermal air currents were minimized by gently blowing air from a fan over the setup. The fan was mounted adjacent to the macroscopic fatigue system on a tripod that was not in contact with the optical table of the DIC system, to avoid introducing vibrations [95].

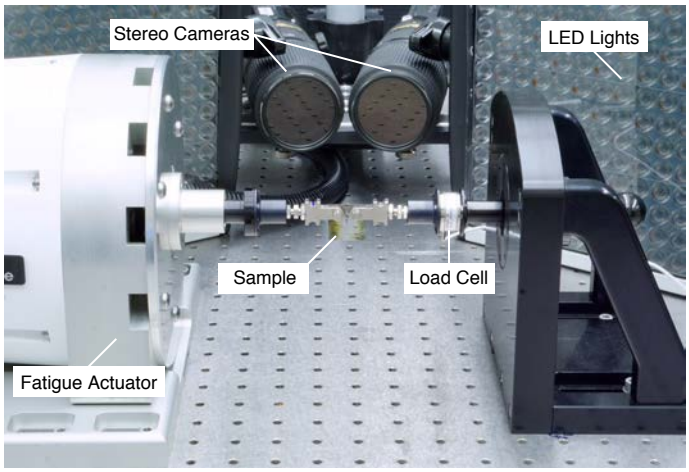


Figure 5.2: Macroscopic fatigue crack growth setup with optical 3-D DIC system, including cross polarization for enhanced optical DIC [3]. For a length scale reference, the grid spacing of the optical table breadboard was 1 inch.

Cracks can be identified and tracked using various DIC quantities, including the correlation residual (also called correlation confidence interval, or CCI), displacement field discontinuities, and strain field discontinuities [153]. CCI is a statistical measure of the local quality of correlation, with units of pixels, and a value of zero indicating perfect correlation. Using CCI to locate crack tips was introduced by Cady, Liu, Rae, and Lovato [154, 155]. The major principal strain field (ε_1) has also been used to locate multiple cracks in the complicated fracture of large masonry walls [156]. Regardless of the quantity used to track the crack tip, only the discontinuities in the field are needed, not their precise values. For the NiTi samples used in this study, the crack tip position was measured by either the CCI or the in-plane strain perpendicular to the crack path, ε_{yy}, whichever provided the clearest discontinuity at the crack tip. For high stress intensity factors (ΔK) during fatigue precracking, the CCI was used because the high strains (process zone) around the crack tip blurred the discontinu-

ity. For low stress intensity factors, the discontinuity from the correlation residual was not large enough to reliably track the crack tip, and the false strains from the separating crack faces (ε_{yy}) were used instead. The threshold ranges for crack tip detection were 5×10^{-3} to 15×10^{-3} px for CCI, and 0.4 to 0.8 % for ε_{yy}.

5.2.3 SEM-DIC measurements

Following the macroscopic fatigue measurements (just described), each sample was subjected to a single high ΔK loading cycle during which crack opening displacements were measured at the microscale by SEM-DIC. Measurements were performed with an FEI Teneo SEM equipped with an in-situ mechanical testing stage (Kammrath & Weiss 5 kN tension/compression module) and a custom, high-precision external scan controller to lower SEM-DIC error. Scanning was controlled by a C++ script running a digital-to-analog converter (DAC, National Instruments NI USB-6251). The output voltage of the DAC board was connected to the input of a pair of analog amplifiers in the SEM, which controlled the SEM scan coils. The secondary electron detector's analog voltage signal was then fed back as an input to the DAC board and was recorded for each image.

The external scan controller provided acceptably low SEM-DIC errors, without the spatial and drift distortion corrections that are normally required for high-magnification SEM-DIC [151, 58]. Before each experiment, baseline noise and error measurements were performed using static and translated images captured away from the crack tip. Each image had an HFW of 200 µm (6400 px). Static SEM-DIC images (no sample displacement) captured drift distortion and noise errors, while translated images captured spatial distortion and noise errors [151]. The noise in the displacements (u, v) and major principal strain ε_1 from a representative set (the 80 nm GS sample) is provided in Figure 5.3, showing small but noticeable image shift errors. Averaging over all five sets (one for each GS) of static images, the false $\overline{u} = 0.21 \pm 0.74$ px $= 0.007 \pm 0.023$ μm (mean $\pm$ standard deviation) and the false $\overline{v} = -0.19 \pm 0.31$ px $= -0.006 \pm 0.010$ μm. The translated images were captured with a sample displacement of $u = 10$ μm (5 % of the HFW), resulting in average false $\overline{\varepsilon}_1 = 0.021 \pm 0.007$ % and standard deviations $\overline{\Delta u} = 0.32$ px $= 0.010$ μm and $\overline{\Delta v} = 0.44$ px $= 0.014$ μm. These displacement errors in terms of pixel units are somewhat larger than desired, but are still small in μm, and quite satisfactory considering the typical values (several microns) measured in this study. For comparison, baseline measurements from a recent SEM-DIC study [157] with very low SEM-DIC errors after drift and distortion corrections, had standard deviations of $\overline{\Delta u} = 0.023$ px $= 1.3$ nm (for their 57 µm, 1024 px HFW) and $\overline{\Delta \varepsilon_{xx}} = 0.020$ % after translation. Here, the $\overline{\Delta u}$ and $\overline{\Delta v}$ noise levels correspond to 10 and 14 nm (0.1 % and 0.14 % of the 10 µm translation), larger than theirs but accept-

able for our purposes. Our average false horizontal strain was $\overline{\Delta\varepsilon_{xx}} = -0.005 \pm 0.015$ %, which is comparable to theirs.

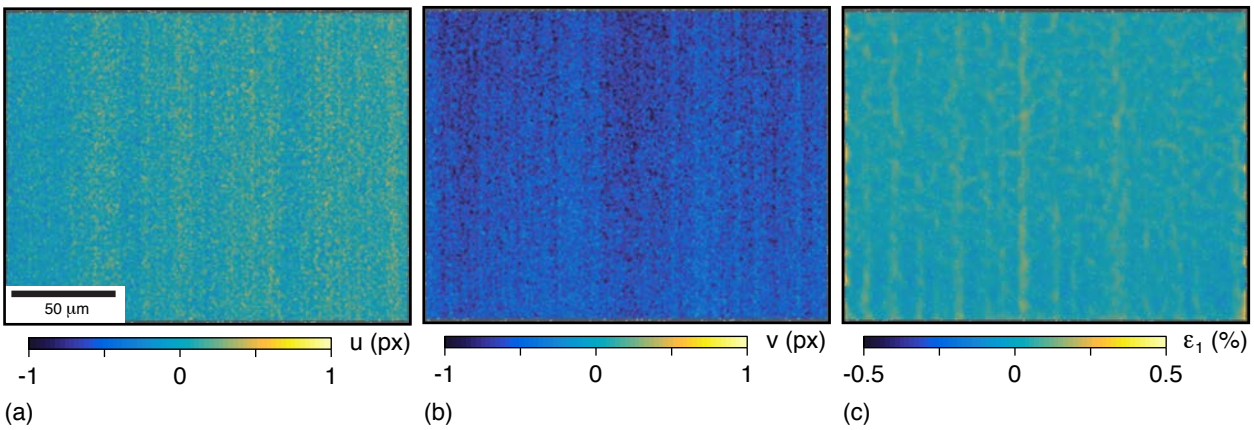


Figure 5.3: A representative set of baseline SEM-DIC images is shown: (a,b) false horizontal and vertical displacements (u, v) by static image pairs; and (c) false major principal strain ε_1 by rigid translation of the sample. The vertical streaks are sub-pixel image shift errors, yet the errors from drift, spatial distortions, noise, and scanning shifts are low relative to the strains and displacements from the crack tip measurements (shown in Figures 5.9 and 5.10). For this set, the respective means and standard deviations are $u = -0.05 \pm 0.11$ px (-0.002 ± 0.004 μm), $v = -0.59 \pm 0.11$ px (-0.018 ± 0.004 μm), and $\varepsilon_1 = 0.024 \pm 0.037$ %.

During quasistatic loading of the samples, force P readings were monitored and the loading was periodically paused to image the crack tip vicinity. After waiting about five minutes for the sample to stress relax, secondary electron images were captured at 5 µs/px dwell time, giving a total scan time per image of about 4 minutes. No line or image averages or integrations were performed. The SEM images were correlated using Vic2D 6 commercial DIC software (Correlated Solutions, Inc.) with a 15 px step size, yielding about 140,000 DIC data points per image. The correlation used low-pass filtering, Gaussian subset weights, eight-tap interpolation, and the normalized squared differences method, with 0.02 px consistency threshold, 0.05 px maximum confidence interval, and 0.1 px matchability threshold.

Only one fatigue loading cycle was executed on each sample for the in-situ SEM-DIC crack opening displacement measurements. The maximum force applied to each sample during testing was adjusted to account for specimen variations (slight differences in fatigue precrack lengths and sample thicknesses due to polishing) to produce a consistent maximum stress intensity factor $K_{max} = 8.0 \pm 0.1$ MPa$\sqrt{m}$. The reference image for each experiment was taken at the minimum stress intensity factor $K_{min} = 0.8$ MPa$\sqrt{m}$, following Ref. [144]. Images were captured in 0.8 MPa$\sqrt{m}$ increments during loading to K_{max}, giving a total of 10 SEM-DIC images (including reference image) for each sample.

The relative crack displacements Δv were measured similar to the approach in [144], although no standard method exists for this new SEM-DIC technique. Since there is a discontinuity in displacement field across a crack and since DIC has a finite step size and subset size, a small region next to each crack face is unmeasurable. Therefore, crack opening measurements were taken at an offset distance ($y_0 = \pm 10$ µm) on either side. The origin ($x = y = 0$) was defined at the crack tip, with $x > 0$ in front of the crack and y pointing normal to it. The relative displacements were taken as $\Delta v(x) = v(x, y_0) - v(x, -y_0)$. The crack tip opening displacement (CTOD) is then $\Delta v(0)$, when measured across the crack tip. The crack opening level, or the stress intensity at which the crack faces open ($\mathrm{K_{open}}$), is indicated by $\Delta v(0) > 0$.

The crack length difference between front and back faces of the sample was verified to be within the ASTM E647 [141] criterion of 25 % of the thickness. Between the two faces of the samples, the difference between the crack lengths averaged 140 µm for the five samples. The largest discrepancy was in the 1500 nm GS sample, where the crack length difference was 200 µm, or 21% of the thickness. The crack length for the stress intensity calculation was measured by an optical microscope on the same sample face that was imaged for SEM-DIC.

5.2.4 Macroscale fatigue measurements at low stress intensity

After the SEM-DIC experiment, each sample was again cyclically loaded with the LM1 ElectroForce system, but cycling at 50 Hz and at a lower ΔK to probe the fatigue thresholds and low ΔK crack growth rates. To determine the fatigue threshold $\mathrm{\Delta K_{th}}$, the stress intensity factor range ΔK (with $R = 0.1$) was decreased incrementally until the crack growth rate was below one Angstrom (10^{-10} meters) per cycle, as measured over the span of at least $250,000$ cycles. Upon reaching $\mathrm{\Delta K_{th}}$, the minimum and maximum cyclic forces were held fixed until the crack had propagated from the precrack length of $a \approx 10.0$ mm (2.0 mm from the 8.0 mm notch) to a final length of $a \approx 14$ to 15 mm (the edge of the DIC field of view).

5.3 Results and discussion

Experimental results are presented and discussed below to quantify the fatigue behavior of the five GS samples and explore their trends. The experimental program subjected each sample to a three stage procedure: (1) high stress intensity fatigue loading for thousands of cycles with optical DIC measurements; (2) one high stress intensity loading cycle with SEM-DIC measurements; and (3) low stress intensity fatigue loading for hundreds of thousands of cycles with optical DIC. Microstructural observations of the samples are first provided to characterize the material (Section 5.3.1). Both sets of macroscale fatigue results, (1) and

(3), are presented together (Section 5.3.2). SEM-DIC results (2) of micro-scale crack opening displacements are then discussed (Section 5.3.3), followed by postmortem fractography results (Section 5.3.4).

5.3.1 Microstructure

In the as-rolled sheet without any subsequent heat treatment (10 nm GS), severe plastic deformation from a 42 % thickness reduction during cold rolling generated significant texture, defects, and residual martensite. High-resolution transmission electron microscopy (HRTEM) imaging of the as-rolled (10 nm GS) sample in Figure 5.4a revealed austenite nanocrystals (blue arrows) with a high density of defects and lattice strains (red arrows). Residual martensite (also known as retained or deformation-induced martensite), as evident from the twin lamellae of martensite (green arrows), existed heterogeneously in the as-rolled sheet. An important consequence of residual martensite is that the material may exhibit significant in-plane anisotropy, both in the mechanical response [158] and thermal expansion [40], due to martensite's low (monoclinic) crystal symmetry. Interestingly, the volume fraction of residual martensite varies non-monotonically with the amount of cold rolling, with a maximum near 20 % thickness reduction, beyond which the residual martensite fragments [159]. The corresponding selected-area electron diffraction (SAED) pattern, taken from the red circle area in Figure 5.4a, shows the existence of weak satellites belonging to martensite 002B19' just within the ring of the main 110B2 spots. From a previous study [40] with the same material and processing used here, the sample with 42 % cold-rolling had complete γ-fiber $\{111\}\langle uvw\rangle$ in the surface normal direction and a relatively weak α-fiber $\{331\}\langle 110\rangle$ texture component in the rolling direction. In that study, the volume fraction of residual martensite in the as-rolled sample was estimated to be about 25 % from chi-scan XRD rotations.

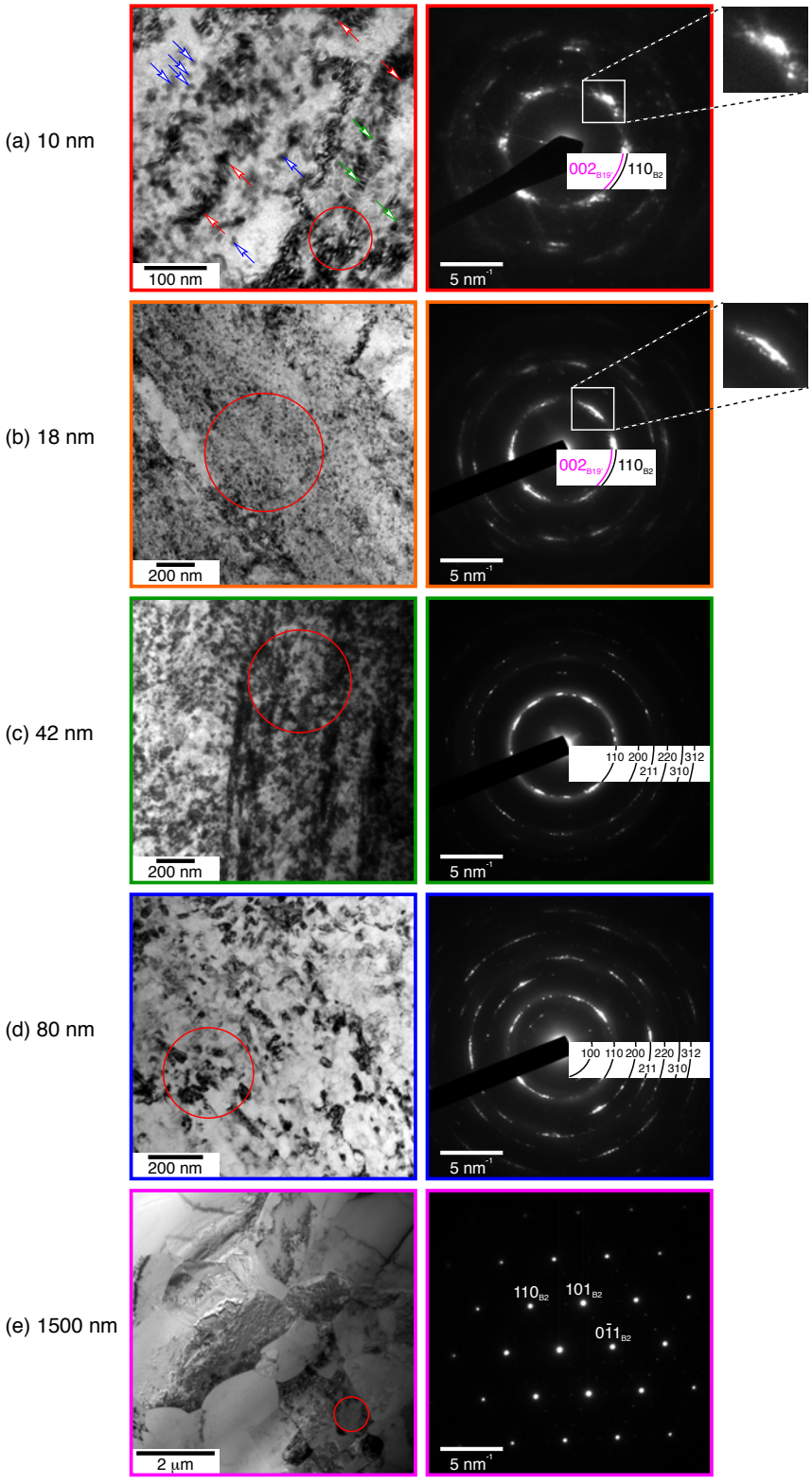


Figure 5.4: High resolution transmission electron micrographs (HRTEM) and selected-area electron diffraction (SAED) patterns are shown for each grain size: (a) 10 nm, (b) 18 nm, (c) 42 nm, (d) 80 nm, and (e) 1500 nm.

The 18 nm GS also exhibited residual martensite, but only about half as much as the 10 nm GS. The 18 nm sample was annealed at 250 °C and underwent slight grain growth from the as-rolled condition. The diffraction pattern of the 18 nm sample still exhibited residual martensite, as inferred from the extra ring near the 110_{B2} in the SAED pattern of Figure 5.4b. Previous work measured the volume fraction of residual martensite for the 18 nm condition to be 11 % [40].

All other GS samples showed no evidence of residual martensite. The 42 nm sample was annealed at 520 °C for two minutes and was fully austenitic at room temperature, as shown by the distinct 110_{B2} ring and spots in Figure 5.4c. Further annealing continued to grow the grains, with 600 °C for 6 minutes yielding 80 nm GS and 600 °C for 45 min yielding 1500 nm GS. The 1500 nm GS sample had relatively coarse austenite grains, with the diffraction from one grain shown in Figure 5.4e.

5.3.2 Macroscale fatigue responses

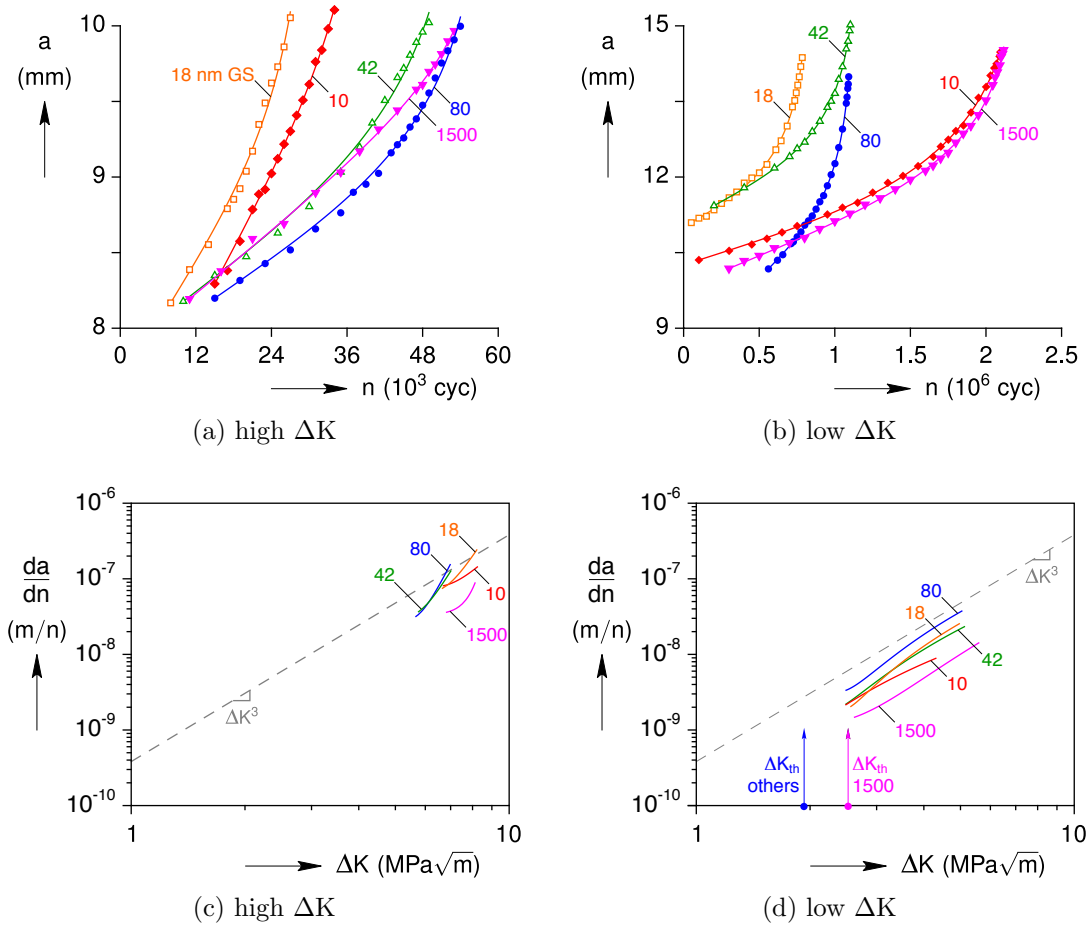
Macroscopic crack length (a) versus cycle (n) results for all five GS samples are summarized in Figures 5.5a and 5.5b for high ΔK and low ΔK, respectively, showing typical exponential-like growth trends for each sample but irregular trends across the grain sizes. We discovered a simple, three-parameter function to fit the data in Figures 5.5a and 5.5b (shown by the curved lines). The function has the form

$$a(\tilde{n}) = A\tilde{n} + B\tanh^{-1}\left[C\tilde{n}\right], \tag{5.2}$$

where $\{A, B, C\}$ are fitting constants, and $\tilde{n}$ is a normalized cycle variable according to

$$\tilde{n} = (n - n_0)/n_f. \tag{5.3}$$

It is well known that obtaining accurate rate information from noisy experimental data is problematic, so some form of filtering or fitting of the data is usually needed before taking any derivatives. ASTM E647 [141] proposes an incremental polynomial method to provide better $\mathrm{d}a/\mathrm{d}n$ results than taking differences of successive data points. Equation (5.2) fit the crack growth data quite nicely over all cycles, without the need for "sliding polynomial" domain shifting. All of the fits shown in Figure 5.5a (high ΔK) and Figure 5.5b (low ΔK) had coefficients of determination $R^2 \geq 0.9920$ and $R^2 \geq 0.9992$, respectively. These functions were subsequently used to generate smooth $\mathrm{d}a/\mathrm{d}n$ curves.



(a) high ΔK (b) low ΔK

(c) high ΔK (d) low ΔK

Figure 5.5: Macroscopic crack lengths a in the five compact tension samples were measured by 3-D DIC during cycling n at (a) high ΔK then (b) low ΔK. Data points were fitted with an inverse hyperbolic tangent function (lines), which were used to generate respective growth rate da/dn curves shown in (c) and (d). The dashed gray lines denote a Paris Law exponent of $m = 3$. At both low and high ΔK, the 1500 nm GS sample had the lowest crack growth rates, and the 80 nm sample had the highest crack growth rates (although the 42 nm sample was quite close at high ΔK).

The corresponding crack growth rate (da/dn) versus ΔK curves (Figures 5.5c and 5.5d) show a complicated ordering with respect to grain size. For reference, the dashed gray line shows the Paris Law with an exponent of $m = 3$, which aligns reasonably well with the slopes of the rate curves at low ΔK in Figure 5.5d. For both low and high ΔK experiments, the 1500 nm GS sample exhibited the slowest crack growth and the 80 nm GS sample exhibited the fastest. At high ΔK, the apparent ordering with increasing crack growth rates was $\{1500, 10, 18, 42, 80\}$ nm GS, based on where curves overlapped at a common ΔK. At low ΔK, the ordering was $\{1500, 10, 42, 18, 80\}$ nm. For high ΔK, the fatigue crack growth rates of the 10 and 18 nm samples were slightly lower than those of the 42 and 80 nm samples.

These differences may be due to the presence of residual martensite in the 10 and 18 nm conditions, about 25 % and 11 % volume fractions [40], respectively. The slight improvement in the high ΔK regime for the 10 nm GS agrees with recent total-life fatigue results [137] of tensile bars of NC NiTi, although no differences were observed in moderate-cycle fatigue life among the 10, 42, and 80 nm GS samples.

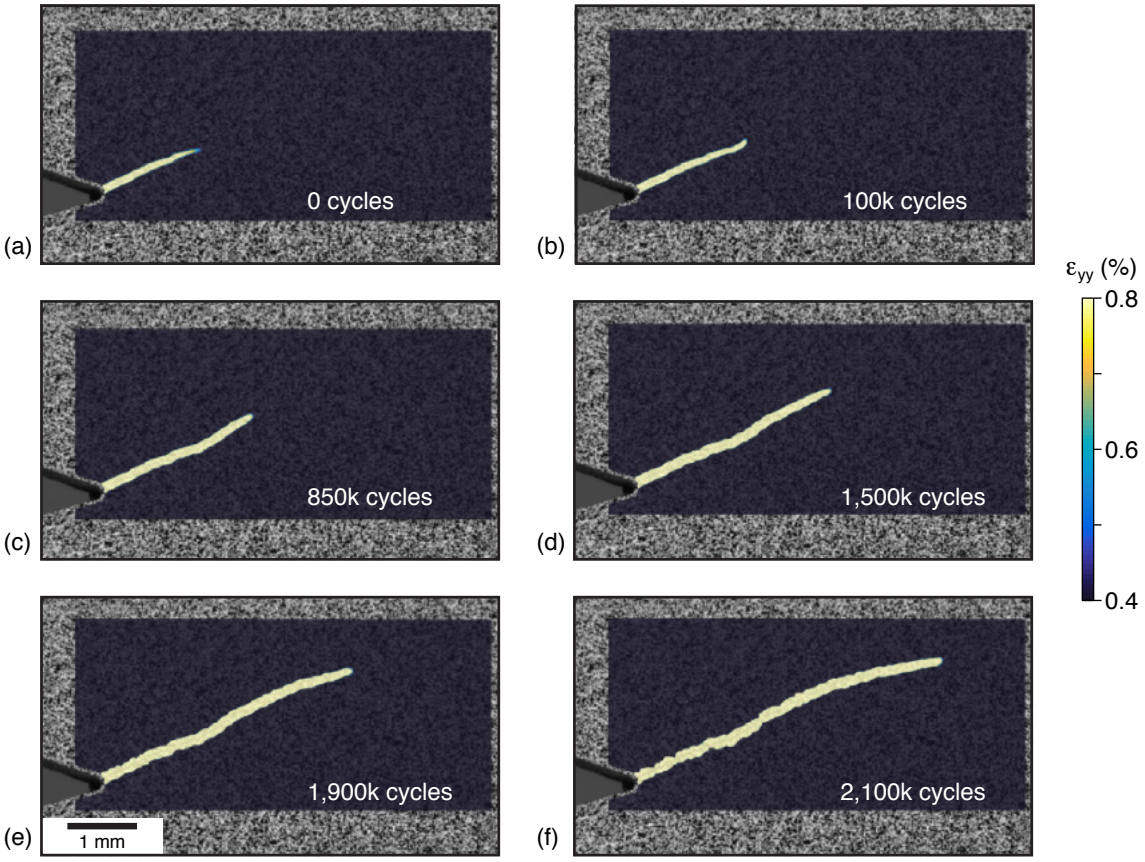


Figure 5.6: In the 10 nm GS sample, the crack grew at an oblique angle during both high and low ΔK cracking (shown here in the low ΔK stage), deviating from the expected mode I path perpendicular to the tensile axis. Its crack length a was measured along the crack path and the sample was subjected to mixed mode I and mode II fatigue, thereby obscuring a direct comparison to the other grain sizes in Figures 5.5a and 5.5b.

Fatigue cracks all grew in the expected (mode I) x-direction, perpendicular to the loading axis, except for the 10 nm GS sample. The 10 nm GS sample exhibited a macroscopic crack path at an oblique angle (Figure 5.6), and since this created a mixed mode I (opening) and mode II (sliding) condition, its interpretation is less clear. In this case, the crack length a was measured along the oblique crack path, underscoring the importance of using spatially-resolved techniques to track crack growth. The source of the anomalous crack path may be due to one or both of the following reasons. First, it may be related to a positive 'T-stress'

σ_{xx}, which is known to promote crack path instability [160]. While a small amount of ϵ_{xx} was detected near the crack tip, the difference between the ϵ_{xx} fields of the 10 and 18 nm GS samples was negligible. Second, it may be due to a strong crystallographic texture in the 10 nm GS that directed the crack along a path of lower resistance. This type of unexpected crack path has been observed in NiTi before, in experiments on compact tension samples produced from NiTi tube [124], where the notches in compact tension samples were oriented in three different directions with respect to the tube axis: longitudinal, circumferential, and 45°. The samples with longitudinal and 45° notches exhibited crack paths in the expected mode I direction, but the circumferential samples exhibited crack paths that were angled about 25° from the notch. Previous measurements on NiTi tube [113] and sheet [161] have also shown that phase transformation is most easily activated when the tensile axis is 45° with respect to the drawing direction.

Crack branching was observed in the 1500 nm GS, presenting another challenge in characterizing fatigue behavior. This occurred at both microscopic (not shown) and macroscopic length scales (Figure 5.7). More than 100,000 cycles were necessary to initiate the first crack, yet within about 24,000 cycles this crack arrested and a second crack grew that dominated the failure process during the remainder of the 260,000 total cycles (Figure 5.7). DIC captured this bifurcation, which likely would have been missed by indirect crack growth measurements that would not distinguish one stagnating crack from a second growing crack.

