NIST Special Publication 260
NIST SP 260-227-upd1

Certification of Standard Reference Material® 3461

Reference Cantilevers for AFM Spring Constant Calibration

Richard S. Gates
William A. Osborn
Mark J. McLean
Gordon A. Shaw
James J. Filliben

This publication is available free of charge from:
https://doi.org/10.6028/NIST.SP.260-227-upd1

NIST NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY
U.S. DEPARTMENT OF COMMERCE

NIST Special Publication 260
NIST SP 260-227-upd1

Certification of Standard Reference Material® 3461

Reference Cantilevers for AFM Spring Constant Calibration

Richard S. Gates
William A. Osborn
Mark J. McLean
Materials Measurement Science Division (MMSD)
Material Measurement Laboratory (MML)

Gordon A. Shaw
Quantum Measurement Division
Physical Measurement Laboratory (PML)

James J. Filliben
Statistical Engineering Division
Information Technology Laboratory (ITL)

This publication is available free of charge from:
https://doi.org/10.6028/NIST.SP.260-227-upd1

October 2023
INCLUDES UPDATES AS OF 10-16-2023; SEE APPENDIX E

U.S. Department of Commerce
Gina M. Raimondo, Secretary

National Institute of Standards and Technology
Laurie E. Locascio, NIST Director and Under Secretary of Commerce for Standards and Technology

NIST SP 260-227-upd1
October 2023

Publication History
Approved by the NIST Editorial Review Board on 2022-06-03
Supersedes NIST 260-227 (August 2022) https://doi.org/10.6028/NIST.SP.260-227

How to Cite this NIST Technical Series Publication
Gates RS, Osborn WA, McLean MJ, Shaw GA, Filliben JJ (2022) Certification of Standard Reference Material® 3461 Reference Cantilevers for AFM Spring Constant Calibration. (National Institute of Standards and Technology, Gaithersburg, MD), NIST Special Publication 260 (NIST SP 260) NIST SP 260-227-upd1. https://doi.org/10.6028/NIST.SP.260-227-upd1

NIST Author ORCID iDs
Richard S. Gates: 0000-0003-1097-8494
William A. Osborn: 0000-0003-4238-4893
Mark McLean: 0000-0003-0336-5972
Gordon A. Shaw: 0000-0002-5881-5443
James J. Filliben: 0000-0002-2388-3198

Abstract

A batch of 94 silicon chips containing arrays of uniform rectangular cantilevers was microfabricated from a single silicon wafer and their stiffnesses and resonance frequencies calibrated using a laser Doppler vibrometer (LDV). The cantilevers are 50 μm wide, (about the width of a human hair), and vary in length from 600 μm to 300 μm. They have stiffness values nominally ranging from 0.03 N/m to 0.2 N/m and nominal first flexural resonance frequencies in the range of 5 kHz to 21 kHz respectively. The LDV cantilever stiffness calibration measurements proved very reliable and combined expanded uncertainties better than ± 3% in stiffness were obtained for individual devices. These devices can be used as reference artifacts to validate atomic force microscopy (AFM) cantilever stiffness measurement methods as well as directly calibrating AFM test cantilever stiffness using the reference cantilever method.

This NIST SP-260 provides details on the source and preparation of the devices as well as the analytical approach of using a laser Doppler vibrometer to accurately measure both stiffness (k), resonance frequency (f_0), and Quality factor (Q) of each cantilever. The statistical analysis applied to the data sets is described. An example of the reference cantilever method is also provided for individuals wishing to use these devices to make direct calibrations of AFM test cantilevers.

Keywords

Atomic Force Microscopy (AFM); Calibration; Cantilever; Electrostatic Force Balance (EFB); Laser Doppler Vibrometer (LDV); Microfabrication; Reference Cantilever Method; Thermal Calibration Method; Resonance Frequency; Standard Reference Material (SRM); Stiffness, Système International (SI).