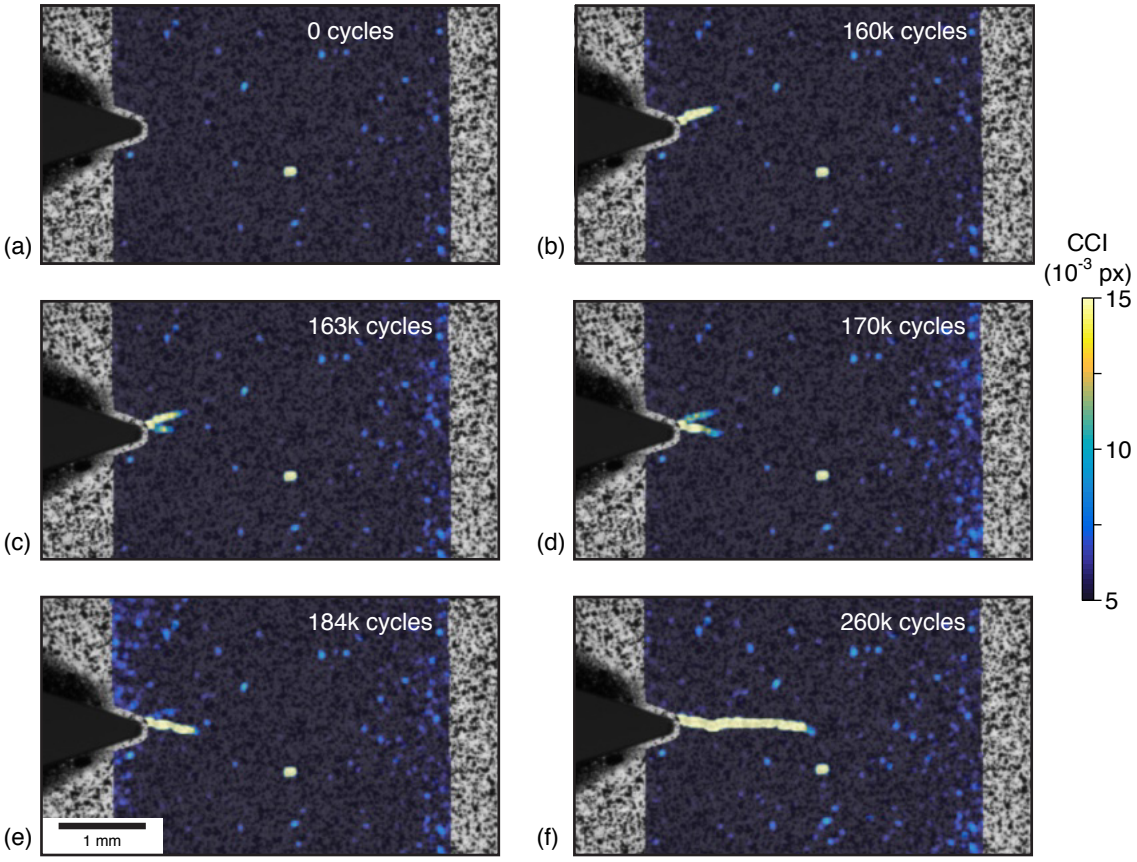


Figure 5.7: Correlation confidence interval (CCI) fields in the 1500 nm GS sample at selected instances during high $\Delta K \approx 6$ MPa$\sqrt{\text{m}}$ cycling show early crack branching. (a) The high CCI spot near the middle of each image should be ignored, likely due to dust on the camera lens observed before cracking. (b) Initially, a crack developed at a positive angle from the notch root. (c) With further cycling, a second crack formed at a negative angle from the notch, and the first crack stopped. (d,e) Subsequently, the second crack continued to grow, while the CCI signal of the first crack progressively faded. (f) The second crack grew further along the x-direction, and the first crack was no longer detectable via CCI since it had completely closed.

In fact, crack branching was observed in another sample with 120 nm GS, which was so extensive that it was removed from the data set here. It had multiple secondary cracks extending laterally from the nominal crack path, which prohibited measuring the relative crack displacements (Figure 5.8). Similar crack branching was observed in a 1500 nm GS sample, but not in the field of view used for the Δv measurements presented below. .

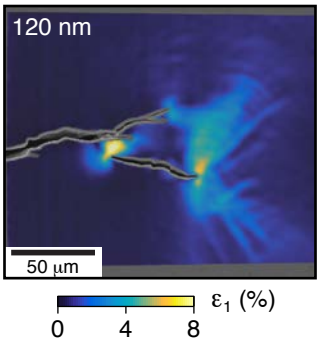


Figure 5.8: For a sample with 120 nm GS that was not included in the results of the main text, a significant degree of crack branching barred valid Δv measurements from the SEM-DIC displacements.

The threshold stress intensities ΔK_{th} of the NC NiTi samples in this work are somewhat smaller than average, but within the expected range, for coarse-grained NiTi. The fatigue threshold for superelastic NiTi with $R = 0.1$ has an average reported value of $\Delta K_{th} \approx 2.8 \pm 1.3$ MPa$\sqrt{\text{m}}$, from aggregated measurements of several experimental studies compiled in a review [123]. Here, the 1500 nm GS sample had a fatigue threshold $\Delta K_{th} = 2.4$ MPa$\sqrt{\text{m}}$, while the 10, 18, 42, and 80 nm GS samples all had a $\Delta K_{th} = 1.9$ MPa$\sqrt{\text{m}}$.

Similar to other intermetallic alloys, NiTi's fatigue threshold decreases with increasing stress intensity ratio R, meaning that fatigue crack growth rates depend on both ΔK and ΔK_{max}. Robertson and Ritchie [124] measured a decrease of ΔK_{th} from 2.48 MPa$\sqrt{\text{m}}$ for $R = 0.1$, down to 1.15 MPa$\sqrt{\text{m}}$ for $R = 0.7$. They proposed that this decrease in ΔK_{th} with increasing R was associated with crack closure mechanisms, and their SEM images of tortuous crack paths and microcracks supported this hypothesis.

5.3.3 Microscale fatigue responses

Between the macroscopic fatigue measurements just described, each sample was subjected to one additional loading cycle at high $K_{max} = 8.0$ MPa$\sqrt{\text{m}}$ and imaged at the microscale by SEM-DIC. The major principal strain (ε_1) fields near the crack tip in each sample at K_{max} are provided in Figure 5.9. With increasing GS from 10 nm to 80 nm, the strain fields exhibited a progressive broadening of process zones (high strain regions) around the crack tip. However, this trend was reversed in the 1500 nm sample that had a small process zone size, between the 10 and 18 nm cases. The 42 and 80 nm samples had large regions of high strain ($\varepsilon_1 \approx 4$ %) roughly ≥ 10 μm around the crack tip, while the 10, 18, and 1500 nm samples had relatively small regions of high strain, < 5 μm around the crack tip. This trend is unexpected, as previous measurements of fracture in NC NiTi demonstrated a

monotonically decreasing fracture toughness K_{IC} with decreasing grain size [136]. Fracture toughness, therefore, does not seem to correlate directly to fatigue behavior (crack growth and crack opening response) in NC NiTi.

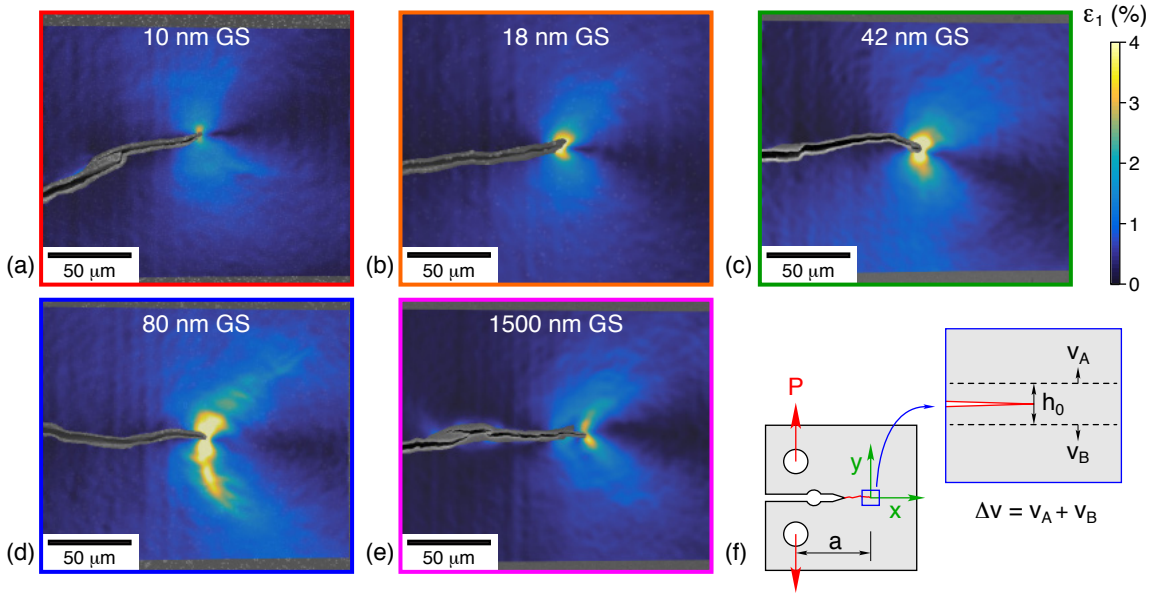


Figure 5.9: Microscale measurements of major principal strain fields $\varepsilon_1(x, y)$ from SEM-DIC are shown at the peak stress intensity K_{max} during a single fatigue cycle (performed between high ΔK and low ΔK measurements shown in Figure 5.5). Smaller regions of high major principal strain ($\varepsilon_1 \approx 4$ %) are observed around the crack tip for the 10, 18, and 1500 nm GS samples (a,b,e) than in the 42 and 80 nm samples (c,d), and the 1500 nm sample (e) exhibits the most microcracking. (f) The origin of coordinate axes (as used in Figure 5.10) is placed at the crack tip with positive x in the cracking direction. The relative crack displacement Δv indicates the elongation between the two lines symmetrically offset about the crack by $h_0 = 20$ μm.

One can observe in Figure 5.9 that the cracks in the 10 and 1500 nm samples actually consisted of multiple microcracks. While the 18 and 42 nm samples did not show evidence of microcracks (at least within this FOV), the cracks do show occasional slight kinks along their length. Only the 80 nm sample exhibited a smooth and continuous crack. Again, recall that the crack did not grow perpendicular to the loading axis in the 10 nm sample, making it a mixed mode fracture case and thus an outlier to the other cases.

The strain state around the crack tip exhibited unequal biaxial strain, with the minor principal strain (not shown) generally close to negative values of the major principal strain, $\varepsilon_2 \approx -\varepsilon_1$, for all grain sizes. This indicates that large in-plane shear strains were present.

A non-monotonic grain size dependence was also observed in the relative crack displacement profiles $\Delta v(x)$ shown in Figure 5.10a-d, and the stress intensity responses with respect to relative crack displacements shown in Figure 5.10e-g. The 1500 nm sample consistently

exhibited the lowest Δv. It also exhibited a bump in the Δv profile about 30 µm behind the crack tip that corresponded to a small branched crack (only visible upon very close inspection) in the SEM-DIC images (Figure 5.9). The 10 and 18 nm samples had the next lowest Δv after the 1500 nm sample. At 90 µm behind the crack tip, the 10 nm sample generally had about 0.1 µm lower Δv than the 18 nm sample, with the largest difference occurring during unloading (Figure 5.10e). The 42 and 80 nm samples had the highest and nearly coincident Δv profiles, with only a slightly higher Δv in the 80 nm sample just behind the crack tip ($-45\ \mu\text{m} < x < 0$). This small difference is also apparent in the CTOD responses shown in Figure 5.10g. Given the irregular ranking among the grain sizes, however, the $\Delta v(x)$ profiles generally grew during loading in a self-similar manner (Figure 5.10a-d).

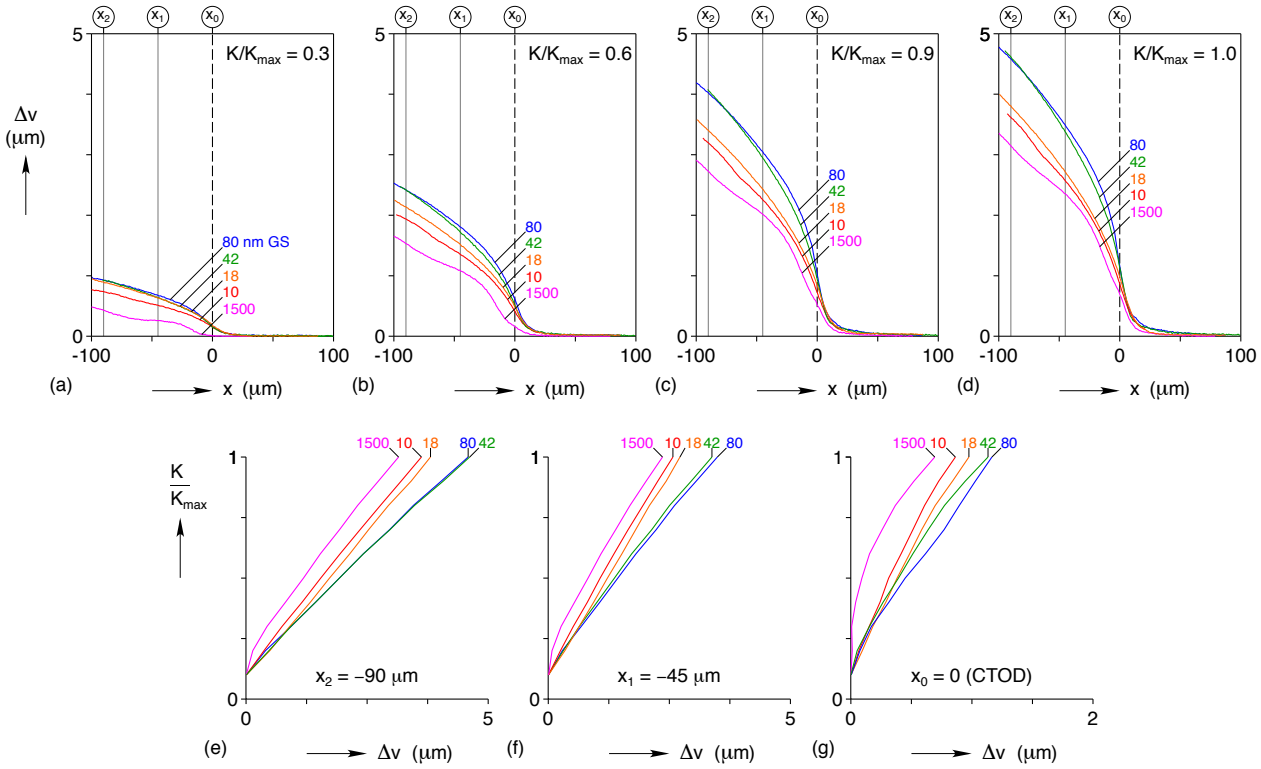


Figure 5.10: In-situ SEM-DIC (Figure 5.9) enabled crack opening responses to be characterized at the microscale, including: (a-d) relative crack displacement profiles $\Delta v(x)$ at selected stress intensities; and (e-f) stress intensity response curves at selected positions. For clarity, the x-axis scale for (g) is magnified compared to (e) and (f). The 1500 nm GS sample consistently exhibited the smallest Δv profile (a-d) and the largest crack opening level ($\text{K}_{\text{open}}/\text{K}_{\text{max}}$ between 30 and 40 %) at $\Delta v = 0$ (e-f).

The elevated crack opening level at $v(0) = 0$, shown in Figure 5.10e-g in the 1500 nm GS (Kopen/Kmax of 30-40 %) versus the others (Kopen/Kmax < 10 %) is consistent with the higher fatigue threshold noted previously ($\Delta\text{Kth} = 2.4\ \text{MPa}\sqrt{\text{m}}$ for the 1500 nm GS

and ΔKth = 1.9 MPa$\sqrt{\text{m}}$ for the others). Interestingly, however, the 10 and 18 nm samples exhibited smaller Δv (nearly 20 % smaller at K_{max} in Figure 5.10e) than the 42 and 80 nm samples, yet all had the same crack opening level.

In attempting to explain the reasons for the irregular dependence on GS in the fatigue crack growth results (Figure 5.5) and the microscale deformation results (Figure 5.10), we note that the Young's modulus of cold-rolled NC NiTi is known to depend on GS in a non-monotonic way [162]. Table 5.2 shows that a minimum exists in the austenite modulus (EA = 39 GPa) at intermediate grain sizes near 48 to 60 nm. Although not listed in the table, coarse-grained NiTi has an even larger Young's modulus, typically $E_A > 60$ GPa [83]. The cases (1500 and 10 nm) that had slow crack growths and small crack opening displacements have relatively large Young's moduli ($E_A = 50$ to ≈ 65 GPa), while the intermediate GS samples that had faster crack growths and larger crack opening displacements have smaller Young's moduli. The crack opening displacement data showed that the 80 nm GS sample had the largest Δv, although it was quite close to that of the 42 nm GS sample. While the correlation with elastic modulus is suggestive, we cannot conclusively say that the variations in elastic modulus are directly responsible for the trends in our fatigue results. The defect density increases monotonically with decreasing grain size, yet the amount of residual martensite increases with decreasing grain size and only exists in the smallest two GS samples. Perhaps, these competing factors are responsible for both non-monotonic trends in elastic modulus and fatigue performance, meaning both may be manifestations of the same microstructural factors.

Table 5.2: The Young's modulus (rolling direction) of nominally austenite NiTi has a non-monotonic relationship with grain size (data from [162])

Average GS (nm)	amorphous	10	24	38	48	60	80	100
Young's modulus (GPa)	54	50	48	43	39	39	45	55

Other complications exist in the superelastic behavior of NiTi that warrant mentioning. First, the macroscopic uniaxial tension behavior of coarse-grained NiTi exhibits unstable phase transformation that manifests as a nucleation peak near the onset, accompanied by strain (and phase) localization. This is followed by a stress plateau, during which transformation fronts (near discontinuities in strain) propagate along the length of the test sample. In other words, the strain field exhibits Lüders-like bands during phase transformation. These phenomena have been extensively studied in NiTi (see, for example, [131, 7, 133, 163, 164]), and the important conclusion is that the apparent uniaxial response is a 'structural' response, not a local stress-strain response at a material 'point', since the strain field is not uniform during the test. In fact, the actual stress-strain relationship has an up-down-up

character, although a special setup that suppresses strain localization is needed to reveal it [165]. The strong strain (and stress) gradients in the vicinity of a crack also likely suppress any such instabilities, and macroscopic phase fronts are not observed in fracture and fatigue experiments. One must be cautious in using uniaxial data to interpret structures with more complex stress fields. Nevertheless, we note that macroscopic uniaxial responses of NC NiTi (see Figure 2a in [44]) demonstrate a gradual progression with decreasing GS, from superelastic responses that exhibit distinct stress plateaus (1500 nm, 80 nm GS) to those that maintain a positive and successively larger tangent modulus (64 nm down to 10 nm GS) during phase transformation. This at least points to an increasing resistance to phase growth and an overall stabilizing effect on the mechanical behavior with grain refinement.

Second, phase transformation in NiTi involves enthalpy changes in the sample that can produce unusual sensitivities to loading rates. This is caused by a combination of interrelated factors: latent heat being released (or absorbed) during phase transformation, the strong thermomechanical coupling of transformation stresses (Clausius-Clapeyron relation), and the vagaries of the heat transfer environment in the test setup [131]. The overall load-unload response in (as-received) coarse-grained superelastic NiTi is a flag-shaped loop if conditions are isothermal (slow loading), yet a smoother loop with elevated transformation stresses and positive tangent moduli if conditions are adiabatic (fast loading). The strain rates that encompass these two extremes depend on the environment, thermal boundary conditions, and the size and geometry of the sample, according to heat transfer scaling laws. In stagnant room air with millimeter scale specimens, the transition between these two extremes occurs at strain rates between 10^{-5}/s to 10^{-2}/s [133], a range normally considered quasi-static for the testing of conventional metals. For NC NiTi, however, the enthalpy change, stress hysteresis, and Clausius-Clapeyron slope have been shown to decrease monotonically to almost zero with grain refinement from 1500 to 10 nm [44], meaning the overall thermomechanical coupling and rate effects are greatly diminished with decreasing GS. The macroscopic fatigue measurements involved relatively fast loading rates (10 Hz for high ΔK and 50 Hz for low ΔK), probably near adiabatic conditions; whereas, the microscopic measurements here were performed slowly, allowing the sample to stress relax and reach thermal equilibrium (isothermal conditions). Thus, one should recognize that any differences in loading rate may be important for the larger grain sizes, but not for the smaller grain sizes.

5.3.4 Fracture surfaces

SEM fractography revealed a distinctly rougher failure surface in the 1500 nm sample compared to the other grain sizes (Figure 5.11). The rougher fatigue fracture surface for the 1500 nm sample can promote crack closure [166], which is consistent with the 1500 nm

sample's larger crack opening level (K_{open}) compared to the other samples. The 1500 nm sample also exhibited small pockets of smooth fracture surfaces that were not observed in the other grain sizes. The samples with 10, 18, 42, and 80 nm grain size exhibited similar morphologies, with transgranular fatigue fracture surfaces and tear ridges in the crack growth direction. Overall, none exhibited the dimpled surfaces that one might see during tearing of more ductile metals, which is consistent with the generally low fracture and fatigue resistance of intermetallic alloys like NiTi.

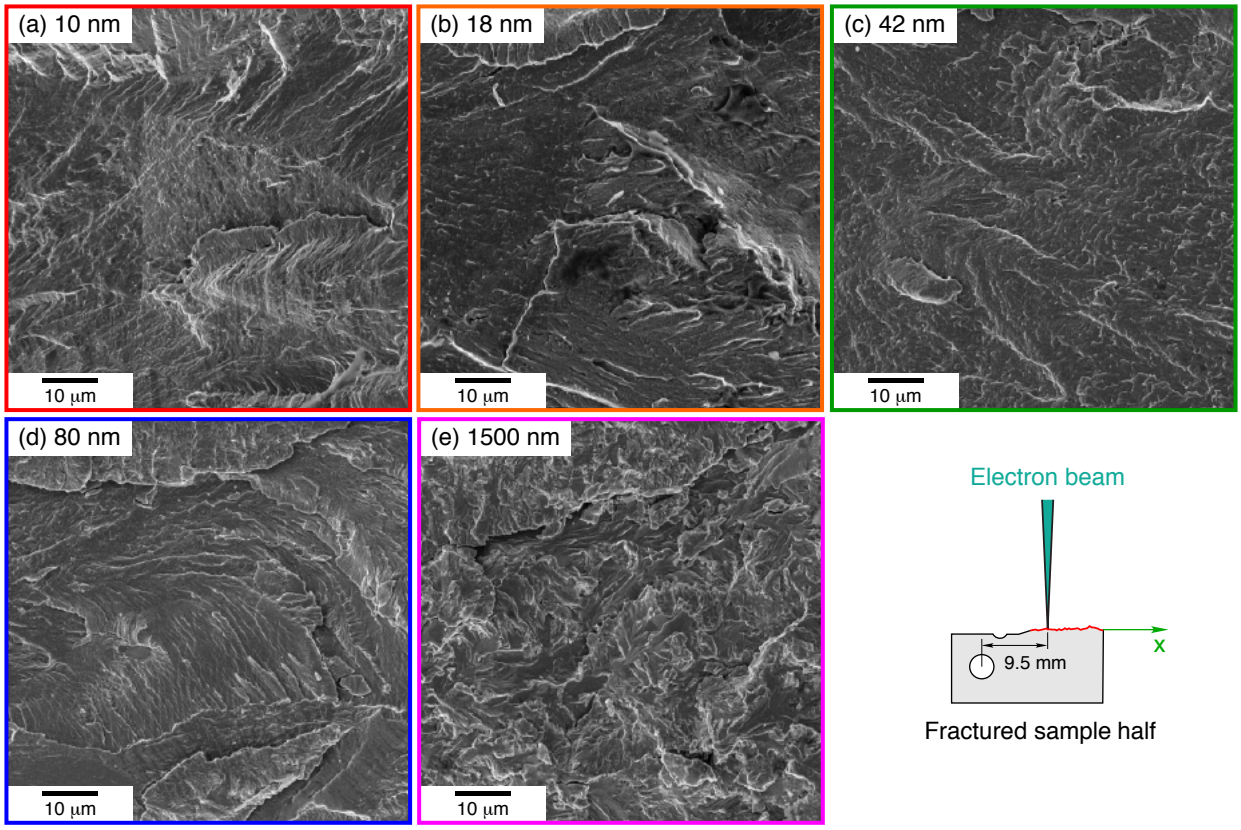


Figure 5.11: SEM micrographs of fracture surfaces were captured post-mortem for the five samples at $a = 9.5$ mm, corresponding to prior cracking at high stress intensity $\Delta K \approx 7$ MPa$\sqrt{m}$. In general, the fatigue fracture surfaces were transgranular, and tear ridges existed in the direction of crack propagation. Surface features in the 10, 18, 42, and 80 nm samples (a-d) were similar. The 1500 nm sample (e) exhibited the most roughness, consistent with expectations from the crack growth rate and crack opening displacement measurements.

5.4 Summary and conclusions

For the first time, the *structural* fatigue behavior of NC NiTi was characterized at low and high stress intensities, at the millimeter and micron length scales, and for test samples

with average grain sizes of 10, 18, 42, 80, and 1500 nm, with the following major results.

- High resolution SEM-DIC, using a novel external scan controller, largely eliminated pernicious drift, distortion, and scanning errors that otherwise require lengthy scan times (line or image averaging) and/or extensive corrections during image post-processing to achieve accurate results.

- Crack paths and deformation fields were measured in-situ during fatigue cycling at both the macroscale (by optical 3D-DIC) and microscale (by SEM-DIC) to connect the fatigue responses at disparate length scales. Optical DIC allowed crack growth to be accurately tracked, occasionally detecting anomalies in crack growth (crack branching in the 1500 nm GS sample and an oblique crack path in the 10 nm GS sample), as well as providing details of the strain field around the crack. SEM-DIC allowed displacement and strain fields to be examined in the local vicinity of the crack tip.

- Macroscopic fatigue performance correlated well with microscopic observations. The grain sizes with relatively fast macroscopic crack growth rates exhibited relatively large crack displacements at the microscale, and vice versa.

- The sample with the largest grain size studied (1500 nm) exhibited the slowest crack growth rates and largest threshold stress intensity at the macroscale, while exhibiting the minimum crack opening displacements, largest crack opening stress intensity level, and roughest fracture surfaces at the microscale. The sample with an intermediate GS (80 nm) generally exhibited the fastest crack growth rates and largest crack opening displacements.

- The structural fatigue behavior measured here did not follow the monotonic fracture trends previously reported on similar NC NiTi. The remarkable *functional* fatigue resistance previously reported for the smallest grain size samples (10 and 18 nm) also was not reflected in improved structural fatigue behavior (especially at low stress intensity). Trends in fatigue behavior among the grain sizes were non-monotonic, and the cause is as yet unknown. Various underlying mechanisms, including the presence of residual martensite (only in the 10 and 18 nm GS samples) and previously observed non-monotonic trends in Young's modulus with GS, were proposed and discussed.

CHAPTER 6

Texture dependence on functional and structural fatigue in NiTi shape memory alloy

6.1 Introduction

With repeated mechanical and/or thermal cycling, shape memory alloys exhibit two types of fatigue: structural fatigue and functional fatigue. *Structural* fatigue is the accumulation of damage that leads to crack initiation and growth. *Functional* fatigue is the decay of properties related to the phase transformation. Structural fatigue is common to other engineering materials, but functional fatigue is typically not (because most structural materials do not exhibit reversible phase transformation). Examples of functional fatigue include "ratcheting", or a gradual length increase during thermal cycling (shape memory effect), and "shakedown," or a reduction in the energy absorbed per during mechanical cycling.

This work builds upon the large body of literature on NiTi fatigue. Particularly, it combines measurements of texture, functional fatigue, and structural fatigue on the same material to extract conclusions about the coupled damage from tensile cycling (functional fatigue) and cyclic crack growth (structural fatigue). The methods build upon the multiscale approach developed in Chapter 5 for examining grain size effects, and then are later extended in Chapter 7 for studying temperature effects.

6.2 Introduction to texture dependence of NiTi in tension

The mechanical response of NiTi is highly anisotropic [42], both for polycrystalline materials [167] and single crystals [168]. Preferential textures commonly form in different production processes and heat treatments of NiTi [111, 112, 113]. The development of texture in NiTi has been experimentally observed from in-situ neutron diffraction. During compression, (100) texture formed parallel to the compression axis, and (110) formed perpendicular

to the compression axis [110].

Generally, NiTi has the highest transformation strain when the (111) and (110) crystal orientations are aligned with the tensile axis [42]. Coincidentally, most drawn wire/bar NiTi stock has a strong (111) fiber texture in the drawing direction [169, 42]. From tension tests of NiTi single crystals, the (111) plane developed large transformation strains (greater than 10 %), while the (100) and (110) single crystals fractured at $\approx$ 2 % and 7 % strain, respectively. Localized phase transformation was only apparent in the macroscopic stress-strain responses of the (110) and (111) single crystals [168].

Early work on the texture dependence of polycrystalline NiTi identified detwinning as an important deformation mechanism that changes with texture [167]. For the rolling direction (RD) of a sheet of NiTi, the martensite phase with $\langle 011 \rangle$ type II twins was partially or completely detwinned during deformation to 6 % strain. For the transverse direction (TD), detwinning was not observed. Rather, the TD sample had significant dislocation density in the $\langle 011 \rangle$ type II twins. Texture measurements suggested that RD was more favorable to the shear direction of $\langle 011 \rangle$ type II twins, while TD was more favorable to the shear direction of (001) compound twins [167].

The role of texture on NiTi deformation was advanced with full-field observations of deformation through the use of digital image correlation [164, 83, 84]. With increased strain rates, TD showed higher propensity to nucleate new bands of martensite on the macroscopic length scale. Also the bands of martensite in TD had lower strains relative to the bands in RD and 45, while the region of austenite in TD had *higher* strains relative to the austenite in RD and 45 [84].

6.3 Introduction to texture dependence of NiTi fatigue and fracture

Continuing from their previous work on single crystals [168], Gall, et al., showed that (111) NiTi crystals that are favorably oriented for martensitic transformation under tension are also oriented favorable to inhibit local cleavage fracture [169]. They observed cleavage on the (110) and especially the (100) planes. Thus, not only is the transformation favorable for NiTi wires/bars with (111) fiber texture in the drawing direction, but their ductility is high because their (100) and (110) cleavage planes are oriented unfavorably for local cleavage fracture [169].

Grain orientation and crystallographic texture influenced the path of fatigue cracks in superelastic NiTi, as evidenced by in-situ X-ray diffraction observations [170]. $A \rightarrow M$ transformation was not observed in certain grains within the transformation zone. However, there

was no clear trend for which specific grain orientations caused crack growth suppression. Furthermore, grain reorientation was not observed in the transformation zone around the crack tip (after $A \rightarrow M$ forward and $M \rightarrow A$ reverse transformation, the orientation of austenite grains did not change) [170].

From fracture experiments on single crystal NiTi, stable crack growth was observed in samples with the tensile axis along the [100] direction, while samples with [111] in the tensile direction had unstable crack growth. In perhaps the most beautiful photographs of any work on NiTi fracture, transformation in the [100] sample localized to the sides of the crack, while transformation in the [100] sample localized directly ahead of the crack [171]. Interestingly, although the [111] sample had higher ductility, it had much less resistance to fracture. For typical materials, increased ductility is associated with increased fracture toughness.

Recently, the fatigue crack growth of [012] single crystal NiTi was examined at both low and high temperature (as stable martensite and stable austenite, with the same composition but only at different temperatures) [172]. The effective elastoplastic stress intensity factor was estimated with a numerical procedure, similar to the correction from Irwin for modifying linear elastic fracture mechanics [173]. The fatigue crack thresholds were $\Delta K_{\mathrm{Ie,th}} = 1.4$ MPa$\sqrt{\mathrm{m}}$ for stable austenite, and $\Delta K_{\mathrm{Ie,th}} = 3.3$ MPa$\sqrt{\mathrm{m}}$ for stable martensite. Also, the local elastic properties of martensite changed significantly (from ≈ 45 GPa to ≈ 85 GPa effective modulus) and had to be factored into the effective stress intensity calculation [172].

Also recently, site-specific microstructures were connected with NiTi fatigue crack initiation and growth around pore defects [174]. The microstructure with the least fatigue resistance was the presence of large (110) grains with respect to the tensile axis within the area of stress concentration from the pores/holes [174].

6.4 Materials and methods

6.4.1 Superelastic NiTi sheet material

NiTi sheet ("BB" alloy, 50.8 at% Ni, ≈ 1 mm thickness, lot 22180-100815; Memry Corporation, Hartford, CT) was obtained with nominally superelastic properties at room temperature. From the manufacturer, the sheet was cold worked with rolling to approximately 30 % thickness reduction, flat annealed in air with a hot press furnace, and then pickled to remove rough surface oxides. The transformation temperatures were measured with differential scanning calorimetry (DSC; instrument model Discovery DSC 2500, TA Instruments, New Castle, DE) on a piece of the sheet that was cut with a low-speed diamond saw. The experimental method used helium purge gas and calculated the heat capacities directly after

calibration with sapphire with the following method:

1. equilibrate at 75 °C,
2. isothermal hold for 1 minute,
3. ramp at 10 °C/min to -150 °C,
4. isothermal for 1 minutes,
5. ramp at 10 °C/min to 75 °C, and
6. isothermal hold for 1 minute.

Figure 6.1 shows the heat flow response of the NiTi sample from DSC. The latent heat peak upon heating, for the $A \rightarrow M$ transformation, was at 0 °C. The austenite finish temperature was 3 °C, confirming that the material will exhibit superelasticity at room temperature.

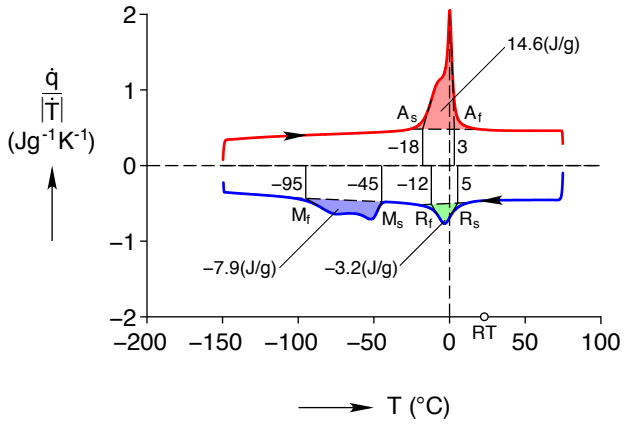


Figure 6.1: Differential scanning calorimetry (DSC) of the NiTi material showed the latent heat for the $A \rightarrow M$ transformation was at 0 °C, and the austenite finish temperature was 3 °C.

The microstructure of the material was examined with transmission scanning electron microscopy (TSEM) in a field-emission gun (FEG) scanning electron microscope (Teneo; Thermo Fisher Scientific, formerly FEI, Hillsboro, OR). A foil for transmission microscopy was prepared with focused-ion beam (FIB) liftout in SEM. The foil's surface was orthonormal to the rolling direction of the sheet. Nominally, grain sizes on the order of 20 to 200 nm were observed, although only a relatively small area of the sample was able to produce quality TSEM images, due to the thickness limitation for electron transparency with the relatively low beam accelerations in TSEM (30 kV used in this work). Nonetheless, the TSEM results revealed distinct grain boundaries for grain size estimates, as well as dislocations in the material (Figure 6.2). Using a sequence of tilts with TSEM images, dislocations can be found as distinct lines that appear at some tilts but disappear at others [175]. Upon tilting in a

sequence of 1, 2, and 3° from the orthonormal position to the beam, the distinct boundaries of a grain were visible at 1° tilt, but no dislocation lines were apparent. At 2° tilt, however, a dislocation line appeared in the grain (indicated by the yellow arrow). The dislocation extended into the neighboring grain to the right. Then at 3° tilt, the dislocation was still visible, although the grain appearance was darkened due to electron channeling contrast.

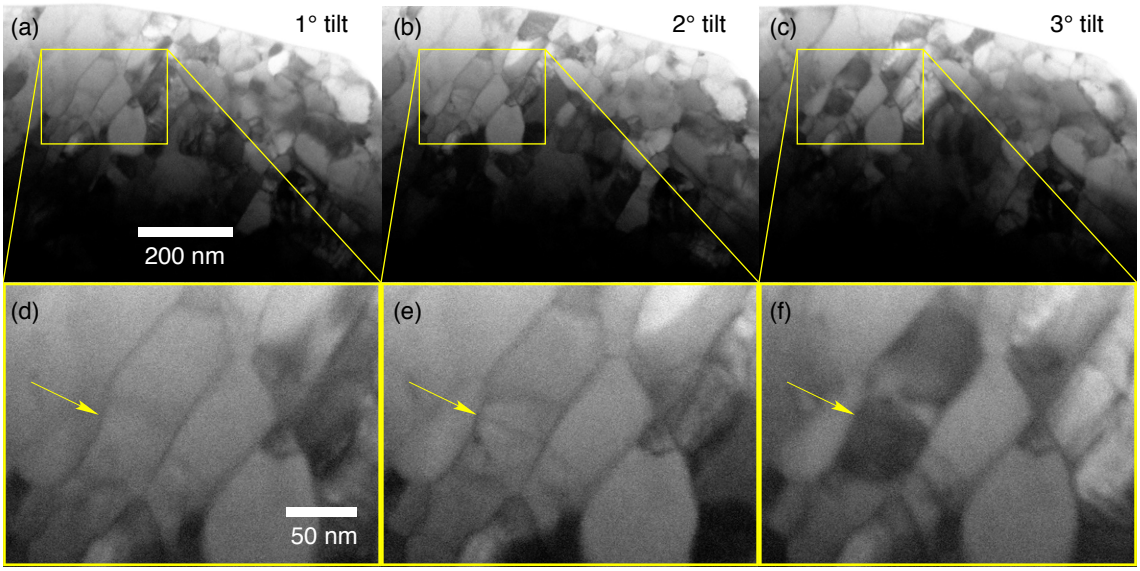


Figure 6.2: TSEM tilt sequence showing nanograined structure (grain sizes on the order of 20 to 200 nm), as well as dislocations present in the as-received material. The TSEM foil sample was prepared with the surface normal in the rolling direction of the bulk sheet material. During a tilt sequence of 1, 2, and 3° (a-c), many grains changed brightness due to electron channeling contrast. Also, a dislocation that was not visible at 1° tilt (d) appeared at 2° tilt (e), indicated by the yellow arrow. The dislocation extended across at least two grains, and is particularly visible in the second grain in the 3° tilt (f).

On the same foil sample from the TSEM imaging (Figure 6.2), transmission Kikuchi diffraction (TKD) was also attempted for further characterizing the material's microstructure. TKD is performed in SEM with the same detector as electron backscatter diffraction (EBSD), although TKD uses transmission electrons that have passed through a thin, electron-transparent sample. For as-received NiTi with cold working, such as the sheet material in the present work, the grain sizes are generally too small for EBSD measurements. However, TKD can have spatial resolutions as small as 10 nm [176, 177]. TKD has been attempted before for NiTi, although with limited results, especially for sub-micron grains [99]. Our TKD measurements (Figure 6.3) were largely unsuccessful, with low indexing confidence for most points along the sample. Nonetheless, the TKD measurements indicated the presence of micron-sized grains. Combined with the TSEM images (Figure 6.2), these results

indicate a broad distribution of grain sizes from tens of nanometers to about 1 µm.

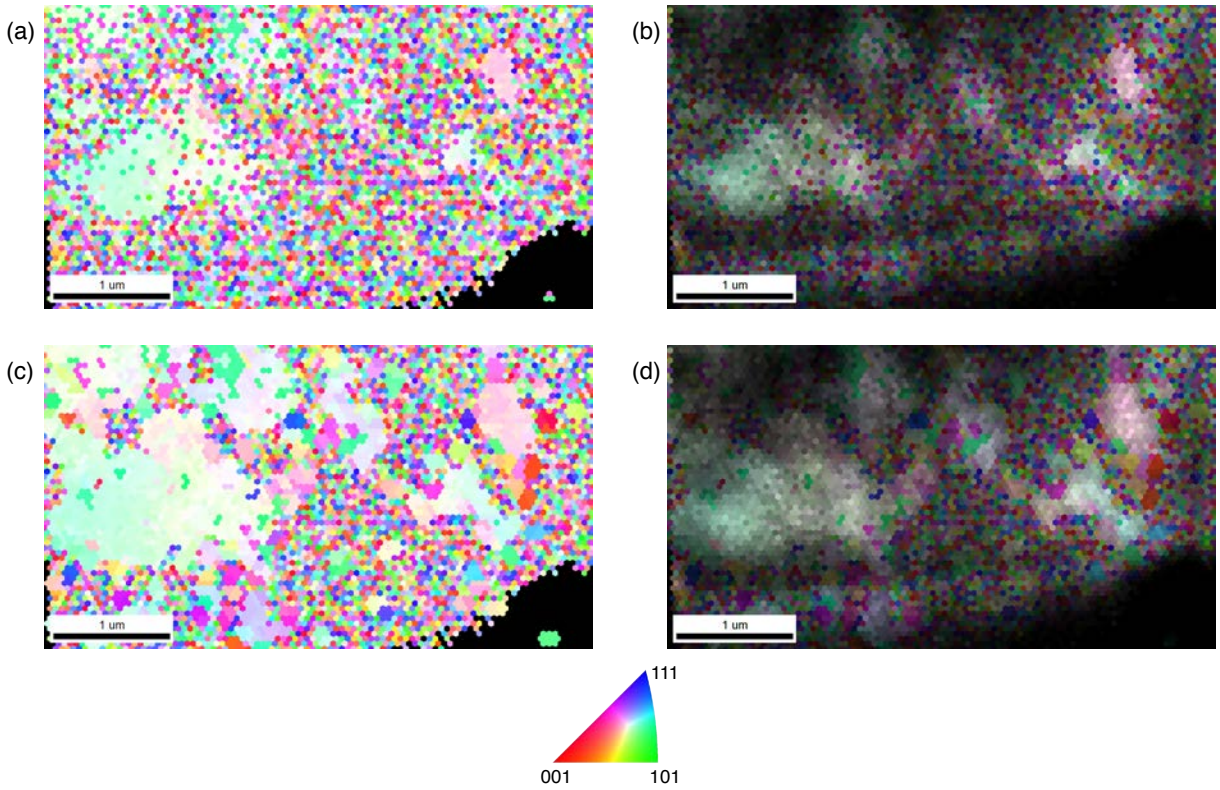


Figure 6.3: Transmission Kikuchi diffraction (TKD) revealed micron-sized austenite grains, although most of the region of interest had poor diffraction quality. The raw diffraction pattern is shown in (a). The raw diffraction pattern is shown again in (b), except with each data point's brightness proportional to the diffraction confidence index (CI). Likewise, (c) and (d) are the diffraction pattern with a grain dilation filter, and (d) has brightness of each data point scaled to CI.

The crystallographic texture of the NiTi material was characterized with X-ray diffraction (XRD). For a full introduction to texture measurements for NiTi, see Chapter 4. A piece of the sheet material (20 × 20 mm) was cut with low-speed diamond sawing, then ground and polished with a progression of 600, 800, and 1200 grit silica papers with flowing water. The final polishing step was a 5:1 volume ratio mixture of colloidal silica (Buehler MasterMet 2, 0.02 µm) and 30 % hydrogen peroxide (Fisher H325) on a microfiber cloth (Buehler MicroCut). The combined mechanical and chemical polishing produced a mirror finish to aid diffraction for the NiTi material.

In-plane pole figures were measured with a Rigaku SmartLab XRD for the (110), (211), and (200) poles for both the sheet material and for a powder sample (NiTi B2 cubic powder, $Ni_{51}Ti_{49}$; Special Metals, New Hartford, NY; lot 3601). The powder sample measurements were used for defocusing correction on the sheet sample measurements to enable calculating

a quantitative orientation distribution function (ODF). The XRD conditions were 40 kV and 44 mA source, parallel beam (PB) collimator slit, 0.5° in-plane parallel slit collimator (PSC), 10 mm length limiting slit, and 0.5° in-plane parallel slit analyzer (PSA). The attenuator was open, the divergence slit was 1 mm, and the Soller slit and receiving slit were 2 mm. The maximum X-ray penetration in NiTi at the experiment conditions was calculated to be 19 µm (Section 4.3.3).

Figure 6.4 shows the pole figure measurements, the fitted ODF, and the error between the ODF and the measurements. Relatively strong texture was observed in both the rolling direction (RD), and transverse direction (TD) of the sheet, with several regions having multiples of uniform distribution (MUD) of about 10. Although the sample was polished to be as flat as possible, and the sample and XRD optics were carefully aligned, there was asymmetry in the transverse direction of the sample. This asymmetry may be due to processing of the sheet material. Nonetheless, for calculating the ODF, orthorhombic sample symmetry was assumed because this symmetry is common for rolled metal sheets.

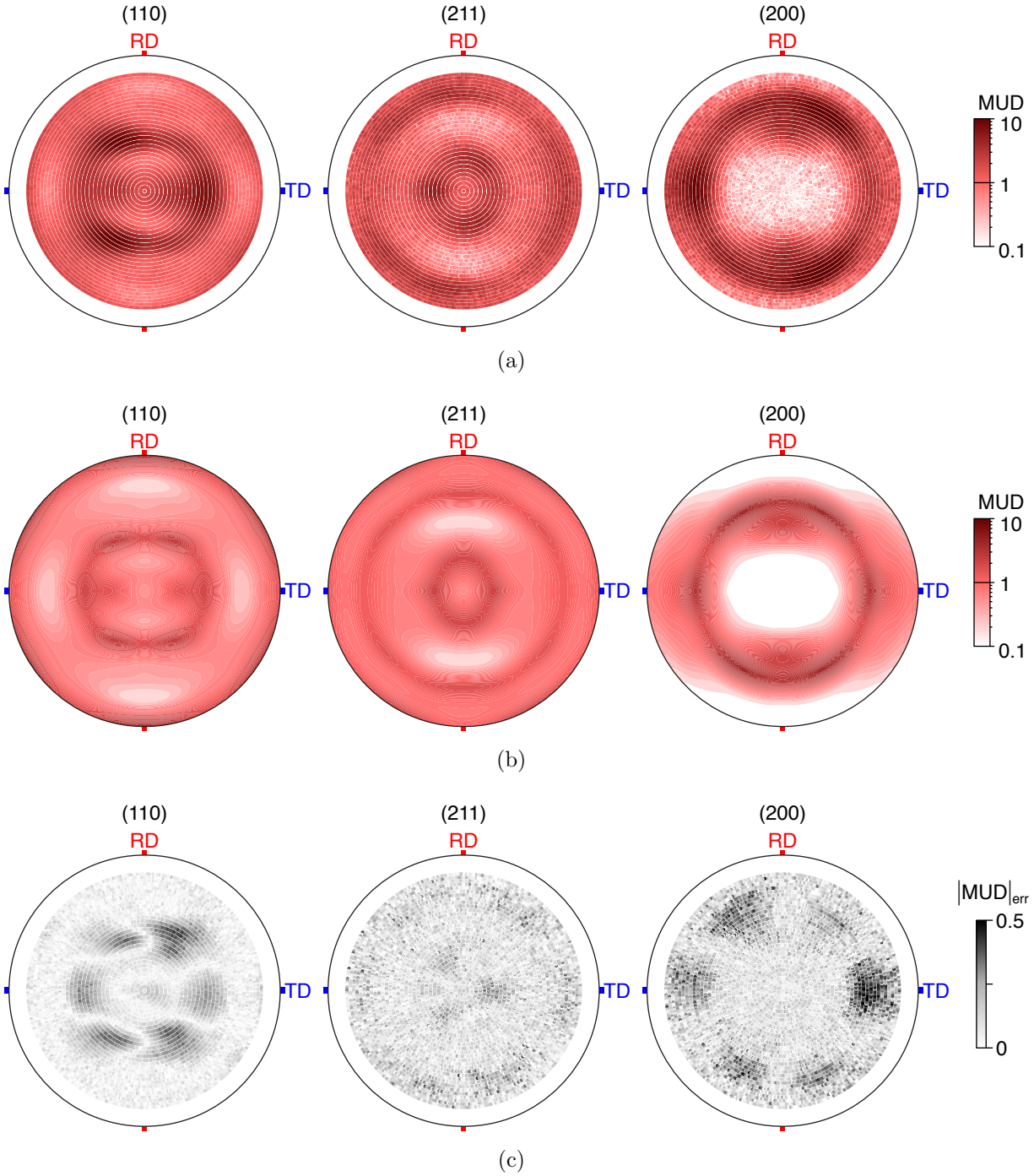
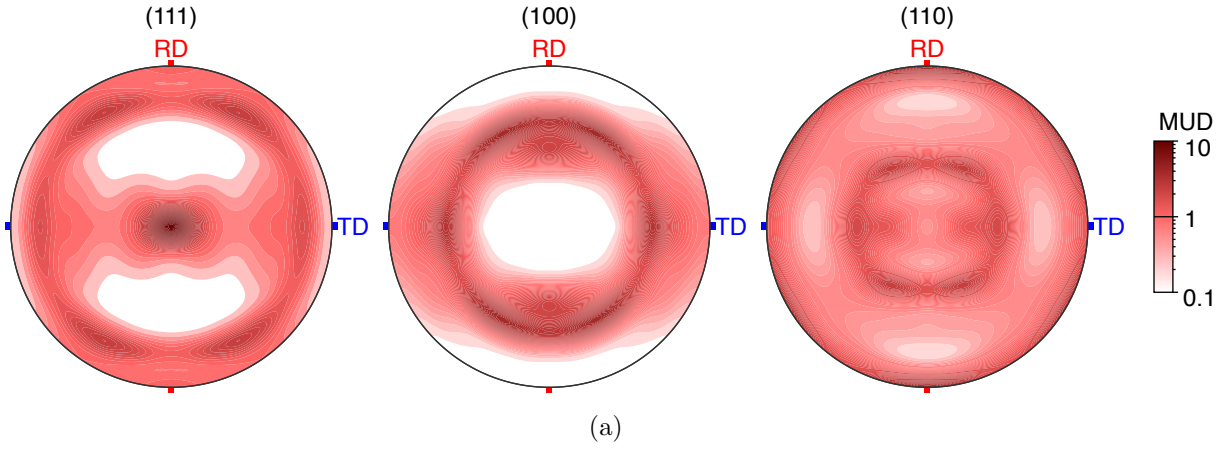


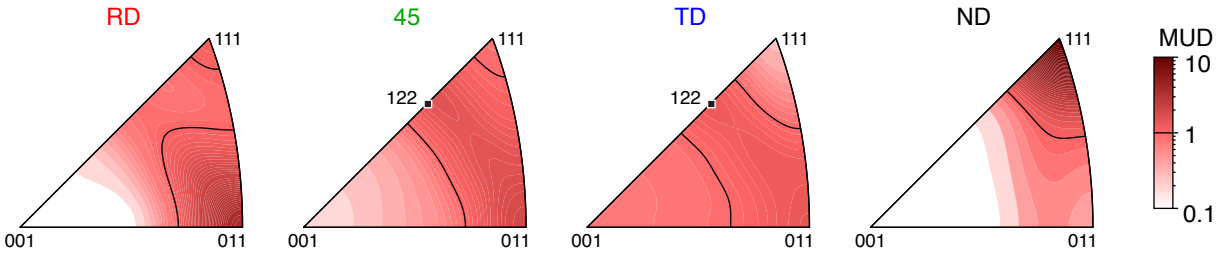
Figure 6.4: (a) The measured pole figures. (b) The fitted orientation distribution function. (c) The error in ODF, calculated as the absolute value of the difference between the pole figure measurements and the fitted ODF.

Figure 6.5 shows the pole figures and inverse pole figures (IPF) for the (111), (100), and (110) poles. For the (111) pole figure in Figure 6.5a, more texture was observed in RD than TD, indicated by the greater variation along the vertical direction (corresponding

to RD) of the pole figure compared with the horizontal direction (corresponding to TD). This observation is paralleled by higher MUD in the IPFs in Figure 6.5b. In (100), slightly more texture was observed in TD than for RD, while in the (110), slightly more texture was observed in RD than for TD. The strongest texture by far, however, was (111) in the normal direction (ND), shown at the center of the (111) pole figure in Figure 6.5a, and the upper right corner of Figure 6.5b. This (111)//ND texture is commonly called the γ-fiber texture, and is often observed in rolled sheets of BCC metals. Furthermore, the γ-fiber texture has been observed in other NiTi sheets [120, 40].



(a)



(b)

Figure 6.5: (a) The orientation distribution function (ODF) at the (111), (100), and (110) poles. (b) Inverse pole figures for the NiTi sheet's rolling direction (RD), 45 ^{o} from the rolling direction (45), transverse direction (TD), and normal direction (ND). The strongest texture was (111) in the normal direction (MUD=9.4), and the second strongest texture was (110) in the rolling direction (MUD=3.7). The maxmima for 45 and TD were MUD=2.0 and 1.4, respectively.

When the orientation distribution function was visualized as ϕ_2 slices (Figure 6.6), following the convention of texture analysis of metals, there was a clear (111)//ND fiber (the γ-fiber).

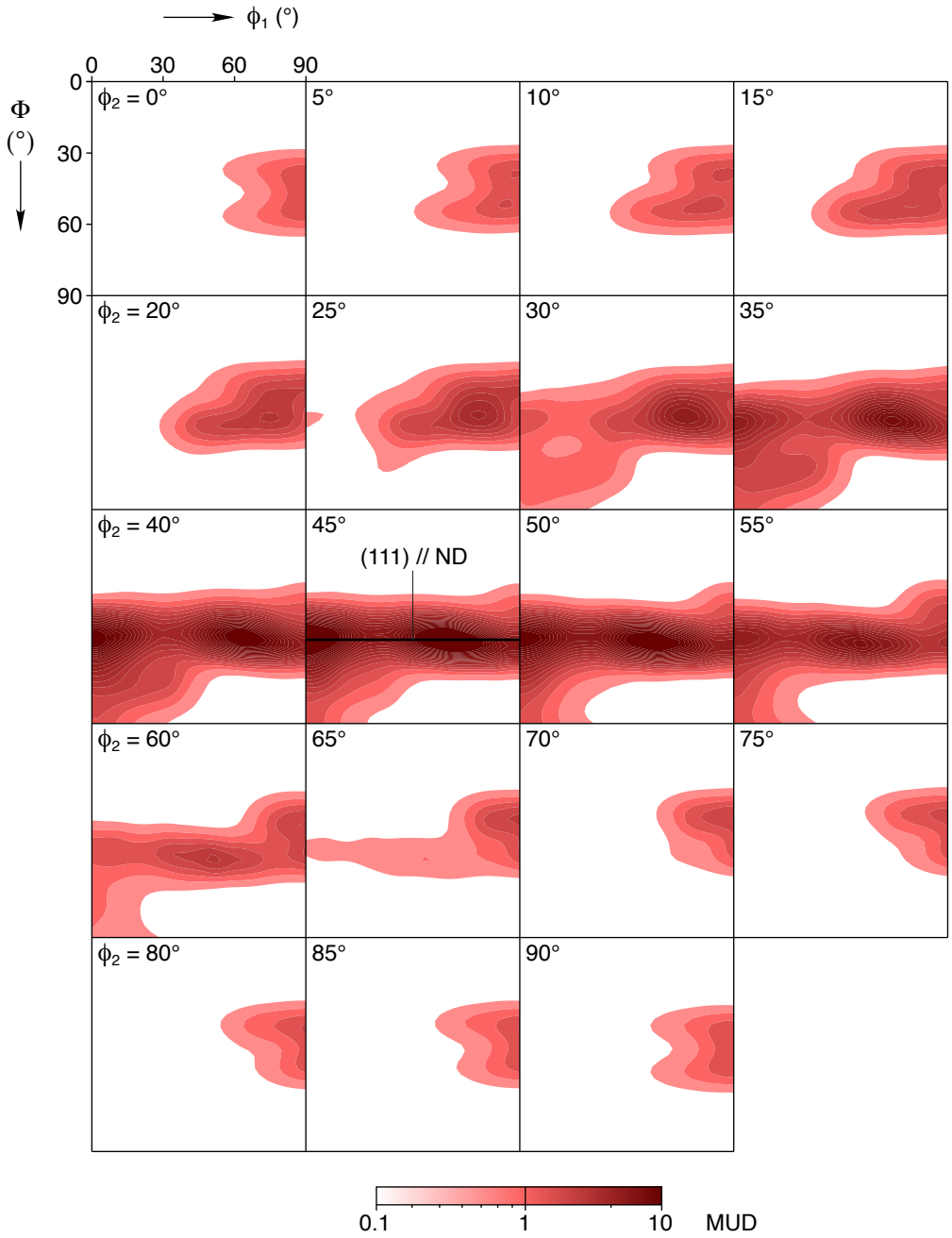


Figure 6.6: A strong γ-fiber texture, (111)//ND, was observed in the orientation distribution function with respect to Bunge Euler angles (defined in Figure 4.1b) with constants ϕ_2 slices.

6.4.2 Macroscopic tension experiments

Uniaxial tension and compact tension samples were produced from the NiTi sheet with wire electrodischarge machining with dimensions shown in Figure 6.7. Samples with the rolling direction parallel to the tensile axis are denoted as RD. The compact tension samples

are approximately plane stress conditions, since they are relatively thin (nominally 1 mm before polishing).

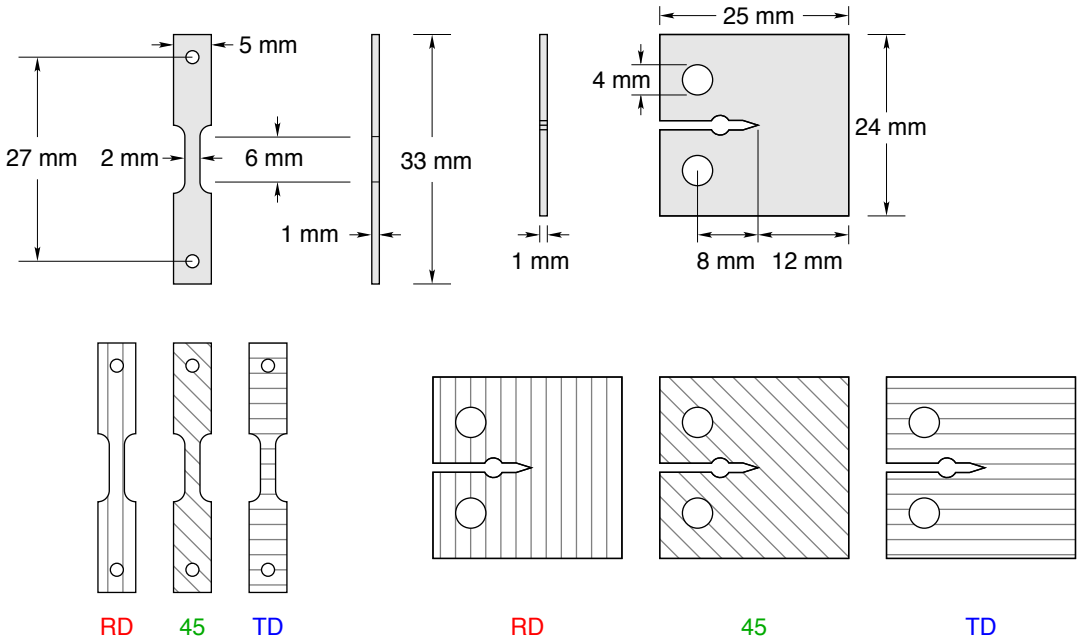


Figure 6.7: Uniaxial tension samples (left) and compact tension samples (right) were produced with the dimensions shown from the NiTi sheet with three different sample directions: tensile axis parallel to the sheet rolling direction (RD), tensile axis 45° from the sheet rolling direction (45), and tensile axis perpendicular to the sample rolling direction (TD).

The macroscopic uniaxial tension responses (Section 6.5.1) were measured with a 200 kN mechanical testing system equipped with a 5 kN load cell (Instron 5585 system with Instron load cell, Norwood, MA), shown in Figure 6.8. The samples were loaded in displacement control with a constant nominal strain rate of 1×10^{-4} 1/s for 25 load-unload cycles. The cycle limits on loading were 450 MPa, and on unloading were 4.5 MPa.

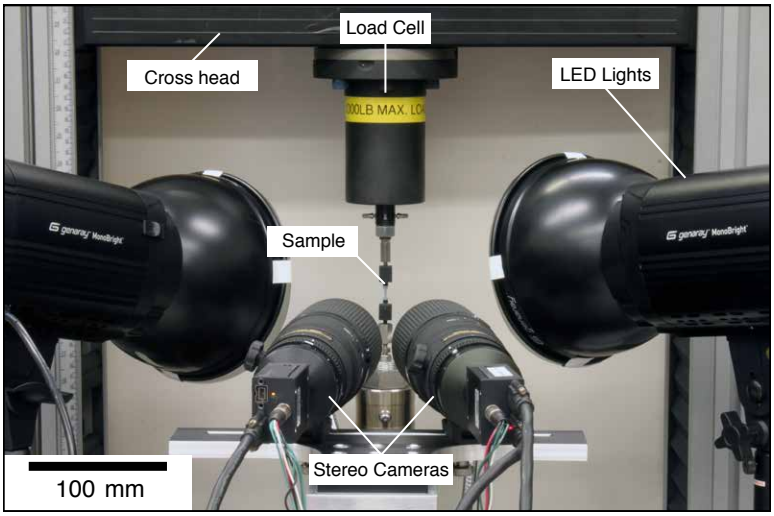


Figure 6.8: The experimental setup used for uniaxial tension experiments.

6.4.3 Macroscopic fatigue cracking experiments

The macroscopic fatigue cracking responses (Section 6.5.2) were measured with an electrodynamic voice-coil fatigue system (ElectroForce LM1, 220 N dynamic cycling capacity, TA Instruments). Two regimes of fatigue cracking were explored: one with a relatively high stress intensity rise ($\Delta K \geq 8.0$ MPa$\sqrt{m}$) for crack initiation and early crack propagation, and one with a relatively low stress intensity ($\Delta K < 8.0$ MPa$\sqrt{m}$) to probe the fatigue threshold and later crack propagation behavior. The cyclic rate for the high stress intensity regime was 10 Hz, and for low stress intensity was 50 Hz, following Ref. [5]. For both the low and high stress intensity regimes, the stress ratio was $R = K_{min}/K_{max} = 0.1$. The compact tension samples were polished with a sequence of 600, 800, and 1200 grit silica papers with water, then final polished to a mirror finish with a 5:1 volume ratio mixture of colloidal silica (Buehler MasterMet 2, 0.02 µm) and 30 % hydrogen peroxide (Fisher H325) on a microfiber cloth (Buehler MicroCut).

6.4.4 Macroscopic digital image correlation

For both the uniaxial tension and fatigue cracking experiments, the sample deformation was measured with a 3-D digital image correlation (DIC) system comprised of two cameras (PointGrey GRAS 50S5M-C) with high-quality lenses (Nikon Nikkor 200 mm, f/16), and two LED lights (Genaray MonoBright). The DIC setup utilized cross polarization, with linear polarizing light filters placed orthogonally on the lights and lenses. Cross polarization improves DIC accuracy by increasing contrast and eliminating saturated pixels [3]. A fan was placed near the setup to blow air gently past the sample to mitigate heat waves that can

be detrimental for DIC measurements [95]. The samples were speckle patterned using an Iwata CM-B airbrush with a basecoat of white paint (Golden High Flow Titanium White, no. 8549-4, lot 366162) and then a speckle pattern of black paint (Golden High Flow Carbon Black, no. 8040-4, lot 345307). This sequence of paints provides greater contrast than the converse sequence because black pigments are more effective at maintaining their contrast over white pigments, since black pigments absorb light and white pigments scatter light [4].