Table of Contents

1. Introduction 1
2. Cantilever Array Design 1
3. Source and Preparation 3
4. Laser Doppler Vibrometry (LDV) Thermal Calibration Method 4
5. Statistical Analysis of Calibration Measurement Uncertainty 7
6. SI Traceability – LDV Comparison to EFB 8
7. Test AFM Calibration using the Reference Cantilever Method 10
References 11
Appendix A. LORENTZIAN PEAK FITTING MODEL USED FOR THE LDV THERMAL CALIBRATION METHOD 13
Appendix B. STATISTICAL ANALYSIS 16
Appendix C. SI TRACEABILITY THROUGH COMPARISON WITH ELECTROSTATIC FORCE BALANCE 22
Appendix D. REFERENCE CANTILEVER CALIBRATION METHOD 26
Appendix E. CHANGE LOG 30

List of Tables

Table 1 Comparison of LDV and EFB cantilever calibrations on the same cantilevers 9

List of Figures

Fig. 1 Plan design view of a single SRM 3461 device 2
Fig. 2 Diagram of an ideal, rectangular, uniform, fixed-free cantilever 2
Fig. 3 Schematic cross section of the production wafer prior to lithography showing the silicon handle, device, and clamping (green) layers interspaced with silicon oxide layers. 4
Fig. 4 SRM 3461 die location identification (left to right) from wafer scale, stepper scale, and individual die. The wafer contains 52 stepper patterns. Each stepper pattern contains 12 dies (devices). 4
Fig. 5 LDV sample laser spot placed at the end of a cantilever 7
Fig. 6 Direct comparison of LDV versus EFB stiffness calibrations on four of the SRM 3461 cantilevers. 9

Glossary

AFM – atomic force microscope

EFB – electrostatic force balance

LDV – laser Doppler vibrometer

f_0 – first flexural mode frequency

k_z – static stiffness in the direction perpendicular to the chip

Q – quality factor of the resonator (the cantilever)

1. Introduction

Atomic force microscopy (AFM) is a popular technique for interrogating surfaces on the micro and nano scales. The most common use for AFM is imaging; however, there are a variety of more specialized AFM techniques that can be used to determine electrical, mechanical, and chemical properties of surfaces. To adequately control the application of forces to surfaces for these techniques (especially mechanical property measurements), accurate stiffness calibrations of test cantilevers should be used. There are a variety of test cantilever stiffness calibration techniques available, based on dimensional, static force and displacement, and dynamic vibrational methods [1], but in general, these have large uncertainties in the range of ± 10% to ± 30% and no Système International (SI) traceability (i.e., unknown accuracy). More rigorous calibrated balance techniques [2], [3] with SI traceability have been pioneered, mostly by National Metrology Institutes (NMIs), but their complexity, expense, and time-consuming operation make them an out-of-reach technique for most AFM researchers. The reference cantilevers represented by NIST SRM 3461 are an accurate and precise force calibration artifact for use in the field.

2. Cantilever Array Design

SRM 3461 comprises an array of seven uniform cross-section, rectangular cantilevers of different lengths attached to the end of a silicon handle chip similar in size and thickness to commercial AFM test cantilever chips. Each cantilever is nominally 50 µm wide and 1.45 µm thick and range in length from 300 µm through 600 µm in steps of 50 µm. This provides seven different reference stiffnesses at the end of each cantilever. The design of the array with the longest cantilever (600 µm) in the middle and decreasing lengths on both left (500, 400, 300) µm and right (550, 450, 350) µm was intended to provide a rough symmetry that might be beneficial in moderating potentially damaging residual stresses that are encountered during microfabrication processing. The cantilevers are spaced laterally at a 150 µm pitch. The handle chip, similar in size to commercial AFM test cantilever chips, was designed with two side supports, thinned near the chip to facilitate removal of the chip from the wafer after processing. Each chip contains a NIST logo, the SRM 3461 label, and a serialization text at the bottom identifying the wafer, stepper die, and device within the stepper field.