For optical 3-D DIC measurements, the images were correlated with Vic3D 7 with low-pass filtering, Gaussian subset weights, eight-tap interpolation, and the normalized squared differences method. The correlation thresholds were 0.02 px consistency threshold, 0.05 px maximum confidence interval, and 0.1 px matchability.

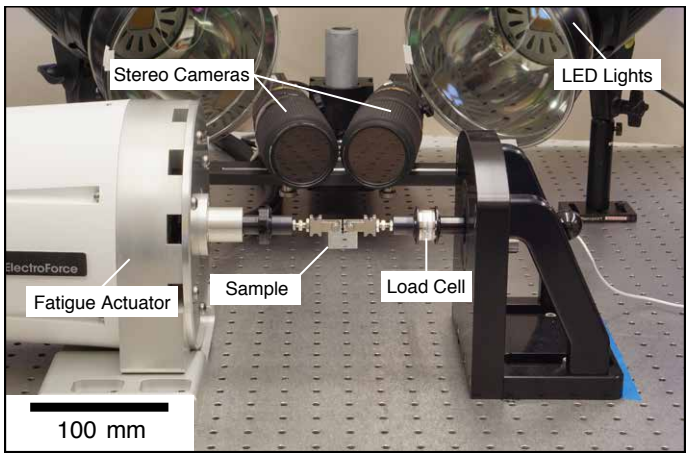


Figure 6.9: The experimental setup used for fatigue cracking experiments.

6.4.5 Microscopic SEM-DIC experiments

Finally, the microscopic deformation around the crack tip (Section 6.5.3) was measured with digital image correlation in a scanning electron microscope (SEM-DIC) for three of the compact tension samples (one each of RD, 45, and TD). The samples for SEM-DIC were fatigue cycled with an initial rise in stress intensity of $\Delta K = 8.0$ MPa$\sqrt{\text{m}}$, based on the initial notch length $a_0 = 8$ mm, until fatigue cracks grew to $a \approx 10$ mm. The painted speckle pattern for optical DIC was removed with methanol, the samples were final polished (again with 5:1 volume ratio mixture of colloidal silica and hydrogen peroxide), and then the samples were speckle patterned for SEM-DIC. The procedure for self-assembled Au nanoparticle patterning from Ref. [58] was followed, with the following exceptions: instead of hydroxylation in a base, the final polishing step with hydrogen peroxide and colloidal silica provided sufficient hydroxyl groups on the sample surface; instead of 24 hours for silanization, only 4 hours were needed; and finally, instead of preparing Au nanoparticles following the

method from Ref. [178], off-the-shelf Au nanoparticles were applied (250 nm, in 0.1 mg/m,L sodium citrate with stabilizer; Alfa Aesar, Ward Hill, MA).

After Au nanoparticle patterning, the compact tension samples were loaded in an in-situ tension/compression stage, shown in Figure 6.10 (10 kN capacity and equipped with 500 N load cell, Kammwrath & Weiss, Dortmund, Germany). The in-situ stage was operated in a field-emission gun scanning electron microscope (Teneo; Thermo Fisher Scientific, formerly FEI). The beam conditions were 30 kV and 0.8 nA, with a working distance of 18 mm. Images from the SEM were recorded with a custom scan controller that significantly reduces the scanning errors that are problematic for SEM-DIC [179]. The details of this method have been recently developed and are detailed further in Ref. [179]. The added precision from this custom scan controller was important for capturing the sub-20 nm differences in the crack separations among the three sample conditions, as presented in Section 6.5.3. The scan controller was programmed to raster "out and back" along the same line of pixels twice (once to the left, and back to the right), then average the two scan lines together for each line in the image. The effective dwell time, with four samples read from the secondary electron detector for each pixel and two pixels averaged together from the out-and-back scan, was 6.4 µs. The images were 300×300 µm and 4096×4096 px for a pixel size of 73 nm/px.

For SEM-DIC measurements, the images were correlated with Vic2D 6 with low-pass filtering, Gaussian subset weights, eight-tap interpolation, and the normalized squared differences method. The correlation thresholds were 0.02 px consistency threshold, 0.05 px maximum confidence interval, and 0.1 px matchability.

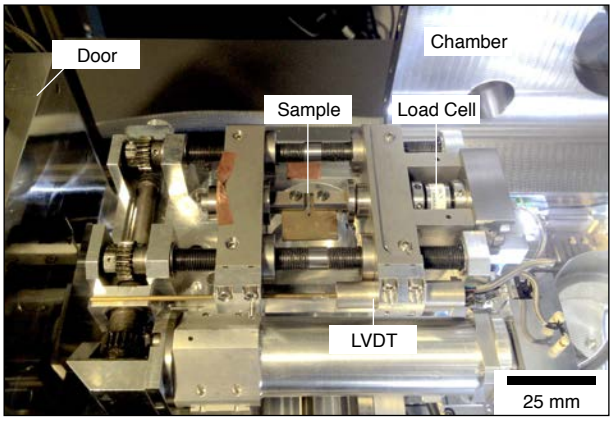


Figure 6.10: A compact tension sample in the in-situ tension/compression stage for SEM-DIC measurements of the microscale deformation at the crack. Although the sample was final polished to a mirror finish and NiTi has a silver appearance when polished, there is an orange tinge on the sample from the self-assembled Au nanoparticle pattern with relatively large nanoparticles (250 nm).

6.5 Results and discussion

6.5.1 Cyclic uniaxial tension

The functional fatigue response of the superelastic NiTi material was investigated with 25 load-unload cycles in tension (Figure 6.11). Strain ($\delta e/L_e$) was measured from DIC virtual extensometers across the gauge section. All three sample conditions (RD, 45, and TD) exhibited the typical tensile response of superelastic NiTi, with an initial elastic loading up to $\approx$ 0.5 % strain, followed by a nucleation peak from localization of the phase transformation, then a relatively constant stress while the phase transformation propagates along the sample gauge section (Figure 6.11a). The phase transformation's propagation completed much sooner for the TD sample, at around 4 % strain for the first cycle, compared with about 6 % strain for the first cycle of the RD sample. Also, there were differences in the region of increasing stress at the end of the loading stress plateau. For the TD sample, there was a more gradual upwards slope as the stress-induced martensite "locked up." The differences between the loading plateaus became more pronounced after 25 cycles. At $n = 25$, the loading plateau for the TD sample no longer had a flat region from propagation of the phase transformation, while the RD sample continued to exhibit localization and propagation of the phase transformation for all 25 cycles. This observation is consistent with the axial strain fields from DIC, described further in Figure 6.13. The stress plateau on loading (Figure 6.11b) was highest for TD (415 MPa for $n = 1$) and lowest for RD (337 MPa for $n = 1$), with the loading stress plateau for 45 approximately splitting the difference (380 MPa for $n = 1$). During the 25 load-unload cycles, the plateau stresses on loading ($P/A_{0,mid}$) consistently decreased at approximately the same rate for all three conditions (3.2, 3.1, and 2.5 MPa/cycle for RD, 45, and TD, respectively, on average for the 25 cycles).

Small amounts of residual strain (ε_r) accumulated with each cycle, with TD accumulating significantly less residual strain than RD and 45 (Figure 6.11c). The residual strains were fitted with respect to the number of cycles (n) using an exponential function with three fitting terms,

$$\varepsilon_r = B_1 - B_2 \exp\left(-n/B_3\right). \tag{6.1}$$

For the RD, 45, and TD conditions, respectively, the coefficients of determination (R^2) for the fits were 0.9987, 0.9938, and 0.9986. Extrapolating the asymptotes of the fits to the limit as $n \to \infty$, TD had one third of the residual strain (0.37 %) compared with RD (1.11 %), while 45 was a bit lower than RD (0.96 %). The rate of residual strain accumulation per cycle (ε_r') was also calculated from the fits (Figure 6.11d). For RD and 45, ε_r' at the first cycle

was about 0.1 %/n, while the value for TD was only about 0.04 %/n. All three conditions reached a residual strain accumulation rate of < 0.02 %/n with the 25 cycles.

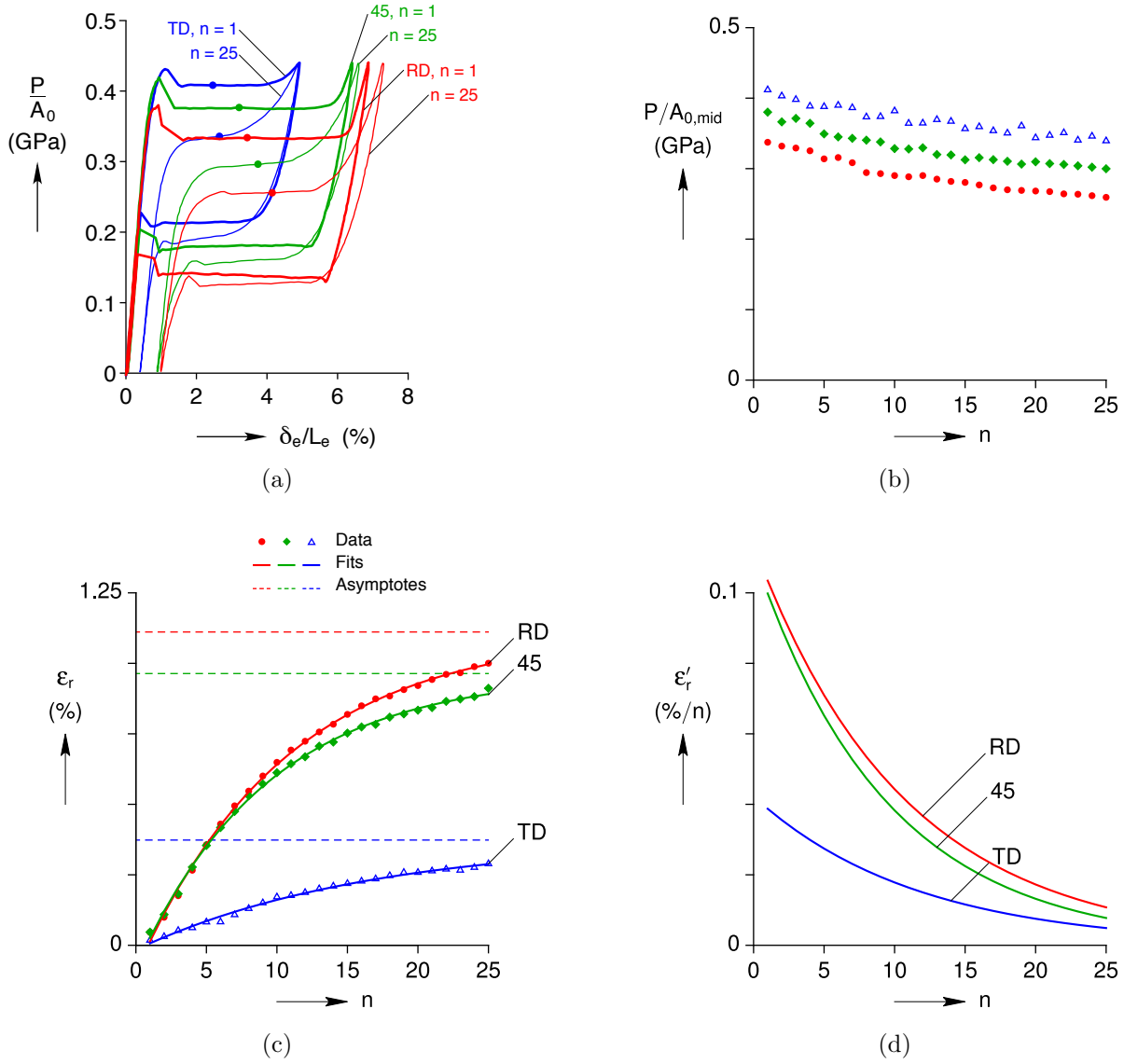


Figure 6.11: (a) The first ($n = 1$) and last ($n = 25$) cycles in tension for the RD, 45, and TD conditions. The plateau stresses on loading ($P/A_{0,mid}$, for Figure 6.11b) are indicated by dots. (b) The cycle-wise values of the plateau stresses on loading ($P/A_{0,mid}$) from Figure 6.11a. (c) The residual strain (ε_r) data, fits, and asymptotes with respect to cycle number (n). (d) The ratcheting rate (ε'_r) with respect to cycle number (n).

From the cyclic tension responses, the cycle-wise strain work was also measured for each condition. The cycle-wise strain work was considered positive during loading (W^+), and

negative during unloading (W^-). The sum of loading and unloading works gives the cycle wise net work, with $W_{net} = W^+ + W^-$. The magnitudes of strain work per cycle were on the order of about 10 MJ/m^3. For each cycle, the RD and 45 samples had nearly identical W^+, but 45 had slightly lower W^-. TD had lower strain work, especially in the first four cycles. By $n > 20$ cycles, the W_{net} for TD was only about 1 MJ/m^3 less than that of RD and 45.

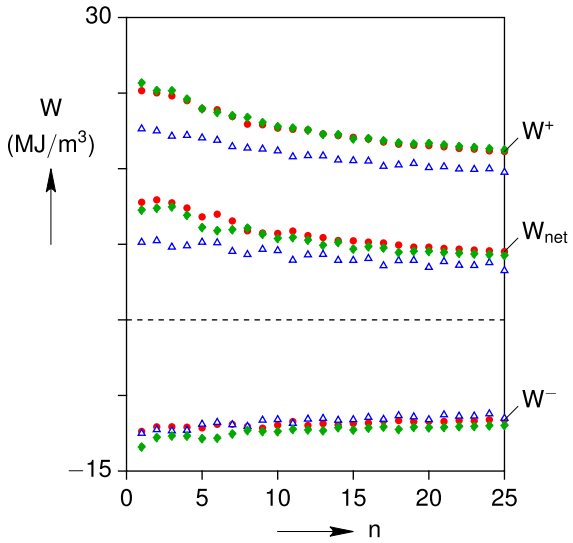


Figure 6.12: The cycle-wise strain work for 25 load-unload cycles in the three sample conditions, with positive strain work during loading (W^+), negative during unloading (W^-), and the cycle wise net work, with $W_{net} = W^+ + W^-$.

Figure 6.13 shows the axial strain fields from 3-D DIC for the images corresponding to the phase transformation's nucleation and coalescence. Cycles 1, 2, 3, 5, 10, 15, and 25 are shown with the axial strain maps during the n^{th} loading cycle (with the phase transformation's nucleation) indicated by "n^+", and during the n^{th} unloading cycle (with the phase transformation's the) indicated by "n^-". Unfortunately, the time resolution of the images was relatively coarse (only about 7,000 image pairs were captured during each samples 25 cycles with a frame rate of 0.1 Hz), so the comparison of the magnitudes of strain among the conditions is limited.

Nonetheless, an interesting trend was found for the location of the phase transformation's nucleation and coalescence. For the RD and 45 samples, the phase transformation nucleated and coalesced in the same spot for each cycle. However, for the TD sample, the phase transformation did not nucleate or coalesce in the same spot among cycles. Furthermore, the location of coalescence for one cycle was often different than the location of the nucleation for that same cycle. Also, the RD and 45 samples continued to have a clear spot of localization

followed by propagation in the last cycles, while the TD sample no longer had distinct localization during loading (but it still localized during unloading). This result is consistent with the engineering stress-strain responses for the three conditions: RD and 45 maintained transformation plateaus, while the plateau in TD dissipated with cycling (Figure 6.11a). Also, the lack of localization in the transverse direction of NiTi sheet has been observed before [167].

A previous report on the texture-dependence of the tensile responses of superelastic NiTi showed that phase transformation tends to localize in the same spot from cycle to cycle [83, 84]. This result is consistent with the RD and 45 samples in the present work, but inconsistent with the TD sample. For the TD case in Ref. [84], TD had stronger cycle-to-cycle similarity in the phase transformation localization. Furthermore, Ref. [84] observed small branches of phase transformation stemming from the primary phase transformation front, but these branches were not observed in the present work. The reason for this is likely the large thickness-to-width ratio of the gauge section in the present work's samples.

One possible explanation for these differences could be the significantly different sample geometries: in Ref. [83], the samples had relatively long and thin gauge sections (12.5 mm long, 3.125 mm wide, 0.254 mm thick). In the present work, the gauge sections are relatively short and thick (6 mm long, 2 mm wide, and 1 mm thick). Sample geometry (specifically, the thickness-to-width ratio) is known to alter the characteristics of the nucleation and propagation of superelastic NiTi's phase transformation [180, 181].

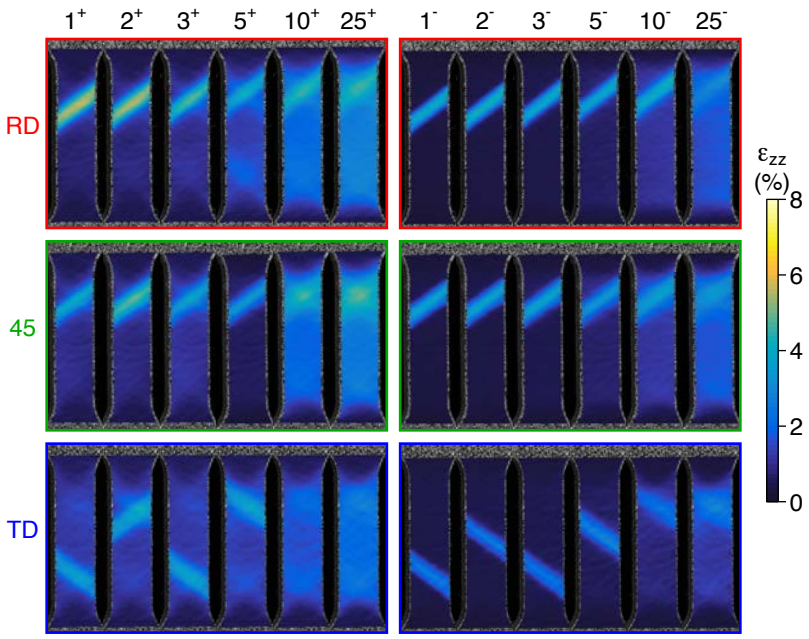


Figure 6.13: Maps of axial strain (ε_{ZZ}) are shown from the 3-D DIC images nearest the nucleation and coalescence of phase transformation for cycles $n = 1$, 2, 3, 5, 10, and 25 for each condition during loading (n^+) and unloading (n^-).

6.5.2 Cyclic fatigue crack growth

Cyclic fatigue experiments were performed on the compact tension samples, and the crack growth rates were measured with optical 3-D DIC. During the fatigue experiments, images were captured with the optical 3-D DIC system during < 1 second pauses at the minimum and maximum cyclic force. The crack tip position was measured from the DIC results, using either the correlation confidence interval (CCI) for the experiments with relatively high stress intensity ($\Delta K \geq 8.0$ MPa$\sqrt{\text{m}}$), or the major principal strain for relatively low stress intensity ($\Delta K < 8.0$ MPa$\sqrt{\text{m}}$). Although the crack is typically removed from the area of interest (AOI) for DIC, this crack tip location method requires keeping the crack in the DIC AOI. Further details about this method are outlined in Section 5.2.2.

To illustrate the efficacy of DIC for tracking crack tips, Figure 6.14 shows the crack path measured from DIC, with a semi-transparent overlay of the crack path (to scale) imaged with optical microscopy after fatigue cracking (and after the speckle pattern was removed). For the RD, 45, and TD samples, DIC captured both the path of the crack and the final position of the crack tip to within 50 µm.

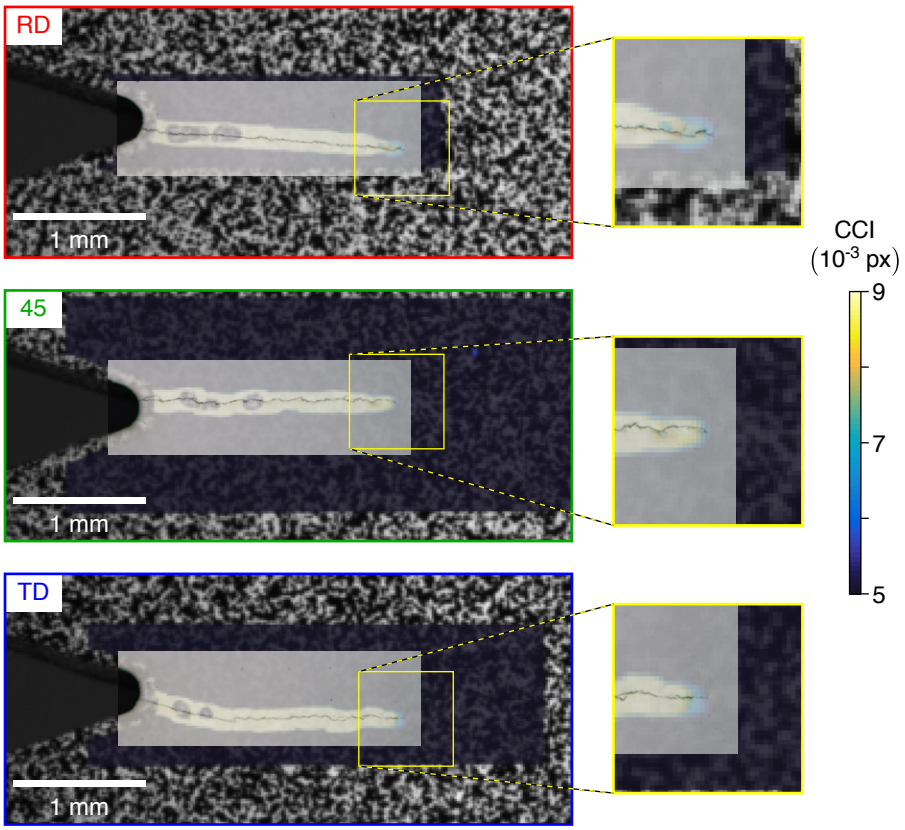


Figure 6.14: The crack path measured from the correlation confidence interval (CCI) of DIC, with a semi-transparent overlay of the crack path (to scale) imaged with optical microscopy after fatigue cracking (and after the speckle pattern was removed).

Using the DIC crack tip tracking method highlighted in Figure 6.14, the crack growth rates were measured for RD, 45, and TD samples in both low and high ΔK regimes (Figure 6.15). For the crack lengths with respect to number of cycles (Figure 6.15a), the starting rise in stress intensity was $\Delta K = 8.0$ MPa$\sqrt{\text{m}}$, and three samples were cracked for each condition.

At the high ΔK regime, the three RD samples had consistently faster crack growth rates than the three TD samples. The behavior of the 45 samples was more complicated, however. One of the 45 samples had a crack growth rate similar to the RD sample with the slowest crack growth.

In the low ΔK regime, the RD sample had faster crack growth rates, and the 45 and TD samples had about the same crack growth rate. For example, at $\Delta K = 3$ MPa$\sqrt{\text{m}}$, the crack growth rate of the RD sample was 2 nm/cycle, twice that of the 45 and TD samples.

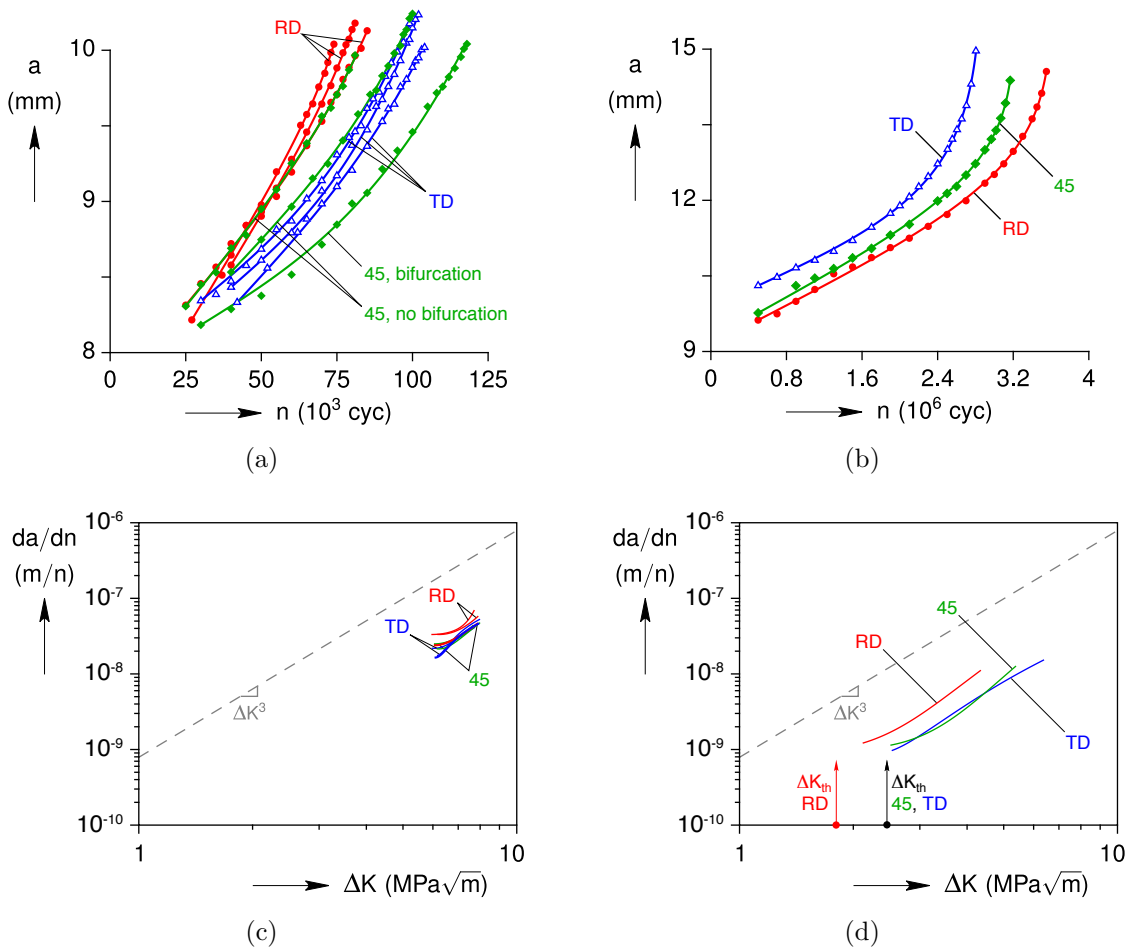


Figure 6.15: (a) The high ΔK regime of crack growth, all with $\Delta K_0 = 6.4$ MPa$\sqrt{\text{m}}$ for $a_0 = 8.0$ mm, including one 45 sample that exhibited a bifurcation (see Figure 6.16). (b) The low ΔK regime of crack growth, with different ΔK_0. (c) The crack growth rates (da/dn) with respect to ΔK for high ΔK. The Paris Law exponent for ΔK^3 is indicated by the gray line. (d) The crack growth rates (da/dn) with respect to ΔK for low ΔK. Again, the Paris Law exponent for ΔK^3 is indicated by the gray line.

Another 45 sample had a crack growth rate similar to the TD sample with the fastest crack growth. The third 45 sample had a bifurcation in its crack path (Figure 6.16), with an initial crack that arrested after about 250 µm of extension, and then a second crack that dominated the remainder of the FCG. Macroscopically-visible bifurcations such as this were not observed in any RD or TD samples.

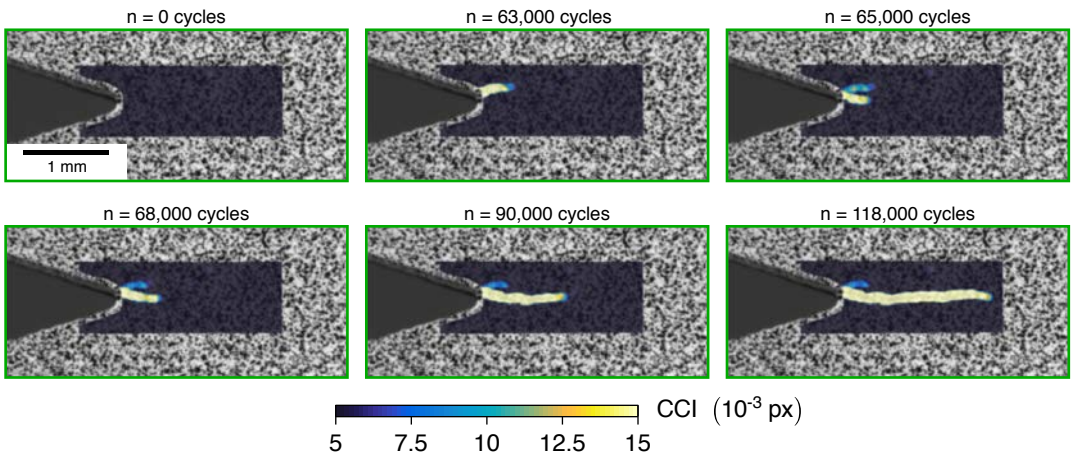


Figure 6.16: Bifurcation was observed in one of the 45 samples at high ΔK (but not in any of the other samples). This 45 sample with a bifurcation had the slower crack growth rate of any of the samples tested, due to the bifurcation, but the other two 45 samples had intermediate crack growth rates between those of the RD and TD samples.

The fatigue thresholds were measured for the three samples in the low stress intensity regime (Figures 6.15b and 6.15d). To find the fatigue threshold, the stress intensity amplitude was progressively decreased until the crack growth rate was just slower than one Angstrom per cycle (10^{-10} m/cycle), measured over at least 50,000 cycles. The fatigue thresholds for RD, 45, and TD samples were 1.8, 2.4, and 2.4 MPa$\sqrt{\text{m}}$. The fatigue thresholds and the crack growth rates at low ΔK are affected by crack closure (at least for fatigue cycling with low K_{min}, which is the case here, with $R = K_{min}/K_{max} = 0.1$). Crack closure was measured on the microscale in Section 6.5.3 and revealed more crack closure in the 45 and TD samples than the RD sample.

6.5.3 Microscopic crack measurements

For one sample of each condition, in-situ SEM-DIC measured the strain around the crack tips and the separation along the faces of the cracks. Figure 6.17 shows the normal strain perpendicular to the crack direction (ε_{yy}) within 300 µm fields of view around the crack tips at three stress intensities (K/K_{max} =0.3, 0.7, and 1.0). Consistently, the TD sample had lower strain at each stress intensity, both near the crack tip and away from the crack tip. For all three conditions, there were two distinct lobes of elevated strain, separated by a valley of low normal strain ($\varepsilon_{yy} \ll 1$ %) that began < 50 µm ahead of the crack tip.

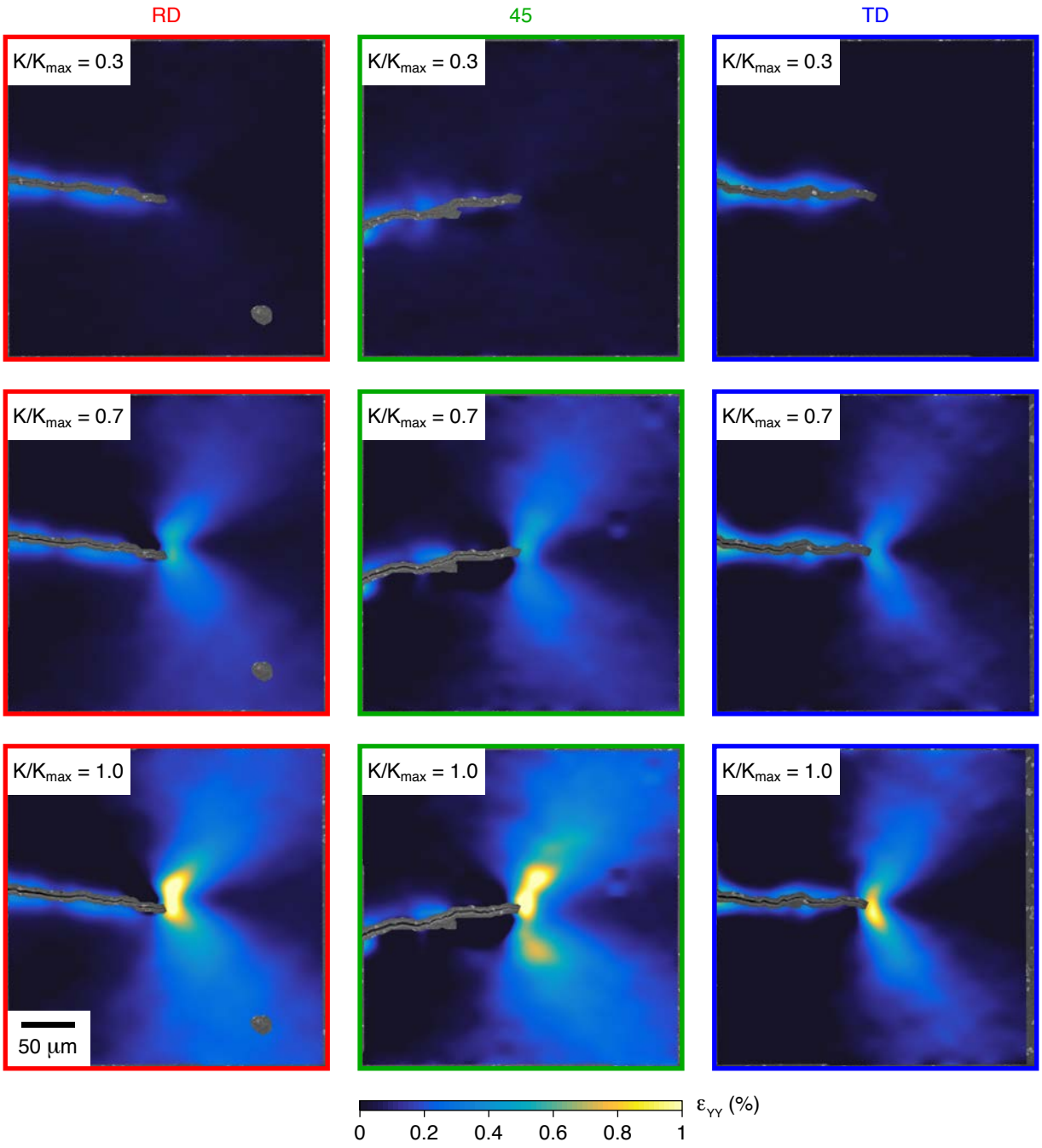
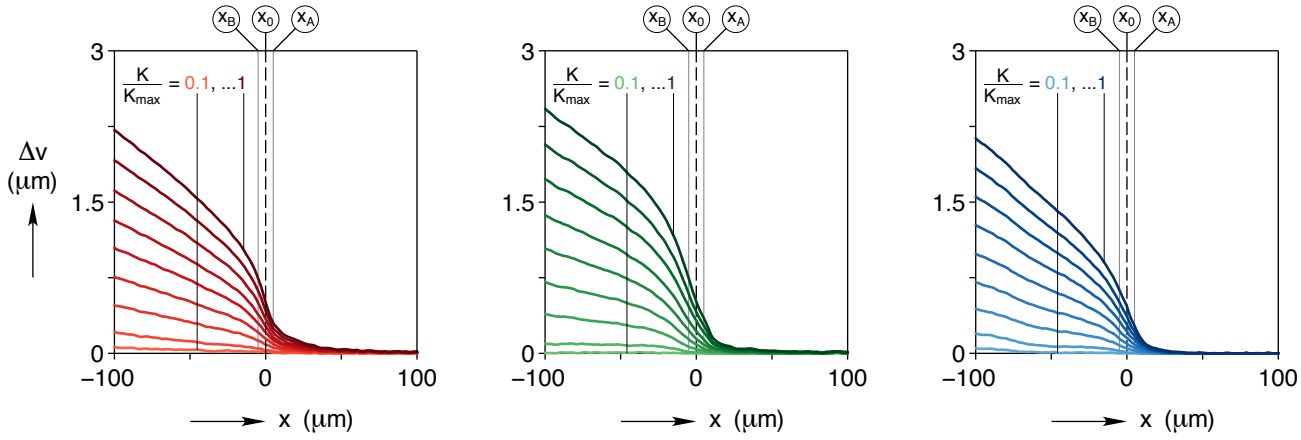


Figure 6.17: The normal strain perpendicular to the crack direction, ε_{yy}, was measured in a 300 µm SEM-DIC field of view for each of the three conditions (RD, 45, and TD). The shape of the high-strain lobes are similar for the three samples, but the strain magnitudes are higher for the RD and 45 samples than for the TD. Although crack appears to be closed in all three samples at the lowest stress intensity shown here, $K/K_{max} = 0.3$, there was slight crack opening (< 200 nm) for all three cases when the crack face separations were measured (Figure 6.18). Note that a small region of data away from the crack tip, coinciding with a poor region in the nanoparticle speckle pattern, was excluded for the RD sample's SEM-DIC calculation.

The crack separation profiles were measured following the method in Ref. [5] (Section 5.2.3). Overall, all three cracks opened in a similar manner with a region behind the crack tip with significant opening ($\Delta v > 1$ µm) and a short region in front of the crack with very small opening ($\Delta v \ll 1$ µm). Within 100 µm ahead of the crack tips, there was approximately no separation. Nonetheless, there were subtle differences in the crack separation behavior of the three samples. The 45 sample had the most crack opening far behind the crack tip. Interestingly, the 45 sample also had the least crack opening at the lowest stress intensity above zero ($K/K_{max} = 0.1$).

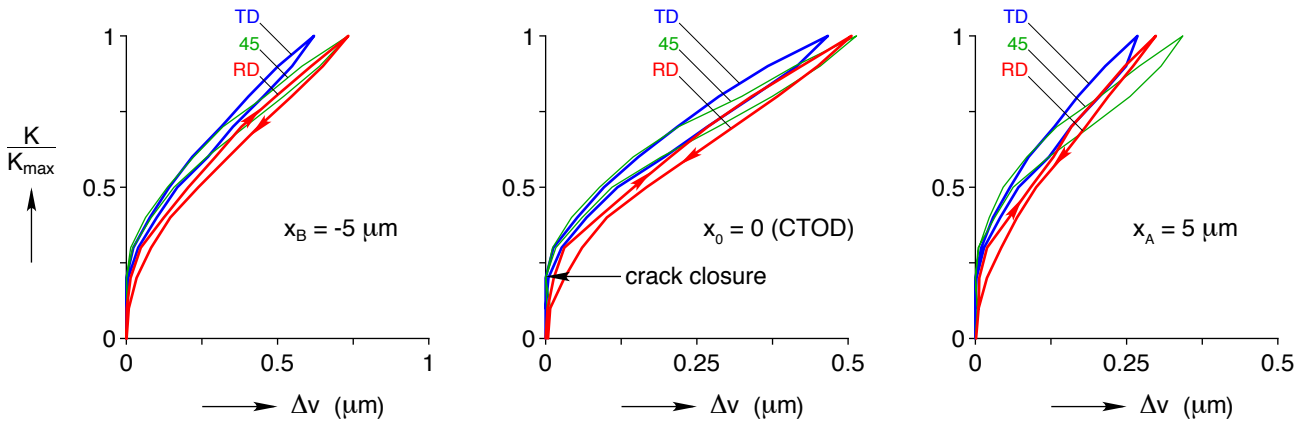


(a) RD sample (b) 45 sample (c) TD sample

Figure 6.18: The profiles of crack separation (Δv) with respect to distance from the crack tip (x) were similar for all three samples. Although the 45 sample had the least opening at the lowest non-zero stress intensity ($K/K_{max} = 0.1$), it had the most opening at the highest stress intensity ($K/K_{max} = 1.0$). For each profile, three positions were inspected: behind the crack tip (x_B =-0.5 µm), at the crack tip (x_0 =0 µm), and ahead of the crack tip (x_A =0.5 µm) for both loading and unloading in Figure 6.19. The crack separation profiles on unloading are not shown, but had similar shapes and only slightly lower magnitudes of separation.

Although the magnitude of crack opening was similar for RD and 45 samples behind the crack tip (x_B, Figure 6.19a) and at of the crack tip (x_0, Figure 6.19b), the 45 sample had more opening ahead of the crack tip (x_A, Figure 6.19c). Importantly, the RD sample exhibited more crack tip opening displacement (CTOD) and less crack closure than the 45 and TD samples. While the 45 and TD samples remained closed at $K/K_{max} = 0.2$, the RD sample exhibited slight opening ($\approx$ 20 nm). By inspecting the point at which the CTOD was non-zero (from Δv in Figure 6.19b), the RD crack tip opened at K/K_{max} between 0.1 and 0.2, the 45 and TD crack tips opened at K/K_{max} between 0.2 and 0.3. This result is consistent with the macroscopic measurements in Section 6.5.2 of the RD sample's lower

fatigue threshold (1.8 MPa$\sqrt{\text{m}}$) compared with that of the 45 and TD samples (2.4 MPa$\sqrt{\text{m}}$). Also, this result is consistent with the slower crack growth rate in the RD sample at low stress intensities (Figure 6.15d). Again, the sub-20 nm differences among the samples here underscored the importance of the high-precision SEM-DIC with custom scan controller, as developed in Ref. [179].



(a) x_B (b) x_0 (c) x_A

Figure 6.19: Three positions were inspected from Figure 6.18 and shown here for loading and unloading. In all cases, there was a slight hysteresis between loading and unloading, forming a clockwise loop (although arrows are only drawn on the RD cases).

Postmortem fracture surfaces were examined in SEM for one sample of each condition, and at two locations along the fracture surfaces: 1 mm from the notch root, $a = 9$ mm (Figure 6.20), and 4 mm from the notch root, $a = 12$ mm (Figure 6.21). The fractography used three magnifications for the three samples. No significant differences were observed among the fracture surfaces of the RD, 45, and TD samples. All three sets of fracture surfaces exhibited relatively high roughness with only small regions of ductile tearing visible at the highest magnification. There were large areas with striations angled ≈ 45 ° from the crack growth direction. Vertical striations (perpendicular to the crack growth direction) were found in only limited areas of all three samples. Overall, these fractography results indicate that there are the differences in macroscopic crack growth rates in Section 6.5.2 and microscopic crack tip measurements in Section 6.5.3 are *not* attributable to differences in crack roughness.

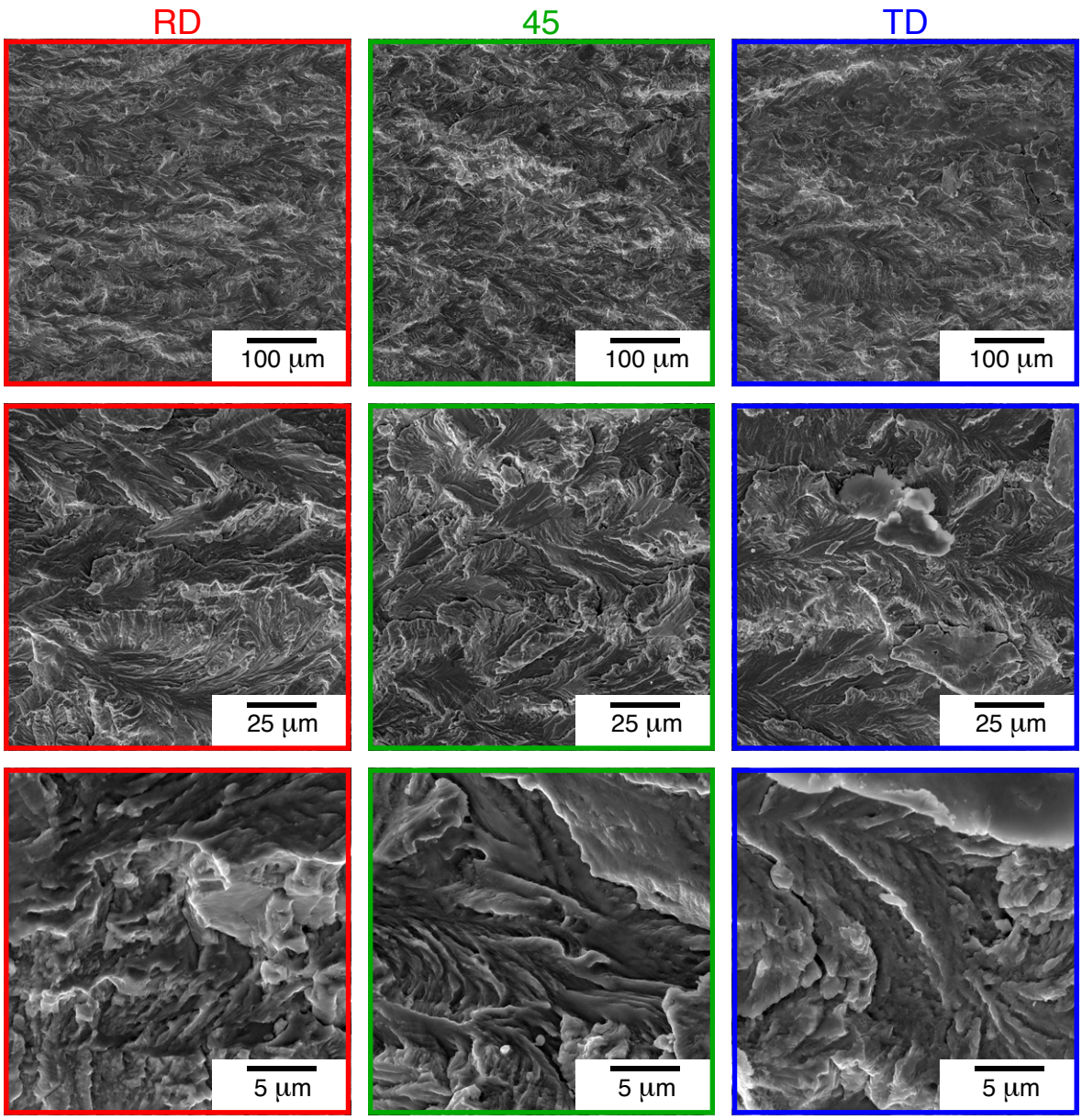


Figure 6.20: Fractography was performed in SEM at 1 mm from the notch root (a = 9 mm). The crack growth direction was from left to right, and the images were captured orthonormal to sample surface.

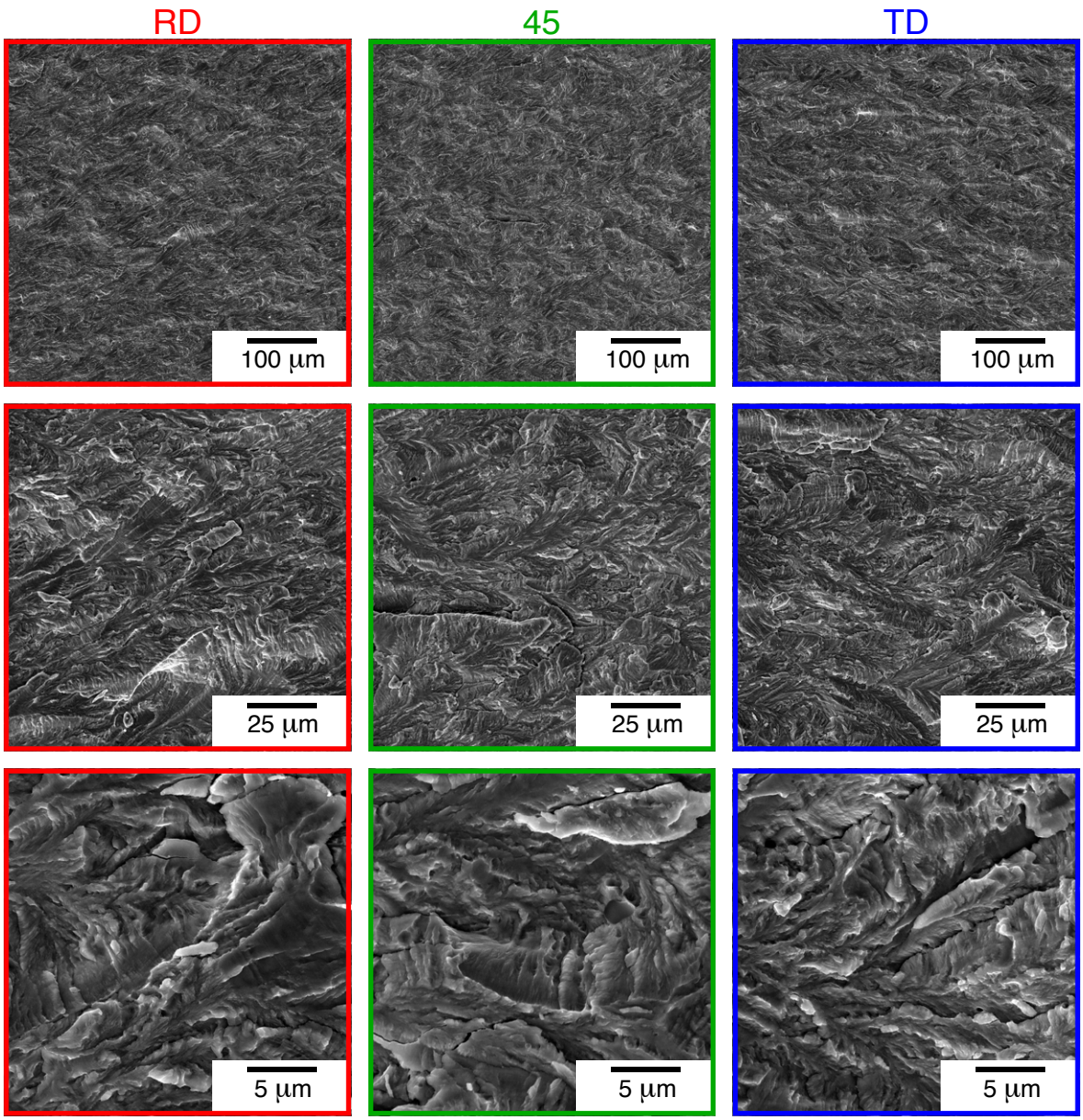


Figure 6.21: SEM fractography was also performed at 4 mm from the notch root ($a = 12$ mm). The crack growth direction was from left to right, and the images were captured orthonormal to sample surface.

6.6 Summary and conclusions

The functional and structural fatigue was characterized for superelastic NiTi with respect to crystallographic anisotropy. Differences between three rolled sheet directions (RD, 45, and TD) were measured in cyclic uniaxial tension, macroscopic crack growth rates and fatigue crack thresholds, and SEM-DIC crack tip displacement and strain fields.

- The NiTi material had an austenite finish temperature of $A_f = 3$ °C, a broad grain size distribution between tens of nanometers up to about 1 µm, and a strong γ-fiber texture, or (111) in the sheet normal direction. The most significant texture difference between the RD, 45, and TD sample directions was about twice as much (110) texture in the RD direction. The 45 and TD directions had relatively similar textures.

- During load-unload cycles in uniaxial tension, there was about three times the residual strain accumulation in the RD sample compared with the TD sample. Although the plateau stress of the 45 sample was about equally between the plateau stresses of the RD and TD samples, the residual strain accumulation of the 45 sample was much closer to that of the RD sample.

- Possibly connected with the TD sample's relatively low strain rate accumulation, there was an interesting observation in the strain maps during cyclic tension. While the phase transformation consistently localized in the same spots on the gauge sections of the RD and 45 samples, the localization in the TD sample did not occur in the same spot from cycle to cycle.

- The crack growth rates of the RD condition were consistently faster than the 45 and TD conditions. For fatigue cracking in a regime with relatively high ΔK, the crack growth rate of the RD samples was about 25 % faster than that of the TD samples, while the 45 sample were in between. Also, a crack bifurcation was observed in one 45 sample, but not in any other samples. At $\Delta K = 3$ MPa$\sqrt{\text{m}}$, the crack growth rate of the RD sample was 2 nm/cycle, twice that of the 45 and TD samples.

- The fatigue crack thresholds for the three sample conditions were 1.8 MPa$\sqrt{\text{m}}$ for RD, and 2.4 MPa$\sqrt{\text{m}}$ for 45 and TD (with stress intensity ratio $R = 0.1$). This macroscopic observation connected with the microscopic observations of crack closure: while the RD crack tip opened at K/K_{max} between 0.1 and 0.2 (between 0.6 and 1.2 MPa$\sqrt{\text{m}}$), the 45 and TD crack tips opened at K/K_{max} between 0.2 and 0.3 (between 1.2 and 1.8 MPa$\sqrt{\text{m}}$).

- No significant differences were observed in the fracture surfaces of the three sample conditions, so the differences in their macroscopic crack growth rates and microscopic crack tip behavior are not attributable to differences in roughness (e.g. roughness-induced crack closure). Rather, their differences in both cracking behavior (structural fatigue) – along with their tensile cycling behavior (functional fatigue) – must be from the inherent anisotropy of the material.

CHAPTER 7

The role of martensitic phase transformation on crack tip energy in shape memory alloys

7.1 Introduction

The complex mechanical responses of NiTi can vary significantly with relatively small changes in temperature. Despite this well-known fact, there has not been an experimental work that measures NiTi's fatigue and fracture response with respect to more than five temperatures for the same sample material and geometry.

Holtz, et al., investigated the fatigue threshold behaviors of NiTI at five temperatures and found two distinct regimes of crack growth with respect to temperature [182]. For thermal martensite (below A_s), there were slower crack growth rates and more evidence of crack closure than the above A_s. Crack closure was not measured directly, but was inferred from the global load-displacement responses. Also, the fracture surfaces of the samples from below A_s were rougher than those from above A_s [182].

Fatigue crack responses with respect to four temperatures were reported by McKelvey and Ritchie. Compared with Ref. [182], consistent observations included an overall enhancement in crack growth resistance and more crack closure with thermal martensite. Unlike the observation in Ref. [182] of similar fatigue cracking for both superelastic and stable austenite, Ref. [183] measured the worst fatigue performance (lowest fatigue threshold and fastest near-threshold crack growth rates) in superelastic austenite. Nonetheless, stable and superelastic austenite had similar crack growth rates at higher stress intensities, implying that superelasticity did not enhance crack growth resistance in NiTi, and two explanations for this were presented [183]. First, although superelastic NiTi can have ≈ 0 residual strain upon cycling to ≈ 5 to 8 % strain, it still has large residual strains upon cycling when the applied strains are greater than about 10 %, which is expected to be the case for material at and very near crack tips. Also, the large triaxiality of stress for the plane strain samples

would inhibit the $A \rightarrow M$ transformation since the transformation has a negative volume change [183].

Ref. [184] measured the fracture toughness of NiTi at effectively three temperatures (from two different NiTi alloys with different transformation temperatures for two experiment temperatures for one of the alloys). For the thermal martensite and superelastic austenite samples, cracks propagated a similar stress intensities (30 ± 5 MPa$\sqrt{\text{m}}$). The similarity was attributed to the cracks growing into detwinned martensite for both cases. Also, SEM observations of crack tips showed that crack blunting did not occur in thermal martensite or superelastic austenite. Although the thermal martensite and superelastic austenite samples showed elastoplastic behavior, the mechanism of their deviation from linearity was attributed to martensite detwinning (and not crack tip blunting). For stable austenite, where stress-induced martensite does not form, the fracture toughness was significantly larger (53 MPa$\sqrt{\text{m}}$).

Importantly, as pointed out in a recent review [185], Ref. [184] observed stress-induced martensite within plane strain samples of superelastic austenite, using in-situ synchrotron X-ray diffraction. This observation debunks the hypothesis from Ref. [183], that large triaxial stress in plane strain samples would inhibit $A \rightarrow M$ transformation at the crack tips of plane strain superelastic samples.

Recently, Ref. [186] measured the fracture toughness of NiTi at three temperatures and developed a model to capture the temperature dependence on NiTi fracture. Consistent with Ref. [184], superelasticity did not enhance the fracture toughness of NiTi.

7.2 On the transformation toughening of NiTi during crack growth

An important distinction about the nature of martensitic phase transformation for cracks was discussed in Refs. [187, 188] and also made clear in Ref. [186]: despite suggestions to the contrary, the $A \rightarrow M$ transformation in shape memory alloys cannot enhance NiTi's fracture toughness (critical stress intensity at the onset of crack extension). Rather, the transformation only acts as a toughening mechanism upon crack *growth*. This concept can be explained in the stress-strain response of superelastic austenite, and the fact that fracture toughness relates to the strain energy release rate *right up until* crack growth. Upon loading (and before unloading), the stress-strain response of superelastic NiTi has a non-linear elastoplastic response with similar strain energy as a non-phase-transforming elastoplastic material (such as NiTi just below the A_s temperature). Upon unloading, a typical elastoplastic material only recovers elastic strain energy. However, superelastic austenite recovers significant strain energy upon unloading (during the $M \rightarrow A$ transformation). In the context of cracks, this

strain energy recovery is manifested in the unloading of material in the wake of an advancing crack. The reverse transformed material in behind the advancing crack tip can place compressive stresses on the crack. Importantly, though, unloading is required (in the material behind an advancing crack) for this transformation toughening on the crack *growth* to occur. Thus, there is not a fracture toughness enhancement from NiTi's martensitic transformation, since fracture toughness corresponds to the maximum strain energy *before* crack extension.

In a broader scope of fracture mechanics, one possible source of confusion for the fracture toughness of NiTi is the phase transformation toughening observed in pure zirconia and yttria-stabilized zirconia. Before discussing zirconia and NiTi, it is worthwhile to examine other toughening mechanisms besides transformation. Fracture toughening mechanisms include those that act primarily ahead of the crack tip (intrinsic mechanisms), and those that act primarily behind the crack tip (extrinsic mechanisms). Common toughness-enhancing mechanisms are identified in Figure 7.1 (adapted from Ref. [189]), including: fiber bridging, grain bridging, oxide wedging, cleavage fracture, crack deflection/bifurcation, and microvoid coalescence (related to higher ductility). An important characteristic of all of these toughening mechanisms is additional energy dissipation *before* crack extension.

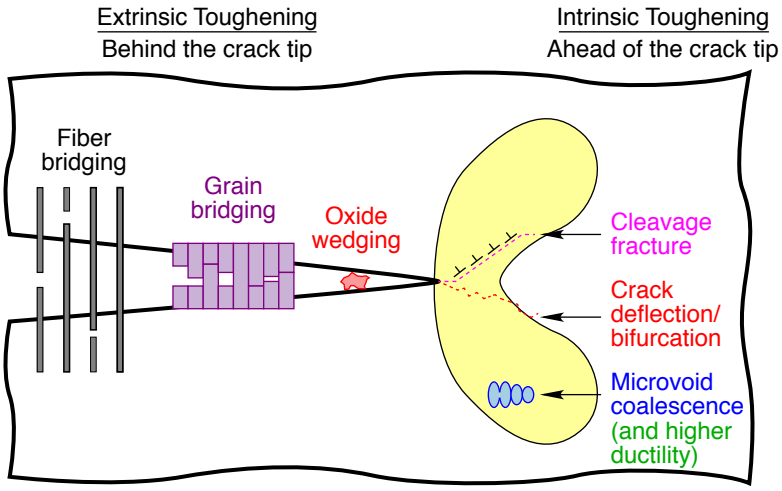


Figure 7.1: Along with certain cases of transformation with large volume expansion (such as the transformation in zirconia, but *not* the transformation in NiTi), other toughening mechanisms include: fiber bridging (composites), grain bridging (metals), oxide wedging, cleavage fracture, crack deflection/bifurcation, and microvoid coalescence (related to higher ductility). Adapted from Ref. [189].

Another toughness mechanism (not shown in Figure 7.1) is transformation toughening. Zirconia can exhibit a stress-induced transformation, similar to superelastic NiTi. However,

although zirconia's transformation imparts toughening, this does not imply that all martensitic transformations increase toughness. Importantly, zirconia's transformation involves a large volume expansion (about 4 %), while NiTi's transformation involves a very small volume contraction (about 0.023 %) [41]. While zirconia's transformation and concurrent volume expansion induce compressive stress at the crack tip *during forward transformation*, superelastic NiTi does not. Although NiTi's transformation enables extra energy dissipation upon unloading, this does not contribute to energy dissipation before crack extension (and before unloading).

7.3 Background on fracture mechanics and crack opening displacement methods

The theory of linear elastic fracture mechanics (LEFM) was presented by Griffith in 1921 [190], although it was not widely utilized until decades later. Griffith developed LEFM to understand why brittle materials failed with stresses on the order of 100 MPa, while the theoretical strength of atomic bonds in the materials were on the order of 10 GPa. Griffith's work showed that the critical stress to fracture (σ_F) in brittle materials, times the square root of flaw length (a), was approximately constant (C), following $\sigma_F\sqrt{a} \approx C$. In the 1940s and 1950s, Irwin extended Griffith's theory to capture the dissipation from crack tip plasticity [173]. While Griffith's theory described fracture behavior in brittle materials very well, it did not agree well with observations of structural materials such as steel with significant plasticity during fracture. Irwin included plastic dissipation into LEFM, and also developed the concept of the stress intensity (K) of a crack. For an applied stress in a structure (σ) with crack length a, the stress intensity is $K = f\sigma\sqrt{\pi a}$, with the term f depending on the geometry of the structure and crack.

In the decades following Irwin's extension of LEFM, continued work on structural materials, especially high-toughness steels, revealed further limitations to LEFM. Irwin's modification was limited to cases with "small-scale yielding," or relatively small regions of plasticity for relatively large cracks. However, Wells and others observed that steels with progressively higher toughnesses would have more rounded crack tips. Also, for the ones with the highest toughnesses, the radius of the plastic zone could be on the order of the length of the crack, and the modifications from Irwin for LEFM did not capture the fracture response well.

Importantly, with significant crack tip plasticity, the two faces of a crack can separate at the crack tip without the crack moving forward. This behavior can be described in terms of a positive crack opening displacement (COD) with respect to an applied stress intensity (K). Theories developed independently and simultaneously by Cottrell and Wells sought

a relationship between a material's fracture toughness and the COD of the material in a fracture test [191, 192, 193, 194]. Further work by in the 1970s confirmed reliable fracture toughness predictions via the crack tip opening displacement (CTOD), or equivalently the crack tip opening angle (CTOA) [195, 196, 197, 198].

Early measurements of COD used rotating paddle apparatuses, such as the "codmeter" in Figure 7.2. The rotation of the paddle corresponded to COD through a calculation of the crack face separations, with the assumption that the crack opens to the tip like a hinge.

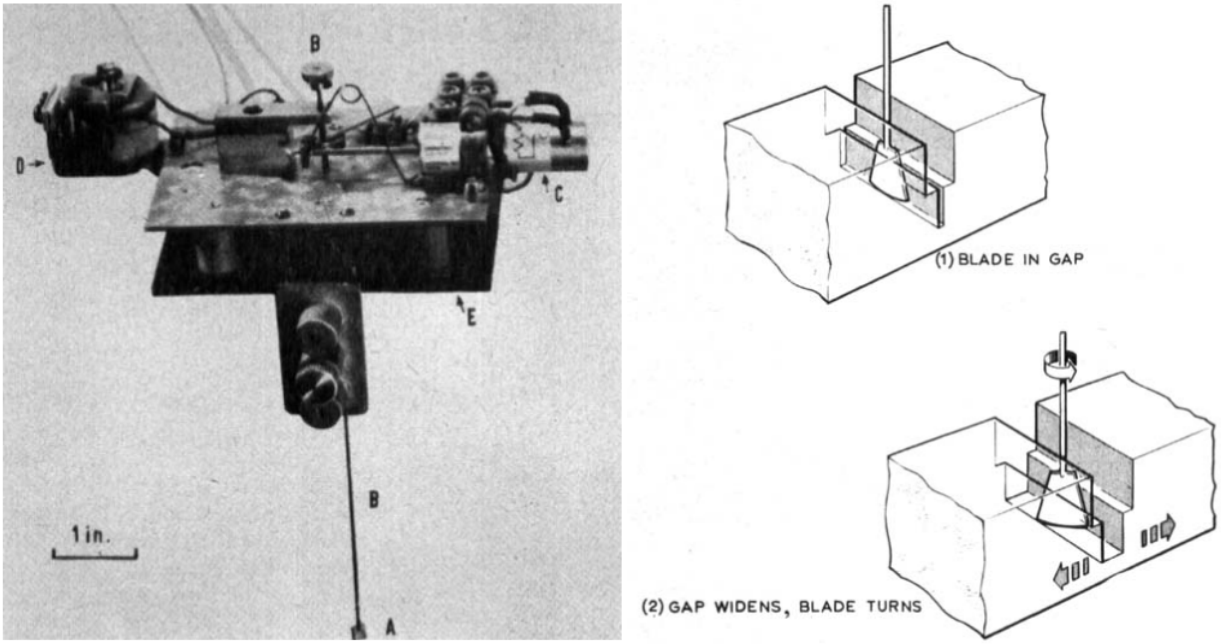


Figure 7.2: A "codmeter" rotating paddle apparatus, used to measure crack opening displacements in the 1960s during early development of elastoplastic fracture mechanics. Figure source: Ref. [191].

However, the accuracy of the paddle-based codmeter measurements were limited due to complications from the roughness of crack surfaces and the inability to plunge the paddle in narrow crack regions. A refined approach to the paddle method uses extensometers on the crack mouth, again with the assumption that the crack opens like a hinge down to the crack tip. Although this method of a crack mouth gauge is still popular today, there are limitations to its accuracy, especially from crack closure. Crack closure, illustrated in Figure 7.3, denotes that a crack remains partially closed with an opening force (for Mode I fracture). Crack closure, first described by Elber [126, 199], improves resistance to fatigue cracking.

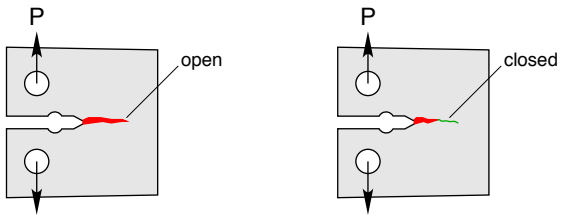


Figure 7.3: Crack closure denotes that a crack remains partially closed during an opening force P (for Mode I fracture). The schematic on the left shows a crack that is opened along its entire length (no crack closure), and the schematic on the right shows a crack that only partially opens (with crack closure).

To circumvent the inaccuracies of indirect COD measurements, optical techniques have been developed that measure crack opening with respect to distance from the crack tip. Notably, measuring COD with digital image correlation (DIC) is a powerful technique. DIC was first applied in the late 1980s for measuring stress intensity factors around cracks [142], and then in the 1990s for measuring crack opening and closing displacements [143]. The work was extended in 2000 to develop a mixed-mode fracture criterion from crack tip opening displacements (CTOD) to predict both the onset and direction of crack growth [200].

7.4 Material and methods

NiTi sheet (0.400 mm thickness, 50.2 at% Ni) was received flat annealed and etched from the manufacturer (Johnson Matthey, lot number 017641, San Jose, CA). The transformation temperatures were measured with differential scanning calorimetry (DSC; TA Instruments model Q200). The heat flow results, shown in Figure 7.4, indicated the following transformation temperatures: M_f of -51 °C, M_s of -11 °C, R_f of 40 °C, R_s of 51 °C, A_s of 35 °C, and A_f of 54 °C.

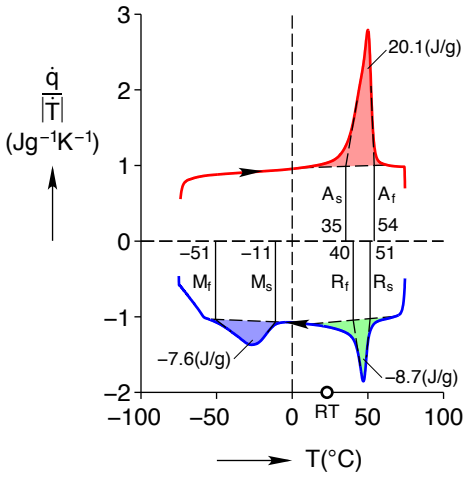


Figure 7.4: The phase transformation temperatures were measured with differential scanning calorimetry (DSC). At room temperature, the material is either thermal M (when heating from below the M_f of -51 °C), or thermal R (when cooling from above A_f of 54 °C).

7.4.1 Uniaxial tension

To measure the material's tensile response with respect to temperature, the samples were heated via conduction with resistive heaters (Minco HAP6944) in a tension instrument (Instron E1000 with 2 kN load cell), as shown in Figure 7.5. The sample temperature was controlled with a low-wattage temperature controller (Thorlabs TC200), with the control temperature signal from a thermistor (100 Ω, Thorlabs TH100PT). For the uniaxial tension experiments only, the gauge section was too small for affixing a thermistor, so the temperature of the gauge section was approximated by a correlation in one sample between the thermistor location outside of the gauge section (shown in Figure 7.5) to the gauge section's temperature measured by an infrared camera. Although other heating methods (especially a thermal chamber) would be more expedient that this custom conductive heating method, the conductive method was designed to be adaptable to vacuum environment (in the chamber of a scanning electron microscope) for future experiments.

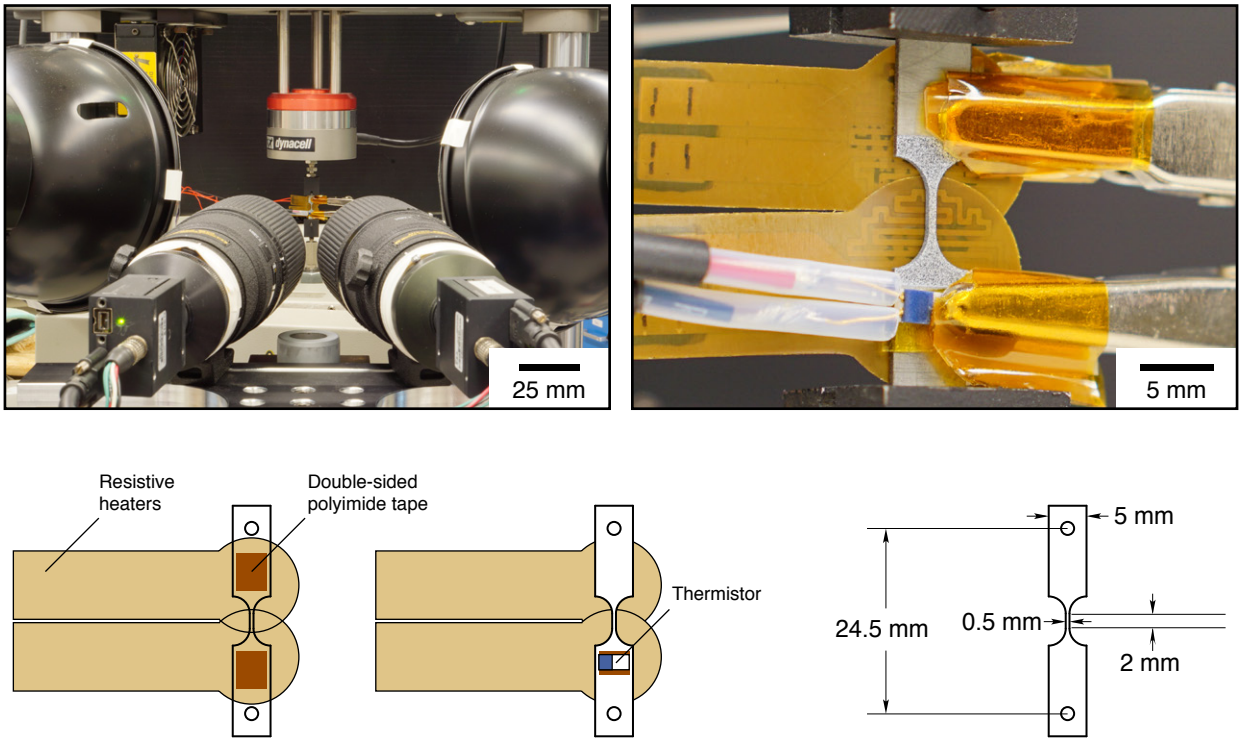


Figure 7.5: Experimental setup for uniaxial tension experiments.

7.4.2 Crack tip opening displacements with constant force and varied temperature

To measure the material's cracking response, compact tension samples were made with the geometry shown in Figure 7.6. A precrack of $a = 4.67$ mm was made by fatigue cycling a compact tension between $P_{min} = 10$ N and $P_{max} = 100$ N. During fatigue precracking, the sample was held at 136 °C (near stable austenite). The sample was used with the setup on the left in Figure 7.6 for the constant-force crack tip opening displacement (CTOD) measurements in Figure 7.8. The heating hardware matched the setup from Section 7.4.1, except with a miniature thermistor (10 kΩ, Epcos B57540G1103F000) taped directly behind the cracking region of the sample. The sample temperature was measured directly from this thermistor instead of the correlation to infrared camera measurements from Section 7.4.1.

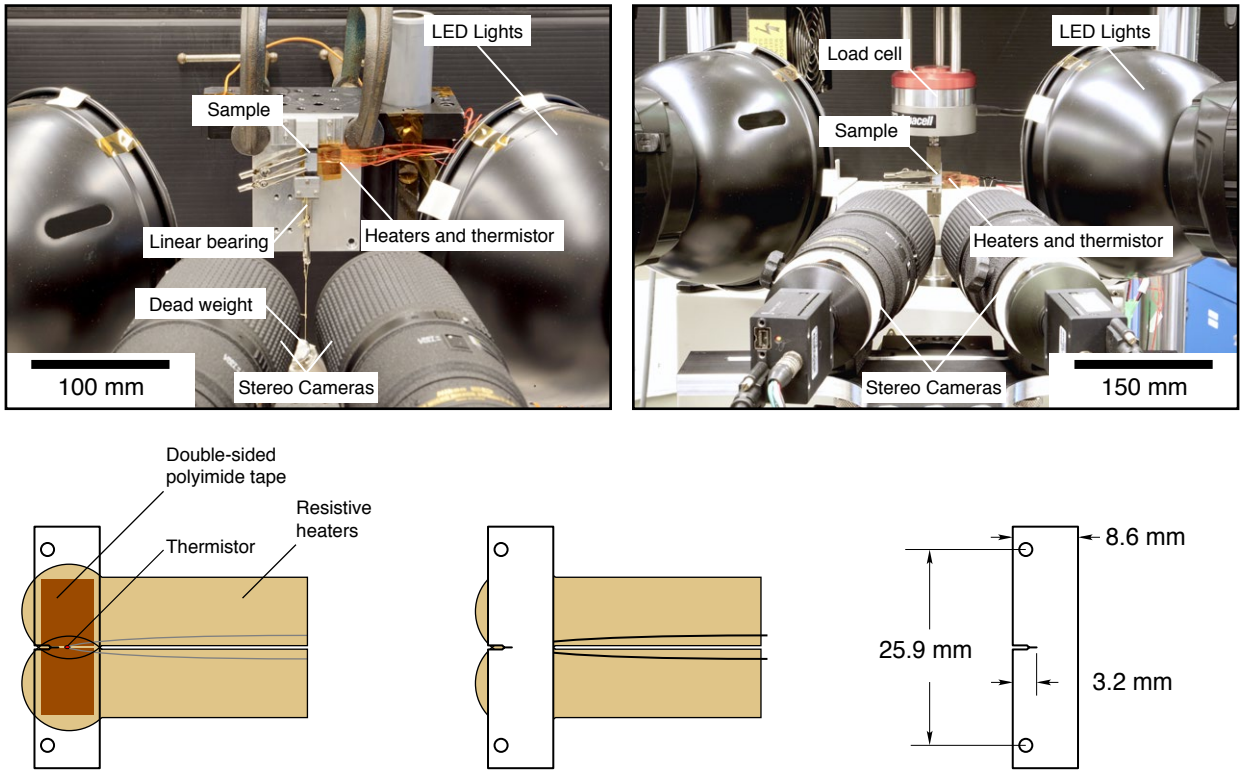


Figure 7.6: Experimental setup for compact tension experiments, with the constant-force setup on the left and the fatigue cracking setup on the right. Each sample was heated with two resistive heaters, and their temperature was measured with a thermistor taped behind the crack.

7.4.3 Fatigue crack growth measurements

Also, fatigue crack growth measurements were conducted on samples with the following temperature conditions (all in degrees Celsius): 194, $57^{(-)}$, $45^{(-)}$, $25^{(-)}$, $57^{(+)}$, $45^{(+)}$, and $25^{(+)}$. The three conditions denoted with "$^{(-)}$" indicate the temperature was reached after heating the sample to above 190 °C and then cooling down to the set temperature. The three conditions denoted with "$^{(+)}$" indicate the temperature was reached after cooling the sample to below −60 °C by placing the sample in contact with dry ice, and then heating the sample up to the set temperature with < 0.5 °C overshoot. The fatigue crack growth measurements followed the methods from Sections 5.2.2, 5.2.4 and 6.4.3, with the addition of the temperature controlling method from Section 7.4.2. The setup for these experiments is shown on the right side of Figure 7.6. The first set of fatigue crack growth measurements were conducted at 10 Hz within a relatively narrow range of moderate stress intensities to initiate a crack from the notch root. The second set of fatigue crack growth measurements were conducted at 50 Hz to examine crack growth rates in a broader range of stress intensities.

7.5 Results

The tensile response was measured for a range of temperatures between 22 and 194 °C, with two load-unload cycles for each temperature (Figure 7.7). The loading increments of each cycle were arbitrarily selected, and the unload points were to < 1 N force. At 22 °C, the sample was thermal martensite. During tension, the material detwins, and there was significant plasticity from each of the two cycles. At 36 °C, the tensile response included a "knee" from the R-phase, consistent with other observations of thermal martensite without cryogenic quenching [37]. At 50 °C, just below A_f (54 °C), there was localization of $A \rightarrow M$ transformation on unloading and $M \rightarrow A$ on unloading, but the reverse transformation did not complete before the stress was removed. At 57 °C, just above A_f (54 °C), there was full superelastic loading and unloading, with $\ll 1$ % residual strain after the two cycles. With progressively increased temperatures, the tensile responses followed the expected behavior of NiTi, including increased plateau stresses (both on loading and unloading) and increased residual stress upon cycling. At 93 °C, there was localization during forward transformation, but not on the reverse. The martensite desist (M_D) temperature was between 150 and 194 °C.

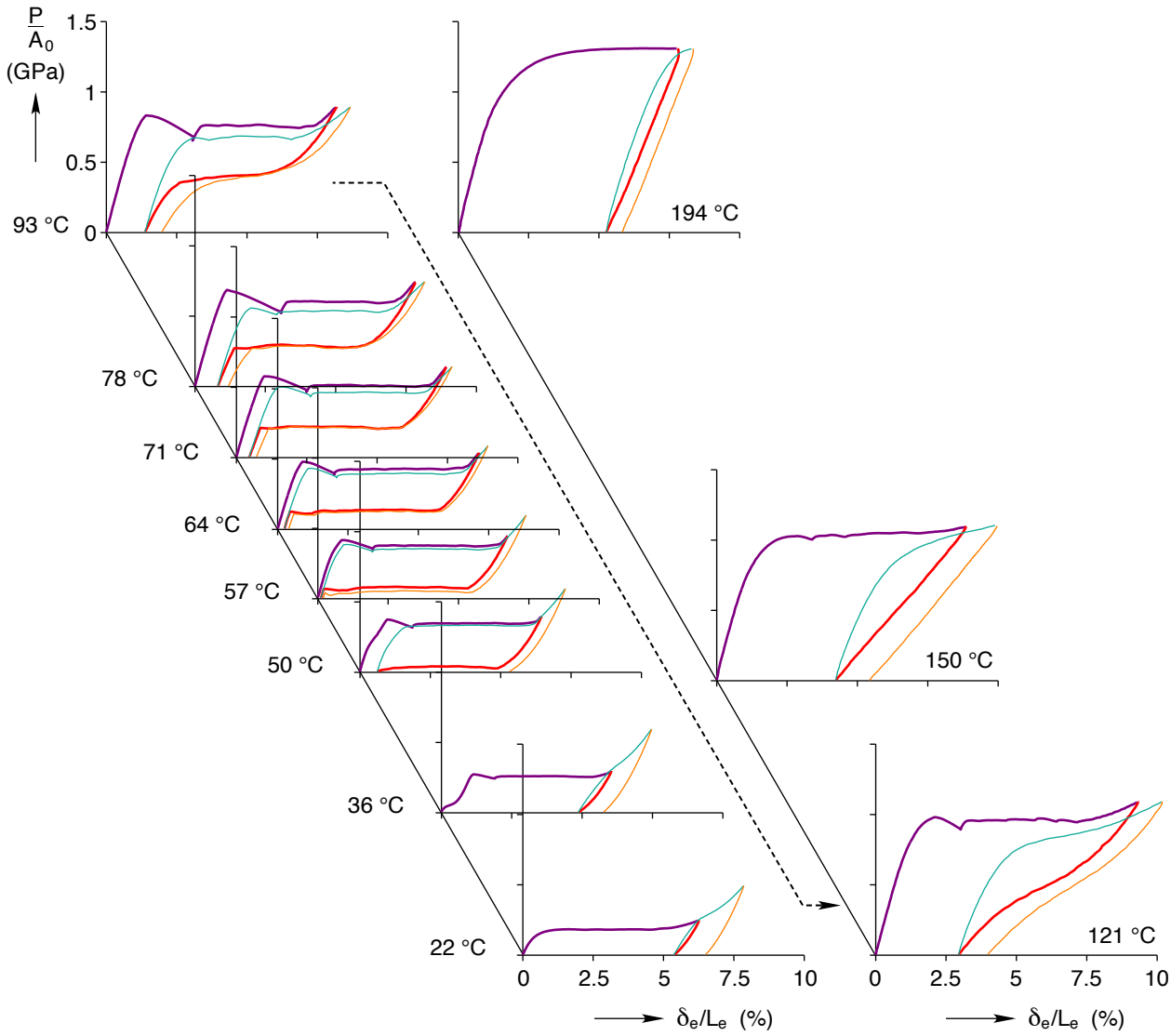


Figure 7.7: The stress-strain responses in tension for temperatures ranging between 22 and 194 °C, with two load-unload cycles at each temperature. The first load-unload cycles have thicker lines (purple and then red); the second load-unload cycles have thinner lines (teal and then orange).

Next, the crack tip opening displacement (CTOD) was measured with respect to temperature. For progressively increasing constant forces, up to a maximum corresponding to $K_{max} = 6.4$ MPa$\sqrt{m}$, there was a consistent increase in CTOD. With respect to temperature, both hysteresis and non-monotonic trends in CTOD were observed. During heating from stable martensite to stable austenite, CTOD was consistently higher than upon cooling from stable austenite. Also, within the $A \rightarrow M$ transformation temperature range, CTOD increased from the value of stable martensite.

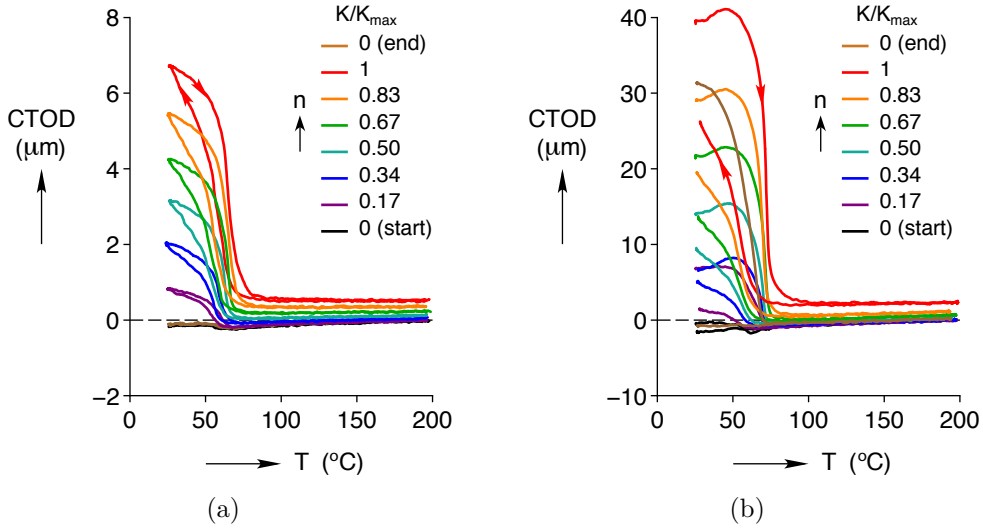


Figure 7.8: The crack tip opening displacement (CTOD) with respect to temperature. (a) The temperature beginning hot (starting in A and cooling to R). (b) The temperature beginning cold (starting in M, heating to A, and cooling to R).

Fatigue crack growth rates also showed two trends with respect to temperature. First, the fatigue crack growth rates were slower for decreasing temperatures. Second, the fatigue crack growth rates for the cooling direction (indicted by "$^{(-)}$") were generally slower than those in the heating direction ("$^{(+)}$").

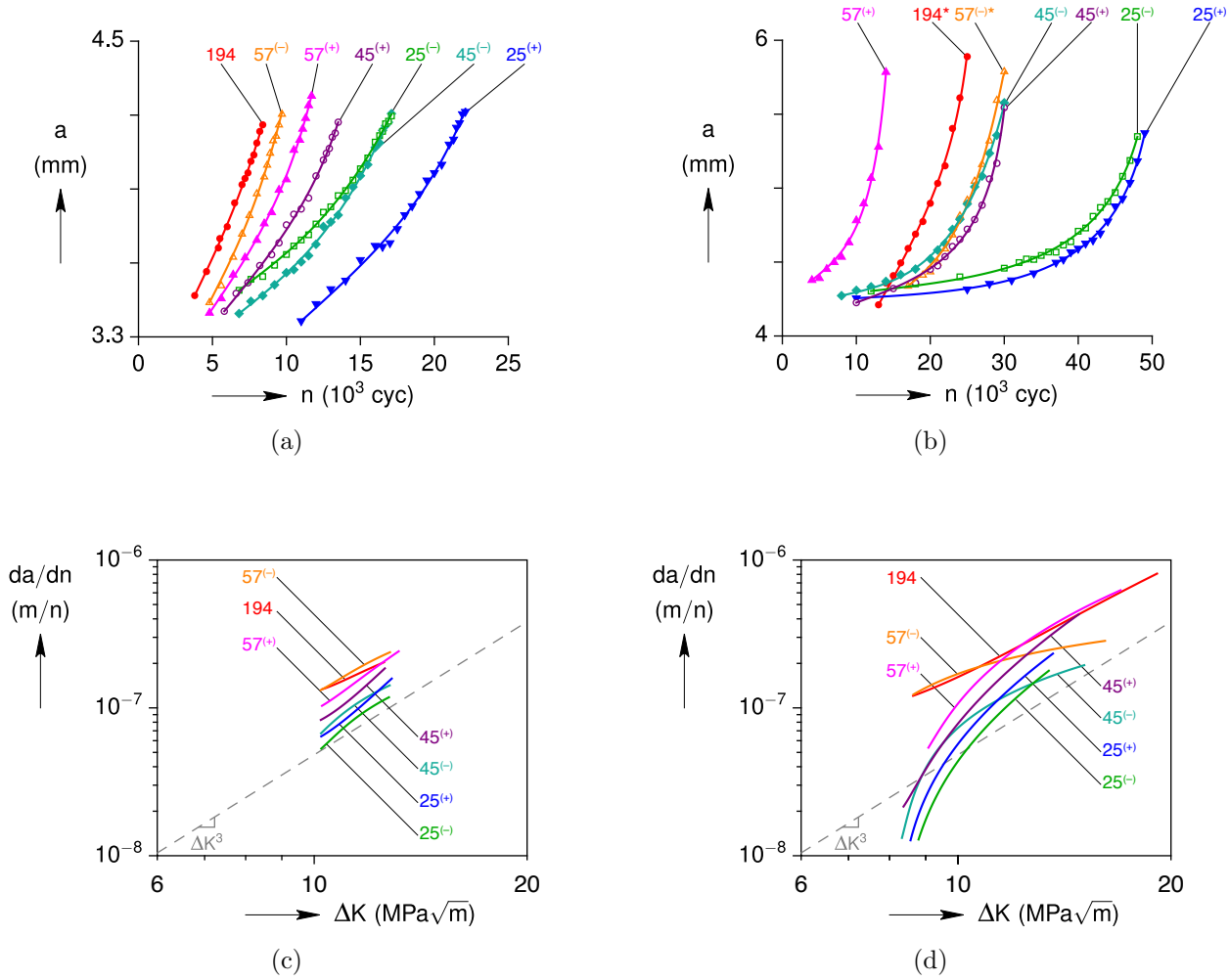


Figure 7.9: (a) The high ΔK regime of crack growth, all with $\Delta K_0 = 8.6$ MPa$\sqrt{\text{m}}$ for $a_0 = 3.2$ mm. Note that the measured ΔK begin past ΔK_0 because of the minimum crack length measurable from the notch root with DIC, which is one-half of a DIC subset width from the notch root. (b) The low ΔK regime of crack growth, with different ΔK_0 (from different crack lengths). Also, the ΔP for all conditions was 45 N, except with $\Delta P = 36$ N for the 194 and $57^{(-)}$ conditions (indicated by an asterisk). (c) The crack growth rates (da/dn) with respect to ΔK for high ΔK. The Paris Law exponent for ΔK^3 is indicated by the gray line. (d) The crack growth rates (da/dn) with respect to ΔK for low ΔK. Again, the Paris Law exponent for ΔK^3 is indicated by the gray line.

Finally, the fracture surfaces were observed with scanning electron microscopy (Figure 7.10). The fracture surface of the stable austenite condition (194 °C) was distinctly less rough than the others, with more cleavage zones and smaller and fewer ductile dimples. Two other trends were observed in the fracture surfaces. First, there was a slight increase in ductile dimpling with respect to decreasing temperatures from 57, 45, and 25 °C(for both

heating and cooling directions). Second, the samples in the heating direction (from stable martensite, indicated by "$^{(+)}$") had more ductile dimpling than those in the cooling direction ("$^{(-)}$").

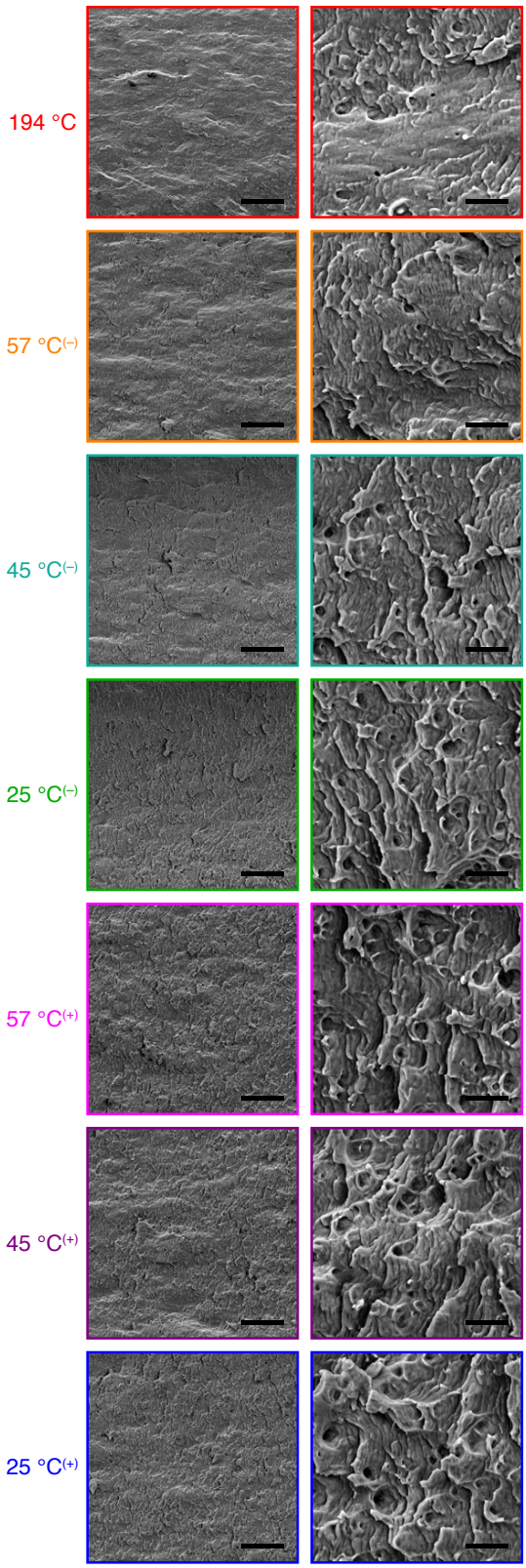


Figure 7.10: Fracture surfaces at $a = 5.7$ mm, with 40 µm scale bars for the left column of images and 4 µm scale bars for the right column of images.

7.6 Conclusions

The crack tip opening displacement (CTOD) and fatigue cracking response of NiTi was characterized with respect to temperature. The smallest CTOD and fastest crack growth rates were measured in the stable austenite phase. With decreasing temperature, crack growth rates decreased. Generally, the cooling-direction samples exhibited slower crack growth rates than the heating-direction samples. The slowest crack growth rates were observed in the $25^{(-)}$ sample with a mixture of R and stress-induced M. These crack growth rate observations were consistent with the fracture surfaces with respect to temperature.

Future work on this topic will address the rate sensitivity of fatigue cracking, which could be accessed with varied cycling rates and heat transfer environments. Also, the constant force setup described in Section 7.4.2 will be adapted for scanning electron microscope digital image correlation (SEM-DIC).

CHAPTER 8

Conclusions and future work

8.1 Conclusions

This work has contributed the following advancements to the fields of shape memory alloys and experimental mechanics.

1. An expedient method was developed for improving optical DIC, a popular and powerful experimental technique. The method, cross polarization, eliminated saturated pixels to preserve sub-pixel displacement resolution. Also, cross polarization improved the correlation confidence interval by about 10 %, and reduced 2-D DIC error by up to 60 % (Chapter 2).

2. Clear guidelines were presented for optimizing painted speckle patterns for optical DIC. Although paints are the most popular speckle pattern technique for optical DIC, there has been limited scientific investigations to inform best practices of painted patterns. This work showed that patterns with black speckles had 24 % lower median normalized false displacements than the patterns with white speckles. Also, the following guideline was offered: the optimum painted speckle pattern fro DIC is a random pattern of black speckles applied with about 50 % density over a basecoat of white paint (Chapter 3).

3. The crystallographic texture of NiTi tube (the most utilized shape memory alloy in its most important product form) was measured (111) in the drawing direction. Also, the axial residual stress in the tube was -29 MPa on the outer radius and 30 MPa on the inner radius (Chapter 4).

4. The grain size effect on NiTi fatigue crack growth was measured with a new multiscale experimental technique (Chapter 5).

   (a) Despite the superior functional fatigue of nanocrystalline NiTi, the largest grain size in the study (1500 nm average) had the best structural fatigue performance.

(b) For all grain sizes, the macroscopic fatigue crack growth correlated well with microscopic crack tip observations: the grain sizes with relatively fast macroscopic crack growth rates exhibited large crack displacements at the microscale, and vice versa.

(c) Crack closure was observed with SEM-DIC on the microscale (not with an indirect measurement like crack mouth gages), and crack closure in the 1500 nm grain size was suspected to slow its crack growth rate. Crack closure was not observed in the other grain sizes.

(d) On the macroscale, optical DIC enabled interesting observations of crack growth, especially a crack bifurcation and an oblique crack growth direction. These observations would have been missed with typical crack growth measurements, such as crack mouth gauges, back-face strain gauges, or electrical resistance drop measurements.

5. The effects of crystallographic texture on NiTi's functional and structural fatigue were characterized (Chapter 6).

(a) NiTi sheet material was utilized. The sheet had a grain size distribution between tens of nanometers up to about 1 µm. The crystallographic texture was a strong γ-fiber (or a (111) texture in the sheet normal). The rolling direction (RD) had about twice as much (110) texture as the directions transverse (TD) and 45 ° (45) from the sheet rolling direction.

(b) During cycles in tension, there was about three times as much residual strain accumulation in the sheet's rolling direction compared to the transverse direction.

(c) Possibly connected with the TD sample's relatively low strain rate accumulation, there was an interesting observation in the strain maps during cyclic tension. While the phase transformation consistently localized in the same spots on the gauge sections of the RD and 45 samples, the localization in the TD sample did not occur in the same spot from cycle to cycle.

(d) The crack growth rates of the RD condition were consistently faster than the 45 and TD conditions. For fatigue cracking in a regime with relatively high ΔK, the crack growth rate of the RD samples was about 25 % faster than that of the TD samples, while those the 45 samples were in between. Also, a crack bifurcation was observed in one 45 sample, but not in any other samples. At $\Delta K = 3$ MPa$\sqrt{\text{m}}$, the crack growth rate of the RD sample was 2 nm/cycle, twice that of the 45 and TD samples.

(e) The fatigue crack thresholds for the three sample conditions were 1.8 MPa$\sqrt{\text{m}}$ for RD, and 2.4 MPa$\sqrt{\text{m}}$ for 45 and TD (with stress intensity ratio $R = 0.1$). This macroscopic observation connected with the microscopic observations of crack closure: while the RD crack tip opened at $\text{K}/\text{K}_{\text{max}}$ between 0.1 and 0.2 (between 0.6 and 1.2 MPa$\sqrt{\text{m}}$), the 45 and TD crack tips opened at $\text{K}/\text{K}_{\text{max}}$ between 0.2 and 0.3 (between 1.2 and 1.8 MPa$\sqrt{\text{m}}$).

(f) Unlike the grain size study that was confounded by roughness-induced crack closure differences, the three texture conditions had similar fracture surface roughnesses. Thus, the differences in their macroscopic crack growth rates and microscopic crack tip behavior are not attributable to differences in roughness (e.g. roughness-induced crack closure). Rather, their differences in both cracking behavior (structural fatigue) – along with their tensile cycling behavior (functional fatigue) – must be from the inherent anisotropy of the material.

6. The role of martensitic phase transformation on cracking was characterized by varying temperature (Chapter 7). A new constant system with temperature control was developed and applied with optical DIC, and the system can also be adopted for SEM-DIC. The crack tip opening displacement (CTOD) and fatigue cracking response of NiTi was characterized with respect to temperature, and the following trends were observed.

   (a) With respect to temperature, both hysteresis and non-monotonic trends in CTOD were observed. During heating from stable martensite to stable austenite, CTOD was consistently higher than upon cooling from stable austenite. Also, within the $A \rightarrow M$ transformation temperature range, CTOD increased from the value of stable martensite.

   (b) Generally, the cooling-direction samples exhibited slower crack growth rates than the heating-direction samples. The slowest crack growth rates were observed in the $25^{(-)}$ sample with a mixture of R and stress-induced M.

   (c) These crack growth rate measurements were consistent with fractography observations: slower crack growth rates for the lower temperature samples corresponded to more ductile dimples and less cleavage on their fracture surfaces.

8.2 Future work

To crack the code on NiTi fatigue and fracture, there are many more scientifically interesting and technologically important questions to answer. The short list of these questions is provided below.

1. What is the structural fatigue and fracture of other SMAs (besides NiTi, or including ternary and quaternary alloys of NiTi, such as NiTiHf)? More SMAs in addition to NiTi should be explored and connected with the existing body of work on NiTi to understand their similarities and differences.

2. What is the crack response of SMAs utilizing the shape memory effect (with actuation driven fatigue that is thermally cycled with bias stress, instead of the mechanical cycling in the present work)?

3. In the present work, the experiments were all isothermal for any given cracking condition. How does varying temperatures change the history-dependent cracking response of NiTi? For example, is there an optimum temperature-stress pathway for an SMA actuator to optimize work output and limit the risk of catastrophic failure?

4. How do complex stress states, along with NiTi's anisotropic properties and tension/compression asymmetry, change the fatigue and fracture response? This will also be related to the other modes of fracture beside the Mode I opening that was explored in the present work. Can classic theories for cumulative damage, such as Miner's Rule, accurately describe NiTi's fatigue in complex cycling?

5. How do complex time intervals on fatigue cycling, especially with NiTi's latent heating and rate sensitivity, affect NiTi's cracking response?

6. environmentally-assisted crack growth, e.g. NiTi is known to exhibit hydrogen embrittlement [201], for harsh fluids with NiTi valves or other challenging environments

7. How can the competing mechanisms of cold working and precipitates be tailored for the optimum combination of strength and functional fatigue resistance? Cold working increases strength (with associated dislocation density and residual martensite), while precipitates of the correct size and density increase functional fatigue resistance (but do not survive severe cold working processes). These processing routes should be understood enough to design NiTi structures for specific applications. For example, if a device needs to prioritize strength over fatigue life (which may be the case for some emerging NiTi applications, such as earthquake-damping structures), then how can such an alloy's processing be rationally designed, modeled, and validate?

Furthermore, there is an immediate need to understand the role of inclusions for the fatigue life of NiTi devices (especially biomedical devices), and Section 8.2.1 addresses this topic.

8.2.1 The role of inclusions on NiTi device life

Inclusions are known to play a major role in the fatigue life of NiTi, and the "micro-cleanliness" that minimizes inclusions is considered the main factor that limits NiTi device life [202, 203]. Minimizing inclusions will be increasingly important for emerging NiTi applications that favor thin features and small structures, such as thin actuator wires that heat quickly, or minimally-invasive stents that deploy steadily small diameters. However, thinner NiTi structures risk a larger ratio between inclusion size and NiTi diameter/thickness.

Broadly, the characteristics of inclusions that affect fatigue life include: inclusion size and shape; inclusion position in material volume (within interior, or at free surface); inclusion mechanical adhesion and atomic coherency (or lack of) with matrix; elastic constants of inclusions and matrix; inclusion chemical composition; size, shape, and frequency of voids around inclusions (called particle/void assemblies, or PVAs [204]), and chemical composition of other species around inclusion (e.g. trapped H_2 gas leads to hydrogen embrittlement in steel, especially at very high cycle fatigue with N$> 10^7$ cycles). Other factors influencing the role of inclusions on fatigue life include surface roughness, residual stresses, and work hardening or other changes in the surface layer's microstructure.

The majority of NiTi endovascular devices are produced from tubes. The axisymmetric processes for making NiTi tubes generate longitudinal inclusions called "stringers" along the tube axis. Such a stringer inclusion is shown in Figure 8.1. If a NiTi device is produced with a stringer across the bend with stresses along the hoop direction (transverse), then the devices fatigue life will certainly be reduced.

An early work reported that carbon levels, inclusion content, unloading plateau strength, and active austenite finish temperature (A_f) did not have significant effects on the fatigue-rupture behavior of NiTi in bending rotary fatigue experiments due to [205]. However, follow-up works have observed fatigue failures that initiate at inclusions.

During bending rotary fatigue, failure initiated at surface flaws (scratches and extrusion defects) as well as near TiC inclusions. Importantly, smaller wire diameters systematically show larger rupture lives, indicating a significant role of volume defects on failure [206]. Similar to other Ti-based alloys, NiTi is sensitive to hydrogen embrittlement. A reduction in ductility was observed for NiTi wires with greater than 400 ppm hydrogen [201]. Nonetheless, the hydrogen was not introduced during normal NiTi processing, but rather was infiltrated with a heated solution of H_3PO_4. In service, NiTi products are known to fail from surface flaws and inclusions, but not hydrogen trapped around interior inclusions.

Although small TiC inclusions can be failure initiation sites [206], TiC grains in a metal matrix of NiTi had similar fatigue crack growth characteristics as unreinforced NiTi, despite markedly different responses in uniaxial tension [207]. The metal matrix composite had

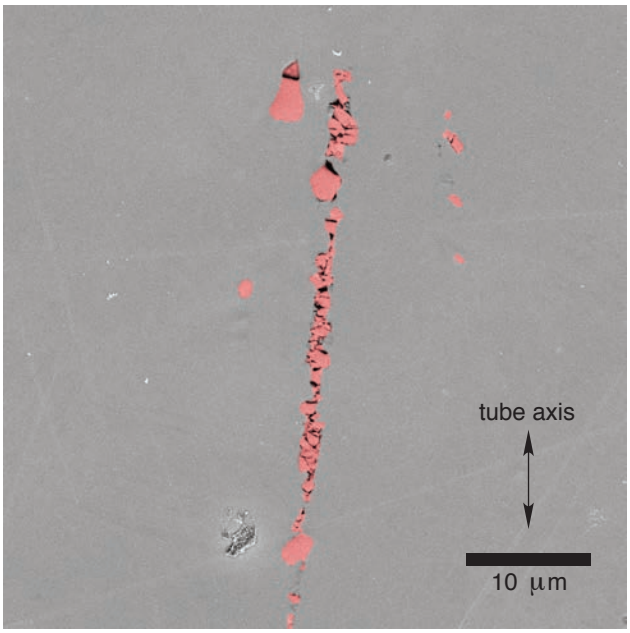


Figure 8.1: Backscattered electron (BSE) image of a stringer along the axis of a NiTi tube. The stringer (forming a particle/void assembly, or PVA) is falsely colored.

a matrix of ≈ 11 µm NiTi grains and 10 and 20 $\%_{\text{volume}}$ TiC with a similar grain size (≈ 13 µm) [207].

Aging of Ni-rich NiTi can nucleate lenticular, coherent Ni_3Ti_4 precipitates [208]. During aging without stress, the precipitates form at grain boundaries. In the presence of relatively small stress (2 MPa), the precipitates form within grains and at grain boundaries [208]. Compressive misfit strain forms at the ends of the precipitates, while tensile strain forms along the long sides of the precipitates [209]. This local stress field around the precipitates enhances the nucleation of martensite and lowers the transformation stress [210]. These precipitates play an integral role in the three-step transformation of aged, Ni-rich NiTi. First, R-phase nucleates at precipitate/matrix interfaces, then B19′ nucleates and grows along precipitate/matrix interfaces, then the matrix between the precipitates transforms to B19′ [211]. With increased aging temperature, the precipitates are larger and eventually switch from being coherent to incoherent with the matrix [212]. The size of Ni_3Ti_4 precipitates can be optimized to balance the competing factors of preventing dislocation motion (with smaller Ni_3Ti_4 precipitates that cannot be sheared/cut by dislocations) and lowering martensitic stress transformation (with larger Ni_3Ti_4 precipitates) [212].

Two common precipitates in NiTi are TiC and $Ti_4Ni_2O_x$. TiC inclusions in NiTi are generally small, uniformly sized, and evenly distributed, while $Ti_4Ni_2O_x$ are generally "blocky," unevenly distributed, and thus present a more significant fatigue initiation site [213]. Smaller diameter wires tend to have smaller inclusions, since more drawing steps break up inclusions.

However, the largest inclusion size was not significantly different between 230 and 610 µm diameter wires, and correspondingly, their fatigue lives were similar [213].

TiC inclusions at the surface were the primary fatigue crack initiation sites for as-drawn NiTi wires, while both surface scratches and inclusions were fatigue crack initiation sites for electropolished NiTi [214]. For both the as-drawn and electropolished wires, the inclusions were 4-6 µm in diameter.

The melting method of NiTi alloys determines the size, composition, and frequency of inclusions. One study [215] compared the inclusion content of two melting methods: vacuum induction melting followed by vacuum arc remelting (VIM/VAR), and induction skull melting followed by vacuum arc remelting (ISM/VAR). Multiple passes of VAR were also attempted, but resulted in unusable material. Although ISM/VAR had lower carbon content and 2.6 times fewer total inclusions than VIM/VAR, the inclusions in ISM/VAR were larger NiTi intermetallics that introduced 2.88 % void fraction (compared with 0.44 % in the VIM/VAR ingot). The void fraction was proportional to the number and distribution of NiTi intermetallic inclusions [215]. Collectively, these voids around particles are called particle/void assemblies (PVAs). Later, a scanning electron microscopy investigation [216] of VIM/VAR inclusions revealed four inclusion types: titanium carbide with a significant amount of substitutional oxygen, TiC(O); intermetallic oxide Ti_4Ni_2O; Ti_2Ni with or without interstitial oxygen; and core-shell inclusions comprised of a TiC(O) core and Ti_2Ni(C,O) shell [216].

The effects of impurity levels on fatigue life were investigated for a high-purity alloy and alloys with the maximum levels of C and O [204]. In low-cycle fatigue, the high-purity alloy had significantly higher fatigue life. In high-cycle fatigue (HCF), the high-purity and C-containing alloy had similarly longer fatigue lives over the O-containing alloy. Most cracks initiated at particle/void assemblies (PVAs), and the high degree of scatter in HCF measurements was attributed to the stochasticity among PVAs. Regardless of alloy purity, surface scratches remain potent fatigue crack initiation sites. The added fatigue life from reduced C and O content is only realized with very smooth surface finishes [204].

New melting methods have been developed to reduce the size and quantity of inclusions, including a combination of melting and thermo-mechanical processing called "Clean Melt" (CM) [202]. Although CM wires had similar C and O content as standard VIM/VAR wires, the inclusions in CM wires were much smaller both on average (1 µm for CM, 29 µm for standard) and for the largest inclusions (< 15 µm for CM, > 70 µm for standard). The improved cleanliness of the CM wires resulted in two orders of magnitude higher average fatigue life. Furthermore, the CM wire with the shortest fatigue life was still one order of magnitude longer fatigue life than the shortest fatigue life among the standard wires [202].

Recently, fatigue experiments were connected with microscale observations to predict the median probability of failure from the median inclusion diameter using an inverse power law [203].

In summary, significant challenges remain from inclusions in NiTi, especially for biomedical device applications. Future work using high-magnification SEM-DIC deformation mapping around inclusions would bring new observations about these important failure sites.

BIBLIOGRAPHY

[1] Alkan S, Wu Y, Sehitoglu H (2017) Giant non-Schmid effect in NiTi. Extreme Mechanics Letters 15:38–43, doi:10.1016/j.eml.2017.05.003

[2] Wu Y, Ojha A, Patriarca L, Sehitoglu H (2015) Fatigue Crack Growth Fundamentals in Shape Memory Alloys. Shape Memory and Superelasticity doi:10.1007/s40830-015-0005-4

[3] LePage WS, Daly SH, Shaw JA (2016) Cross Polarization for Improved Digital Image Correlation. Experimental Mechanics 56:1–27, doi:10.1007/s11340-016-0129-2

[4] LePage WS, Shaw JA, Daly SH (2017) Optimum Paint Sequence for Speckle Patterns in Digital Image Correlation. Experimental Techniques doi:10.1007/s40799-017-0192-3

[5] LePage WS, Ahadi A, Lenthe WC, Sun QP, Pollock TM, Shaw JA, Daly SH (2017) Grain size effects on NiTi shape memory alloy fatigue crack growth. Journal of Materials Research 1–17, doi:10.1557/jmr.2017.395

[6] Reedlunn B (2012) Thermomechanical Behavior of Shape Memory Alloy Cables and Tubes. Ph.D. thesis, University of Michigan

[7] Shaw JA, Kyriakides S (1997) On the Nucleation and Propagation of Phase Transformation Fronts in a NiTi Alloy. Acta Materialia 45:683–700

[8] Kimiecik M, Jones JW, Daly S (2016) The effect of microstructure on stress-induced martensitic transformation under cyclic loading in the SMA Nickel-Titanium. Journal of the Mechanics and Physics of Solids 89:16–30, doi:10.1016/j.jmps.2016.01.007

[9] Pataky GJ, Ertekin E, Sehitoglu H (2015) Elastocaloric cooling potential of NiTi, Ni2FeGa, and CoNiAl. Acta Materialia 96:420–427, doi:10.1016/J.ACTAMAT.2015.06.011

[10] Kauffman GB, Mayo I (1997) The Story of Nitinol: The Serendipitous Discovery of the Memory Metal and Its Applications. The Chemical Educator 2(2):1–21, doi:10.1007/s00897970111a

[11] Lecce L, Concilio A (2014) Shape memory alloy engineering: for aerospace, structural and biomedical applications. Elsevier

[12] Duerig TW, Pelton AR, Stockel D (1996) The use of superelasticity in medicine. Metall-Heidelberg 50(9):569–574

[13] American Association of Endodontists, Endodontic Treatment Statistics. https://www.aae.org/specialty/about-aae/news-room/endodontic-treatment-statistics/

[14] Healthcare Cost and Utilization Project, Hospital Stays with Cardiac Stents, 2009. https://www.hcup-us.ahrq.gov/reports/statbriefs/sb128.jsp

[15] Everett KD, Conway C, Desany GJ, Baker BL, Choi G, Taylor CA, Edelman ER (2015) Structural Mechanics Predictions Relating to Clinical Coronary Stent Fracture in a 5 Year Period in FDA MAUDE Database. Annals of Biomedical Engineering doi:10.1007/s10439-015-1476-3

[16] Reedlunn B, Daly S, Shaw J (2013) Superelastic shape memory alloy cables: Part 1 Isothermal tension experiments. International Journal of Solids and Structures 50(20-21):3009–3026, doi:10.1016/j.ijsolstr.2013.03.013

[17] Reedlunn B, Daly S, Shaw J (2013) Superelastic shape memory alloy cables: Part II–Subcomponent isothermal responses. International Journal of Solids and Structures 50(20-21):3027–3044

[18] Mas B, Biggs D, Vieito I, Cladera A, Shaw J, Martinez-Abella F (2017) Superelastic shape memory alloy cables for reinforced concrete applications. Construction and Building Materials 148:307–320

[19] Biggs DB (2017) Thermo-Mechanical Behavior and Shakedown of Shape Memory Alloy Cable Structures. Ph.D. thesis, University of Michigan

[20] NASA (2018), Reinventing the Wheel. https://www.nasa.gov/specials/wheels/

[21] Duerig TW, Tolomeo DE, Wholey M (2000) An overview of superelastic stent design. Minimally Invasive Therapy & Allied Technologies 9(3-4):235–246

[22] Smart Technology, Space Age Technology That Will Make The Brazed Joint Obsolete. http://www.smartechnology.com/history.htm

[23] AMF, Osteosynthesis Staples. http://www.nitifrance.com/v2/en-osteosynthesis+staples+19m.html

[24] Dentsply Sirona, WaveOne Gold. https://www.dentsplysirona.com/en-us/products/endodontics/waveone-gold-learn-more.html

[25] Boston Scientific, Rebel Stent System. http://www.bostonscientific.com/en-US/products/stents–coronary/rebel-platinum-chromium-coronary-stent-system.html

[26] Miga Motor Company, A revolution in motion. http://www.migamotors.com/

[27] NASA (2018), NASA Tests New Alloy to Fold Wings in Flight. https://www.nasa.gov/centers/armstrong/feature/nasa-tests-new-alloy-to-fold-wings-in-flight.html

[28] NASA, Shape Memory Alloys. https://www.grc.nasa.gov/WWW/StructuresMaterials/AdvMet/research/shape_memory.html

[29] TiNi Alloy Company, State-of-the-art shape memory alloy actuators. http://www.tinialloy.com/act98.htm

[30] Ruffnax, Quibik, The Hubble Space Telescope as seen from the departing Space Shuttle Atlantis. https://commons.wikimedia.org/wiki/File:HST-SM4.png

[31] Aoki J, Nakazawa G, Tanabe K, Hoye A, Yamamoto H, Nakayama T, Onuma Y, Higashikuni Y, Otsuki S, Yagishita A, Others (2007) Incidence and clinical impact of coronary stent fracture after sirolimus-eluting stent implantation. Catheterization and Cardiovascular Interventions 69(3):380–386

[32] Ino Y, Toyoda Y, Tanaka A, Ishii S, Kusuyama Y, Kubo T, Takarada S, Kitabata H, Tanimoto T, Mizukoshi M, Others (2009) Predictors and prognosis of stent fracture after sirolimus-eluting stent implantation. Circulation Journal 73(11):2036–2041

[33] Nakazawa G, Finn AV, Vorpahl M, Ladich E, Kutys R, Balazs I, Kolodgie FD, Virmani R (2009) Incidence and predictors of drug-eluting stent fracture in human coronary artery: a pathologic analysis. Journal of the American College of Cardiology 54(21):1924–1931

[34] Scheinert D, Scheinert S, Sax J, Piorkowski C, Bräunlich S, Ulrich M, Biamino G, Schmidt A (2005) Prevalence and clinical impact of stent fractures after femoropopliteal stenting. Journal of the American College of Cardiology 45(2):312–315

[35] Calkins FT, Mabe JH, Butler GW (2006) Boeing's variable geometry chevron: morphing aerospace structures for jet noise reduction. vol. 6171, 61710O, International Society for Optics and Photonics, doi:10.1117/12.659664

[36] Shaw JA, Churchill C, Iadicola M (2008) Tips and Tricks for Characterizing Shape Memory Alloy Wire: Part 1-Differential Scanning Calorimetry and Basic Phenomena. Experimental Techniques 32(5):55–62, doi:10.1111/j.1747-1567.2008.00410.x

[37] Churchill CB (2010) Experimental techniques for characterizing the thermo-electro-mechanical shakedown response of SMA wires and tubes. ProQuest Dissertations and Theses 305

[38] Stebner A, Vogel S, Noebe R, Sisneros T, Clausen B, Brown D, Garg A, Brinson L (2013) Micromechanical quantification of elastic, twinning, and slip strain partitioning exhibited by polycrystalline, monoclinic nickeltitanium during large uniaxial deformations measured via in-situ neutron diffraction. Journal of the Mechanics and Physics of Solids 61(11):2302–2330, doi:10.1016/J.JMPS.2013.05.008

[39] Watkins RT, Reedlunn B, Daly S, Shaw JA (2018) Uniaxial, Pure Bending, and Column Buckling Experiments on Superelastic NiTi Rods and Tubes. International Journal of Solids and Structures doi:10.1016/J.IJSOLSTR.2018.01.037

[40] Ahadi A, Matsushita Y, Sawaguchi T, Schaffer J, Sun QQ, Tsuchiya K (2017) Origin of zero and negative thermal expansion in severely-deformed NiTi alloy. Acta Materialia 124:79–92, doi:10.1016/j.actamat.2016.10.054

[41] Bhattacharya K (1992) Self-accommodation in martensite. Archive for Rational Mechanics and Analysis 120(3):201–244, doi:10.1007/BF00375026

[42] Bhattacharya K (2003) Microstructure of martensite: why it forms and how it gives rise to the shape-memory effect, vol. 2. Oxford University Press

[43] Bowers M, Gao Y, Yang L, Gaydosh D, De Graef M, Noebe R, Wang Y, Mills M (2015) Austenite grain refinement during load-biased thermal cycling of a Ni49.9Ti50.1 shape memory alloy. Acta Materialia 91:318–329, doi:10.1016/J.ACTAMAT.2015.03.017

[44] Ahadi A, Sun Q (2013) Stress hysteresis and temperature dependence of phase transition stress in nanostructured NiTi. Applied Physics Letters 103:5–9, doi:10.1063/1.4812643

[45] Sutton MA, Orteu JJ, Schreier HW (2009) Image Correlation for Shape, Motion, and Deformation Measurements: Basic Concepts, Theory, and Applications. Springer, New York,˜

[46] Stinville J, Echlin M, Texier D, Bridier F, Bocher P, Pollock T (2015) Sub-Grain Scale Digital Image Correlation by Electron Microscopy for Polycrystalline Materials during Elastic and Plastic Deformation. Experimental Mechanics 197–216, doi:10.1007/s11340-015-0083-4

[47] Sutton MA, Yan JH, Tiwari V, Schreier HW, Orteu JJ (2008) The effect of out-of-plane motion on 2D and 3D digital image correlation measurements. Optics and Lasers in Engineering 46(10):746–757, doi:10.1016/j.optlaseng.2008.05.005

[48] Franck C, Hong S, Maskarinec S, Tirrell D, Ravichandran G (2007) Three-dimensional full-field measurements of large deformations in soft materials using confocal microscopy and digital volume correlation. Experimental Mechanics 47(3):427–438, doi:10.1007/s11340-007-9037-9

[49] Reu P (2014) All about Speckles : Aliasing. Experimental Techniques 38:1–3

[50] Reu P (2015) All about Speckles: Speckle Density 39:1–2

[51] Reu P (2014) All about Speckles : Speckle Size Measurement. Experimental Techniques 44(11):4–5

[52] Reu P (2015) All about Speckles: Edge Sharpness. Experimental Techniques 39(2):1–2, doi:10.1111/ext.12139

[53] Dong Y, Pan B (2017) A Review of Speckle Pattern Fabrication and Assessment for Digital Image Correlation. Experimental Mechanics 57(8):1161–1181, doi:10.1007/s11340-017-0283-1

[54] Kammers AD, Daly S (2011) Small-scale patterning methods for digital image correlation under scanning electron microscopy. Measurement Science and Technology 22(12):125501, doi:10.1088/0957-0233/22/12/125501

[55] Reu P (2015) Points on Paint. Experimental Techniques 39:1–2

[56] Lionello G, Cristofolini L (2014) A practical approach to optimizing the preparation of speckle patterns for digital-image correlation. Measurement Science and Technology 25(10):107001, doi:10.1088/0957-0233/25/10/107001

[57] Jonnalagadda KN, Chasiotis I, Yagnamurthy S, Lambros J, Pulskamp J, Polcawich R, Dubey M (2010) Experimental Investigation of Strain Rate Dependence of Nanocrystalline Pt Films. Experimental Mechanics 50(1):25–35, doi:10.1007/s11340-008-9212-7

[58] Kammers A, Daly S (2013) Self-Assembled Nanoparticle Surface Patterning for Improved Digital Image Correlation in a Scanning Electron Microscope. Experimental Mechanics 53(8):1333–1341, doi:10.1007/s11340-013-9734-5

[59] Bomarito G, Hochhalter J, Ruggles T, Cannon A (2017) Increasing accuracy and precision of digital image correlation through pattern optimization. Optics and Lasers in Engineering 91:73–85, doi:10.1016/j.optlaseng.2016.11.005

[60] Walter T (2011) Structural architecture of the 1980 Mount St. Helens collapse: An analysis of the Rosenquist photo sequence using digital image correlation. Geology 39(8):767–770, doi:10.1130/G32198.1

[61] Wang X, Pan Z, Fan F, Wang J, Liu Y, Mao S, Zhu T, Xia S (2015) Nanoscale deformation analysis with high-resolution transmission electron microscopy and digital image correlation. Journal of Applied Mechanics 82(12):121001

[62] Reu P (2014) Calibration: Care and Feeding of a Stereo-rig. Experimental Techniques 38(3):1–2, doi:10.1111/ext.12083

[63] Jones EM, Reu PL (2017) Distortion of Digital Image Correlation (DIC) Displacements and Strains from Heat Waves. Experimental Mechanics (Dic):1–24, doi:10.1007/s11340-017-0354-3

[64] Reu P (2016) A Realistic Error Budget for Two Dimension Digital Image Correlation. In Advancement of Optical Methods in Experimental Mechanics, vol. 3, 289–297, doi:10.1007/978-3-319-22446-6_24

[65] Peters WH, Ranson WF, Sutton MA, Chu TC, Anderson J (1983) Application of digital correlation methods to rigid body mechanics. Optical Engineering 22(6):738–742

[66] Sutton MA, Mingqi C, Peters WH, Chao YJ, McNeill SR (1986) Application of an optimized digital correlation method to planar deformation analysis. Image and Vision Computing 4(3):143–150, doi:10.1016/0262-8856(86)90057-0

[67] Bay BK, Smith TS, Fyhrie DP, Saad M (1999) Digital volume correlation: Three-dimensional strain mapping using X-ray tomography. Experimental Mechanics 39(3):217–226, doi:10.1007/BF02323555

[68] Sutton M, Orteu J, Schreier H (2009) Image Correlation for Shape, Motion, and Deformation Measurements. Springer, New York, NY, doi:10.1007/978-0-387-78747-3

[69] Hild F, Roux S (2012) Comparison of Local and Global Approaches to Digital Image Correlation. Experimental Mechanics 52(9):1503–1519, doi:10.1007/s11340-012-9603-7

[70] Reu P (2015) All about Speckles: Contrast. Experimental Techniques 51(4):1–2

[71] Kwak JB (2008) Digital image correlation (DIC) methods for small scale measurement. Master's thesis, Mechanical Engineering, State University of New York at Binghamton

[72] Adams A (1948), Basic Photo One: Camera and Lens

[73] Allphin W (1959), Primer of lamps and lighting

[74] Apolinar E, Rowe WF (1980) Examination of human fingernail ridges by means of polarized light. Journal of forensic sciences 25(1):154–161

[75] Robertson A, Toumba K (1999) Cross-polarized photography in the study of enamel defects in dental paediatrics. Journal of Audiovisual Media in Medicine 22(2):63–70, doi:10.1080/014051199102179

[76] Kawara T, Obazawa H (1980) A New Method for Retroillumination Photography of Cataractous Lens Opacities. American Journal of Ophthalmology 90(2):186–189, doi:10.1016/S0002-9394(14)74852-4

[77] Anderson R (1991) Polarized light examination and photography of the skin. Archives of Dermatology 127(7):1000–1005

[78] Cooper M, Skaggs M, Reu P (2015) High-Speed Stereomicroscope Digital Image Correlation of Rupture Disc Behavior. Society of Experimental Mechanics 2015 Annual Conference

[79] LePage W, Shaw JA, Daly S (2015) Thermomechanical Characterization of Shape Memory Alloy Mode {I} Fracture Behavior. Society of Experimental Mechanics 2015 Annual Conference

[80] Hanrahan P, Krueger W (1993) Reflection from layered surfaces due to subsurface scattering. In Proceedings of the 20th annual conference on Computer graphics and interactive techniques, 165–174, ACM

[81] Pan B, Lu Z, Xie H (2010) Mean intensity gradient: An effective global parameter for quality assessment of the speckle patterns used in digital image correlation. Optics and Lasers in Engineering 48(4):469–477, doi:10.1016/j.optlaseng.2009.08.010

[82] Prewitt J (1970) Object enhancement and extraction. Picture processing and Psychopictorics 10(1):15–19

[83] Kim K, Daly S (2010) Martensite Strain Memory in the Shape Memory Alloy Nickel-Titanium Under Mechanical Cycling. Experimental Mechanics 51(4):641–652, doi:10.1007/s11340-010-9435-2

[84] Kim K, Daly S (2013) The effect of texture on stress-induced martensite formation in nickeltitanium. Smart Materials and Structures 22(7):075012, doi:10.1088/0964-1726/22/7/075012

[85] Woodward C (1851), A familiar introduction to the study of polarized light

[86] Kammers A, Daly S (2013) Digital Image Correlation under Scanning Electron Microscopy: Methodology and Validation. Experimental Mechanics 53(9):1743–1761, doi:10.1007/s11340-013-9782-x

[87] Berfield T, Patel J, Shimmin R, Braun P, Lambros J, Sottos N (2007) Micro-and nanoscale deformation measurement of surface and internal planes via digital image correlation. Experimental Mechanics 47(1):51–62, doi:10.1007/s11340-006-0531-2

[88] Yoneyama S, Kitagawa A, Iwata S, Tani K, Kikuta H (2007) Bridge deflection measurement using digital image correlation. Experimental Techniques (February):34–40, doi:10.1111/j.1747-1567.2007.00132.x

[89] Golden, High Flow Paints <http://www.goldenpaints.com/products/color/colors/high-flow/>

[90] Chemours (2016), Titanium dioxide for coatings, Ti Pure overview brochure

[91] Tilley R (2010) Colour due to Scattering. In Colour and the Optical Properties of Materials, chap. 5, 175–196, doi:10.1002/9780470974773.ch5

[92] Laver M (1986) Titanium Dioxide Whites. In RL Feller, A Roy, editors, Artists' pigments: a handbook of their history and characteristics, Volume III, National Gallery of Art, Washington

[93] ASTM (2011), Standard Test Method for Hiding Power of Paints by Reflectometry (D2805-11). doi:10.1520/D2805-11

[94] JCGM (2008), Evaluation of measurement data - Guide to the expression of uncertainty in measurement. doi:10.1373/clinchem.2003.030528

[95] Reu P, Jones E (2016) Distortion of Full-Field Surface Displacements from Heat Waves. In Society of Experimental Mechanics Annual Conference, 1–2

[96] Lecompte D, Smits A, Bossuyt S, Sol H, Vantomme J, Van Hemelrijck D, Habraken AM (2006) Quality assessment of speckle patterns for digital image correlation. Optics and Lasers in Engineering 44(11):1132–1145, doi:10.1016/j.optlaseng.2005.10.004

[97] Kocks U (1998) Anisotropy and Symmetry. In U Kocks, C Tomé, HR Wenk, editors, Texture and Anisotropy: Preferred Orientations in Polycrystals and their Effect on Materials Properties1, chap. 1, 1–43, Cambridge Univeristy Press, Cambridge, 1 edn.

[98] Wenk HR, Van Houtte P (2004) Texture and anisotropy. Reports on Progress in Physics 67(8):1367–1428, doi:10.1088/0034-4885/67/8/R02

[99] Gao Y, Casalena L, Bowers ML, Noebe RD, Mills MJ, Wang Y (2017) An origin of functional fatigue of shape memory alloys. Acta Materialia 126:389–400, doi:10.1016/j.actamat.2017.01.001

[100] Hu H (1974) Texture of Metals. Texture 1(4):233–258, doi:10.1155/TSM.1.233

[101] Rollett A, Wright S (1998) Typical textures in metals. In U Kocks, C Tomé, HR Wenk, editors, Texture and Anisotropy: Preferred Orientations in Polycrystals and their Effect on Materials Properties, chap. 5, 178–238, Cambridge Univeristy Press, Cambridge

[102] Bachmann F, Hielscher R, Schaeben H (2010) Texture analysis with MTEX - free and open source software toolbox. In Solid State Phenomena, vol. 160, 63–68, Trans Tech Publ, doi:10.4028/www.scientific.net/SSP.160.63

[103] Hielscher R, Schaeben H (2008) A novel pole figure inversion method: Specification of the MTEX algorithm. Journal of Applied Crystallography 41(6):1024–1037, doi:10.1107/S0021889808030112

[104] Houtte P (1987) On the Representation of Texture Functions of Cubic Metals in Euler Space. Textures and Microstructures 7:187–205, doi:10.1155/TSM.7.187

[105] Suwas S, Gurao NP (2014) Representation of Texture. In Crystallographic Texture of Materials, chap. 2, 11–38, Springer-Verlag London, doi:10.1007/978-1-4471-6314-5_2

[106] Wenk HR, Kocks UF (1987) The representation of orientation distributions. Metallurgical Transactions A 18(6):1083–1092

[107] Kocks U (1998) The Representation of Orientations and Textures. In U Kocks, C Tomé, HR Wenk, editors, Texture and Anisotropy: Preferred Orientations in Polycrystals and their Effect on Materials Properties, chap. 2, 44–101, Cambridge Univeristy Press, Cambridge

[108] Matthies S, Helming K, Kunze K (1990) On the Representation of Orientation Distributions in Texture Analysis by sigma-Sections, I. General Properties of sigma-Sections. Physica Status Solidi (B) 71:71–83, doi:10.1002/pssb.2221570105

[109] Matthies S, Helming K, Kunze K (1990) On the Representation of Orientation Distributions in Texture Analysis by sigma-Sections, II. Consideration of Crystal and Sample Symmetry, Examples. Physica Status Solidi (B) 157(1):489–507, doi:10.1002/pssb.2221570202

[110] Stebner A, Gao X, Brown DW, Brinson LC (2011) Neutron diffraction studies and multivariant simulations of shape memory alloys: Empirical texture development-mechanical response relations of martensitic nickel-titanium. Acta Materialia 59(7):2841–2849, doi:10.1016/j.actamat.2011.01.023

[111] Robertson SW, Imbeni V, Wenk HRHR, Ritchie RO (2005) Crystallographic texture for tube and plate of the superelastic/shape-memory alloy Nitinol used for endovascular stents. Journal of biomedical materials research Part A 72(2):190–9, doi:10.1002/jbm.a.30214

[112] Robertson SW, Imbeni V, Wenk HR, Ritchie RO (2006) Erratum: Crystallographic texture for tube and plate of the superelastic/shape-memory alloy nitinol used for endovascular stents. Journal of Biomedical Materials Research Part A 78A(2):432–432, doi:10.1002/jbm.a.30841

[113] Robertson SW, Gong XY, Ritchie RO (2006) Effect of product form and heat treatment on the crystallographic texture of austenitic Nitinol. Journal of Materials Science 41(3):621–630, doi:10.1007/s10853-006-6478-y

[114] Palumbo G, Lin PKY, Tomantschger K, Smith F (2012), Low texture, quasi-isotropic metallic stent

[115] Barney MM, Xu D, Robertson SW, Schroeder V, Ritchie RO, Pelton AR, Mehta A (2011) Impact of thermomechanical texture on the superelastic response of Nitinol implants. Journal of the Mechanical Behavior of Biomedical Materials 4(7):1431–1439, doi:10.1016/j.jmbbm.2011.05.013

[116] Ahadi A, Sun Q (2014) Effects of grain size on the rate-dependent thermomechanical responses of nanostructured superelastic NiTi. Acta Materialia 76:186–197, doi:10.1016/j.actamat.2014.05.007

[117] Mao SC, Luo JF, Zhang Z, Wu MH, Liu Y, Han XD (2010) EBSD studies of the stress-induced B2-B19 martensitic transformation in NiTi tubes under uniaxial tension and compression. Acta Materialia 58(9):3357–3366, doi:10.1016/j.actamat.2010.02.009

[118] Ghassemi-Armaki H, Leff AC, Taheri ML, Dahal J, Kamarajugadda M, Kumar KS (2017) Cyclic compression response of micropillars extracted from textured nanocrystalline NiTi thin-walled tubes. Acta Materialia 136:134–147, doi:10.1016/j.actamat.2017.06.043

[119] Wenk HR (1998) Pole Figure Measurements with Diffraction Techniques. In U Kocks, C Tomé, HR Wenk, editors, Texture and Anisotropy: Preferred Orientations in Polycrystals and their Effect on Materials Properties, chap. 4, 126–177, Cambridge Univeristy Press, Cambridge

[120] Laplanche G, Kazuch A, Eggeler G (2015) Processing of NiTi shape memory sheets - Microstructural heterogeneity and evolution of texture. Journal of Alloys and Compounds 651:333–339, doi:10.1016/j.jallcom.2015.08.127

[121] Nagao K, Kagami E (2011) X-ray thin film measurement techniques: VII. Pole figure measurement. The Rigaku Journal 27(2):6–14

[122] Chantler C, Olsen K, Dragoset R, Chang J, Kishore A, Kotochigova S, Zucker D (2005), NIST Standard Reference Database 66: Detailed Tabulation of Atomic Form Factors, Photoelectric Absorption and Scattering Cross Section, and Mass Attenuation Coefficients for Z = 1-92 from E = 1-10 eV to E = 0.4-1.0 MeV

[123] Robertson SW, Pelton AR, Ritchie RO (2012) Mechanical fatigue and fracture of Nitinol. International Materials Reviews 57(1):1–37, doi:10.1179/1743280411Y.0000000009

[124] Robertson SW, Ritchie RO (2007) In vitro fatigue-crack growth and fracture toughness behavior of thin-walled superelastic Nitinol tube for endovascular stents: A basis for defining the effect of crack-like defects. Biomaterials 28(4):700–9, doi: 10.1016/j.biomaterials.2006.09.034

[125] Paris PC, Gomez MP, Anderson WE (1961) A rational analytic theory of fatigue. The trend in engineering 13:9–14

[126] Elber W (1970) Fatigue Crack Closure Under Cyclic Tension. Engineering Fracture Mechanics 2(1):37–45, doi:10.1016/0013-7944(70)90028-7

[127] Sehitoglu H (1985) Crack opening and closure in fatigue. Engineering Fracture Mechanics 21(2):329–339, doi:10.1016/0013-7944(85)90021-9

[128] Sun Q, Aslan A, Li M, Chen M (2014) Effects of grain size on phase transition behavior of nanocrystalline shape memory alloys. Science China Technological Sciences 57(4):671–679, doi:10.1007/s11431-014-5505-5

[129] Waitz T, Tsuchiya K, Antretter T, Fischer FD (2009) Phase Transformations of Nanocrystalline Martensitic Materials. MRS Bulletin 34(11):814–821, doi:10.1557/mrs2009.231

[130] Duerig T, Pelton A, Sto D (1999) An overview of nitinol medical applications. Materials Science & Engineering A 275:149–160, doi:10.1016/S0921-5093(99)00294-4

[131] Shaw JA, Kyriakides S (1995) Thermomechanical Aspects of NiTi. Journal of the Mechanics and Physics of Solids 43:1243–1281

[132] Li ZQ, Sun QP (2002) The initiation and growth of macroscopic martensite band in nano-grained NiTi microtube under tension. International Journal of Plasticity 18:1481–1498

[133] Chang BC, Shaw JA, Iadicola MA (2006) Thermodynamics of Shape Memory Alloy Wire: Modeling, Experiments, and Application. Continuum Mechanics and Thermodynamics 18:83–118

[134] Ahadi A, Sun Q (2015) Stress-induced nanoscale phase transition in superelastic NiTi by in situ X-ray diffraction. Acta Materialia 90:272–281, doi:10.1016/j.actamat.2015.02.024

[135] Petersmann M, Pranger W, Waitz T, Antretter T (2017) An Energy Approach to Determine the Martensite Morphology in Nanocrystalline NiTi. Advanced Engineering Materials doi:10.1002/adem.201600684

[136] Ahadi A, Sun Q (2016) Grain size dependence of fracture toughness and crack-growth resistance of superelastic NiTi. Scripta Materialia 113:171–175, doi:10.1016/j.scriptamat.2015.10.036

[137] Yin H, He Y, Moumni Z, Sun Q (2016) Effects of grain size on tensile fatigue life of nanostructured NiTi shape memory alloy. International Journal of Fatigue 88:166–177, doi:10.1016/j.ijfatigue.2016.03.023

[138] Leitner T, Sabirov I, Pippan R, Hohenwarter A (2017) The effect of severe grain refinement on the damage tolerance of a superelastic NiTi shape memory alloy. Journal of the Mechanical Behavior of Biomedical Materials doi:10.1016/j.jmbbm.2017.03.020

[139] Prokofiev E, Burow J, Frenzel J, Gunderov D, Eggeler G, Valiev R (2011) Phase Transformations and Functional Properties of NiTi Alloy with Ultrafine-Grained Structure. Materials Science Forum 669:1059–1064, doi:10.4028/www.scientific.net/MSF.667-669.1059

[140] Gall K, Tyber J, Wilkesanders G, Robertson SW, Ritchie RO, Maier HJ (2008) Effect of microstructure on the fatigue of hot-rolled and cold-drawn NiTi shape memory alloys. Materials Science and Engineering A 486(1-2):389–403, doi:10.1016/j.msea.2007.11.033

[141] ASTM International (2013), E647 Standard Test Method for Measurement of Fatigue Crack Growth Rates. doi:10.1520/E0647-13E01.2

[142] McNeill SR, Peters WH, Sutton MA (1987) Estimation of stress intensity factor by digital image correlation. Engineering Fracture Mechanics 28:101–112, doi:10.1016/0013-7944(87)90124-X

[143] Sutton MA, Zhao W, McNeill SR, Helm JD, Piascik RS, Riddell WT (1999) Local Crack Closure Measurements: Development of a Measurement System Using Computer Vision and a Far-Field Microscope. Advances in Fatigue Crack Closure Measurement and Analysis, ASTM STP 1343 145–156

[144] O'Connor S, Nowell D, Dragnevski K (2016) Measurement of fatigue crack deformation on the macro- and micro-scale: Uniform and non-uniform loading. International Journal of Fatigue 89:66–76, doi:10.1016/j.ijfatigue.2016.02.032

[145] Sangid MD, Pataky GJ, Sehitoglu H, Rateick RG, Niendorf T, Maier HJ (2011) Superior fatigue crack growth resistance, irreversibility, and fatigue crack growth-microstructure relationship of nanocrystalline alloys. Acta Materialia 59(19):7340–7355, doi:10.1016/j.actamat.2011.07.058

[146] de Matos PFP, Nowell D (2009) Experimental and numerical investigation of thickness effects in plasticity-induced fatigue crack closure. International Journal of Fatigue 31(11-12):1795–1804, doi:10.1016/j.ijfatigue.2008.12.003

[147] Carroll J, Efstathiou C, Lambros J, Sehitoglu H, Hauber B, Spottswood S, Chona R (2009) Investigation of fatigue crack closure using multiscale image correlation experiments. Engineering Fracture Mechanics 76(15):2384–2398, doi:10.1016/j.engfracmech.2009.08.002

[148] Casperson MC, Carroll JD, Lambros J, Sehitoglu H, Dodds RH (2014) Investigation of thermal effects on fatigue crack closure using multiscale digital image correlation experiments. International Journal of Fatigue 61:10–20, doi:10.1016/j.ijfatigue.2013.11.020

[149] Carroll JD, Abuzaid W, Lambros J, Sehitoglu H (2013) High resolution digital image correlation measurements of strain accumulation in fatigue crack growth. International Journal of Fatigue doi:10.1016/j.ijfatigue.2012.06.010

[150] Sutton MA, Li N, Joy DC, Reynolds AP, Li X (2007) Scanning electron microscopy for quantitative small and large deformation measurements Part I: SEM imaging at magnifications from 200 to 10,000. Experimental Mechanics doi:10.1007/s11340-007-9042-z

[151] Sutton MA, Li N, Garcia D, Cornille N, Orteu JJ, McNeill SR, Schreier HW, Li X, Reynolds AP (2007) Scanning electron microscopy for quantitative small and large deformation measurements Part II: Experimental validation for magnifications from 200 to 10,000. Experimental Mechanics 47, doi:10.1007/s11340-007-9041-0

[152] ASTM International (2016), E561 Standard Test Method for KR Curve Determination. doi:10.1520/E0561-15A.2

[153] Hild F, Roux S (2015) Evaluating damage with digital image correlation: b. from physical to mechanical damage. Handbook of Damage Mechanics: Nano to Macro Scale for Materials and Structures 1277–1299, doi:10.1007/978-1-4614-5589-9_24

[154] Cady CM, Liu C, Rae PJ, Lovato ML (2009) Thermal and loading dynamics of energetic materials. In Proceedings of the SEM Annual Conference, Albuquerque

[155] Liu C, Cady CM, Rae PJ, Lovato ML (2010) On the Quantitative Measurement of Fracture Toughness in High Explosive and Mock Materials. 14th International Detonation Symposium 425–434

[156] Wang LW, Li K, Sanusei S, Ghorbani R, Matta F, Sutton MA (2014) Advancement of Optical Methods in Experimental Mechanics, Volume 3. In H Jin, C Sciammarella, S Yoshida, L Lamberti, editors, Conference Proceedings of the Society for Experimental Mechanics Series, vol. 3 of *Conference Proceedings of the Society for Experimental Mechanics Series*, 289–297, Springer International Publishing, Cham, doi: 10.1007/978-3-319-00768-7_23

[157] Guo SM, Sutton MA, Li N, Li XD, Wang LW, Rajan S (2016) Measurement of Local Thermal Deformations in Heterogeneous Microstructures via SEM Imaging with Digital Image Correlation. Experimental Mechanics 1–16, doi:10.1007/s11340-016-0206-6

[158] Sedlak P, Frost M, Kruisova A, Hivrmanova K, Heller L, Sittner P (2014) Simulations of mechanical response of superelastic NiTi helical spring and its relation to fatigue resistance. Journal of Materials Engineering and Performance 23(7):2591–2598, doi: 10.1007/s11665-014-0906-y

[159] Li Y, Li JY, Liu M, Ren YY, Chen F, Yao GC, Mei QS (2015) Evolution of microstructure and property of NiTi alloy induced by cold rolling. Journal of Alloys and Compounds 653:156–161, doi:10.1016/j.jallcom.2015.09.056

[160] Gupta M, Alderliesten RC, Benedictus R (2014) A review of T-stress and its effects in Fracture Mechanics. Engineering Fracture Mechanics 134:218–241, doi:10.1016/j.engfracmech.2014.10.013

[161] Gao S, Yi S (2003) Experimental study on the anisotropic behavior of textured NiTi pseudoelastic shape memory alloys. Materials Science and Engineering A 362(1-2):107–111, doi:10.1016/S0921-5093(03)00585-9

[162] Xia M, Liu P, Sun Q (2017) Grain Size Effects on Young's Modulus and Hardness of Nanocrystalline NiTi Shape Memory Alloy. In Advances in Shape Memory Materials, doi:10.1007/978-3-319-53306-3

[163] Feng P, Sun QP (2006) Experimental investigation on macroscopic domain formation and evolution in polycrystalline {NiTi} microtubing under mechanical force. Journal of the Mechanics and Physics of Solids 54:1568–1603, doi:10.1016/j.jmps.2006.02.005

[164] Daly S, Ravichandran G, Bhattacharya K (2007) Stress-induced martensitic phase transformation in thin sheets of Nitinol. Acta Materialia 55(10):3593–3600, doi:10.1016/j.actamat.2007.02.011

[165] Hallai JF, Kyriakides S (2013) Underlying material response for Luders-like instabilities. International Journal of Plasticity 47:1–12, doi:10.1016/j.ijplas.2012.12.002

[166] Suresh S, Ritchie RO (1983) Near-threshold fatigue crack propagation: a perspective on the role of crack closure. Tech. rep., Brown Univ., Providence, RI (USA). Div. of Engineering; Lawrence Berkeley Lab., CA (USA)

[167] Liu Y, Xie Z, Van Humbeeck J, Delaey L (1999) Effect of texture orientation on the martensite deformation of NiTi shape memory alloy sheet. Acta Materialia 47(2):645–660, doi:10.1016/S1359-6454(98)00376-0

[168] Gall K, Sehitoglu H, Chumlyakov YI, Zuev YL, Karaman I (1998) The role of coherent precipitates in martensitic transformations in single crystal and polycrystalline Ti-50.8at%Ni. Scripta Materialia 39(6):699–705, doi:10.1016/S1359-6462(98)00236-X

[169] Gall K, Yang N, Sehitoglu H, Chumlyakov YI (2001) Fracture of precipitated NiTi shape memory alloys. International Journal of Fracture 109(2):189–207, doi:10.1023/A:1011069204123

[170] Robertson S, Mehta A, Pelton AR, Ritchie R (2007) Evolution of crack-tip transformation zones in superelastic Nitinol subjected to in situ fatigue: A fracture mechanics and synchrotron X-ray microdiffraction analysis. Acta Materialia 55(18):6198–6207, doi:10.1016/j.actamat.2007.07.028

[171] Creuziger a, Bartol L, Gall K, Crone W (2008) Fracture in single crystal NiTi. Journal of the Mechanics and Physics of Solids 56(9):2896–2905, doi:10.1016/j.jmps.2008.04.002

[172] Sgambitterra E, Maletta C, Furgiuele F, Sehitoglu H (2018) Fatigue crack propagation in [0 1 2] NiTi single crystal alloy. International Journal of Fatigue 112:9–20, doi:10.1016/j.ijfatigue.2018.03.005

[173] Irwin GR (1957) Analysis of stresses and strains near the end of a crack traversing a plate. J appl Mech

[174] Paul PP, Fortman M, Paranjape HM, Anderson PM, Stebner AP, Brinson LC (2018) Influence of Structure and Microstructure on Deformation Localization and Crack

Growth in NiTi Shape Memory Alloys. Shape Memory and Superelasticity 1–9, doi: 10.1007/s40830-018-0172-1

[175] Callahan PG, Stinville JC, Yao ER, Echlin MP, Titus MS, De Graef M, Gianola DS, Pollock TM (2018) Transmission scanning electron microscopy: Defect observations and image simulations. Ultramicroscopy 186:49–61, doi:10.1016/J.ULTRAMIC.2017.11.004

[176] Trimby PW (2012) Orientation mapping of nanostructured materials using transmission Kikuchi diffraction in the scanning electron microscope. Ultramicroscopy 120:16–24, doi:10.1016/J.ULTRAMIC.2012.06.004

[177] Keller R, Geiss R (2012) Transmission EBSD from 10 nm domains in a scanning electron microscope. Journal of Microscopy 245(3):245–251, doi:10.1111/j.1365-2818.2011.03566.x

[178] Frens G (1973) Controlled Nucleation for the Regulation of the Particle Size in Monodisperse Gold Suspensions. Nature Physical Science 241(105):20–22, doi:10.1038/physci241020a0

[179] Lenthe WC, Stinville J, Echlin MP, Chen Z, Daly SH, Pollock TM (2018) Advanced Detector Signal Acquisition and Electron Beam Scanning for High Resolution SEM Imaging. Submitted to Ultramicroscopy

[180] Shaw J, Kyriakides S (1997) Initiation and propagation of localized deformation in elasto-plastic strips under uniaxial tension. International Journal of Plasticity 13(10):837–871, doi:10.1016/S0749-6419(97)00062-4

[181] Dong L, Zhou R, Wang X, Hu G, Sun Q (2015) On interfacial energy of macroscopic domains in polycrystalline NiTi shape memory alloys. International Journal of Solids and Structures 80:445–455, doi:10.1016/j.ijsolstr.2015.10.006

[182] Holtz RL, Sadananda K, Imam MA (1999) Fatigue thresholds of Ni-Ti alloy near the shape memory transition temperature. International Journal Of Fatigue 21:S137–S145, doi:10.1016/S0142-1123(99)00065-1

[183] McKelvey AL, Ritchie RO (2001) Fatigue-Crack Growth Behavior in the Superelastic and Shape-Memory Alloy Nitinol. Metallurgical and Materials Transactions A 32(3):731–743, doi:10.1007/s11661-001-0089-7

[184] Gollerthan S, Young M, Baruj A, Frenzel J, Schmahl W, Eggeler G (2009) Fracture mechanics and microstructure in NiTi shape memory alloys. Acta Materialia 57(4):1015–1025, doi:10.1016/j.actamat.2008.10.055

[185] Baxevanis T, Lagoudas DC (2015) Fracture Mechanics of Shape Memory Alloys: Review and Perspectives. International Journal of Fracture 191(1-2), doi:10.1007/s10704-015-9999-z

[186] Maletta C, Sgambitterra E, Niccoli F (2016) Temperature dependent fracture properties of shape memory alloys: novel findings and a comprehensive model. Scientific Reports doi:10.1038/s41598-016-0024-1

[187] Baxevanis T, Landis CM, Lagoudas DC (2013) On the Fracture Toughness of Pseudoelastic Shape Memory Alloys. Journal of Applied Mechanics 81(4):041005, doi:10.1115/1.4025139

[188] Baxevanis T Landis CM LDC (2014) On the Effect of Latent Heat on the Fracture Toughness of Pseudoelastic Shape Memory Alloy. J Appl Mech

[189] Launey ME, Ritchie RO (2009) On the Fracture Toughness of Advanced Materials. Advanced Materials 21(20):2103–2110, doi:10.1002/adma.200803322

[190] Griffith A (1921) The phenomena of rupture and flow in solids. Philiosophical Transactions of the Royal Society of London 163–197

[191] Burdekin FM, Stone DEW (1966) The crack opening displacement approach to fracture mechanics in yielding materials. Journal of Strain Analysis 1(2):145–153, doi:10.1243/03093247V012145

[192] Wells A (1961) Unstable crack propagation in metals: cleavage and fast fracture. In Proceedings of the crack propagation symposium, vol. 1

[193] Cottrell A (1962) Theoretical Aspects of Radiation Damage and Brittle Fracture in Steel Pressure Vessels. Ing Nucleare 4

[194] Wells A (1969) Crack opening displacements from elastic-plastic analyses of externally notched tension bars. Engineering Fracture Mechanics 1(3):399–410, doi:10.1016/0013-7944(69)90001-0

[195] Rice J, Sorensen E (1978) Continuing crack-tip deformation and fracture for plane-strain crack growth in elastic-plastic solids. Journal of the Mechanics and Physics of Solids 26(3):163–186, doi:10.1016/0022-5096(78)90007-8

[196] Shih C, DeLorenzi H, Andrews W (1979) Studies on Crack Initiation and Stable Crack Growth. In J Landes, J Begley, G Clarke, editors, Elastic-Plastic Fracture, ASTM International

[197] Kanninen M, Rybicki E, Stonesifer R, Broek D, Rosenfield A, Marschall C, Hahn G (1979) Elastic-Plastic Fracture Mechanics for Two-Dimensional Stable Crack Growth and Instability Problems. In J Landes, J Begley, G Clarke, editors, Elastic-Plastic Fracture, ASTM International

[198] Shih C (1981) Relationships between the J-integral and the crack opening displacement for stationary and extending cracks. Journal of the Mechanics and Physics of Solids 29(4):305–326, doi:10.1016/0022-5096(81)90003-X

[199] Elber W (1971) The significance of fatigue crack closure. In Damage tolerance in aircraft structures, ASTM International

[200] Sutton MA, Deng X, Ma F, Newman Jr JC, James M (2000) Development and application of a crack tip opening displacement-based mixed mode fracture criterion. International Journal of Solids and Structures 37(26):3591–3618, doi:10.1016/S0020-7683(99)00055-4

[201] Pelton BL, Slater T, Pelton AR (1997) Effects of Hydrogen in TiNi. In Proceedings of the International Conference on Shape Memory and Superelastic Technologies

[202] Coda A, Cadelli A, Zanella M, Fumagalli L (2018) Straightforward Downsizing of Inclusions in NiTi Alloys: A New Generation of SMA Wires with Outstanding Fatigue Life. Shape Memory and Superelasticity 1–7, doi:10.1007/s40830-018-0159-y

[203] Adler P, Frei R, Kimiecik M, Briant P, James B, Liu C (2018) Effects of Tube Processing on the Fatigue Life of Nitinol. Shape Memory and Superelasticity 1–21, doi:10.1007/s40830-018-0153-4

[204] Rahim M, Frenzel J, Frotscher M, Pfetzing-Micklich J, Steegmüller R, Wohlschlögel M, Mughrabi H, Eggeler G (2013) Impurity levels and fatigue lives of pseudoelastic NiTi shape memory alloys. Acta Materialia 61(10):3667–3686, doi:10.1016/j.actamat.2013.02.054

[205] Reinoehl M, Bradley D, Bouthot R, Proft J (2000) The influence of melt practice on final fatigue properties of superelastic NiTi wires. In Proceedings of the International Conference on Shape Memory and Superelastic Technologies, Pacific Grove, CA

[206] Sawaguchi TA, Kausträter G, Yawny A, Wagner M, Eggeler G (2003) Crack initiation and propagation in 50.9 at. pct Ni-Ti pseudoelastic shape-memory wires in bending-rotation fatigue. Metallurgical and Materials Transactions A 34(12):2847–2860, doi: 10.1007/s11661-003-0186-x

[207] Vaidyanathan R, Dunand D, Ramamurty U (2000) Fatigue crack-growth in shape-memory NiTi and NiTiTiC composites. Materials Science and Engineering: A 289(1-2):208–216, doi:10.1016/S0921-5093(00)00882-0

[208] Khalil-Allafi J, Dlouhy A, Eggeler G (2002) Ni4Ti3-precipitation during aging of NiTi shape memory alloys and its influence on martensitic phase transformations. Acta Materialia 50(17):4255–4274, doi:10.1016/S1359-6454(02)00257-4

[209] Tirry W, Schryvers D (2005) Quantitative determination of strain fields around Ni4Ti3 precipitates in NiTi. Acta Materialia 53(4):1041–1049, doi:10.1016/j.actamat.2004.10.049

[210] Gall K, Sehitoglu H, Chumlyakov Y, Kireeva I (1999) Tensioncompression asymmetry of the stressstrain response in aged single crystal and polycrystalline NiTi. Acta Materialia 47(4):1203–1217, doi:10.1016/S1359-6454(98)00432-7

[211] Michutta J, Somsen C, Yawny A, Dlouhy A, Eggeler G (2006) Elementary martensitic transformation processes in Ni-rich NiTi single crystals with Ni4Ti3 precipitates. Acta Materialia 54(13):3525–3542, doi:10.1016/J.ACTAMAT.2006.03.036

[212] Frick CP, Ortega AM, Tyber J, Maksound AEM, Maier HJ, Liu Y, Gall K (2005) Thermal processing of polycrystalline NiTi shape memory alloys. Materials Science and Engineering A 405(1-2):34–49, doi:10.1016/j.msea.2005.05.102

[213] Norwich DW, Fasching A (2009) A study of the effect of diameter on the fatigue properties of NiTi wire. Journal of Materials Engineering and Performance 18(5-6):558–562, doi:10.1007/s11665-009-9415-9

[214] Wong S, Lin ZC, Tahran A, Boylan J, Pike K, Kramer-Brown P (2009) An Investigation of Factors Impacting Nitinol Wire Fatigue Life. Journal of ASTM International 6(8)

[215] Kramer GM (2009) A comparison of chemistry and inclusion distribution and morphology versus melting method of NiTi alloys. Journal of Materials Engineering and Performance 18(5-6):479–483, doi:10.1007/s11665-009-9438-2

[216] Coda A, Zilio S, Norwich D, Sczerzenie F (2012) Characterization of Inclusions in VIM/VAR NiTi Alloys. Journal of Materials Engineering and Performance 21(12):2572–2577, doi:10.1007/s11665-012-0366-1