To determine reasonable cantilever sizes for the target stiffnesses, Euler-Bernoulli beam theory [4] is a useful predictive tool. The flexural stiffness (k_z) of ideal fixed-free cantilevers is given by:

$$k_z = \frac{Ebt^3}{4L^3} \quad (1)$$

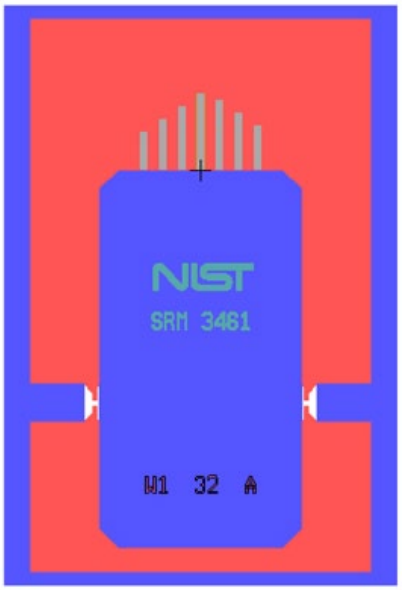

Fig. 1 Plan design view of a single SRM 3461 device

where E is the elastic modulus of the beam in the direction of L, and b, t, and L are width, thickness, and length respectively as indicated in Fig 2.

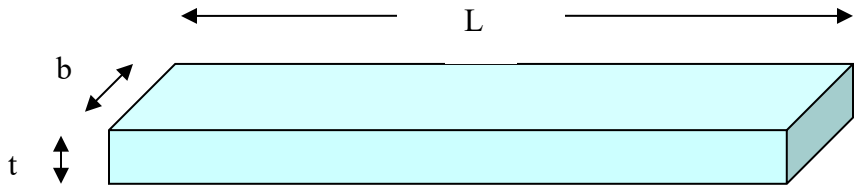

Fig. 2 Diagram of an ideal, rectangular, uniform, fixed-free cantilever

This model points out dimensional dependencies for stiffness and clearly demonstrates that cantilever stiffness variations are particularly sensitive to uncertainties in thickness and length since these terms are cubed. This in turn dictates the need for highly uniform thickness silicon device layers, very precise patterning for cantilever length, and the need for a lower design bound (300 μm) on the cantilever length as methods to minimize cantilever stiffness variation for the SRM 3461 devices.

3. Source and Preparation

The devices were produced from a 150 mm diameter silicon on insulator (SOI) wafer sourced for its extremely high uniformity in thickness of the device layer that would be used to form the main structural element – the cantilever. The wafer manufacturer (Soitec[1], France) specified the device thickness of the particular batch of "SmartCut" SOI wafers purchased as 1,503 nm ± 28 nm (at six sigma). The starting wafers were modified by a series of RCA clean, oxidation, and polysilicon deposition steps to allow the creation of a silicon clamping layer sandwiched by thin oxide films on top of the device layer. This design feature allowed tighter control on the final length plan dimensions of the individual cantilevers in the array by substantially decoupling it from the creation of the handle chip. The handle chip is made from the thick handle portion of the SOI wafer and requires deep Si etching though the back thickness of the handle wafer (several hundred micrometers) with much larger lateral dimensional uncertainty after etching. The final cross-sectional schematic for the modified pre-production wafer is depicted in Fig. 3. The device layer was thinned slightly due to RCA cleaning and controlled silicon oxide growth steps resulting in a final estimated device layer thickness of 1.45 μm. All Si layers are structural elements with the device layer providing the key cantilever component. All oxide layers serve as both etch stops during critical silicon etching processes when they are lithographically exposed and strong bonding elements between silicon layers where they are lithographically protected.

The SRM 3461 devices were defined by an extensive series of lithographic patterning and etching steps performed in microfabrication facilities at the Cornell Nanofabrication Facility (CNF) in Ithaca, NY and at the NIST Center for Nanoscale Science and Technology (CNST) in Gaithersburg, MD.

The front side lithographic process involved the use of very precise step-and-repeat (stepper) technology to repeatably define the key length and width plan dimensions for the cantilevers and is capable of sub-micrometer resolutions. Individual labels for the devices were incorporated on the front side using direct laser writing lithography. The overall scaling of the wafer, stepper, and individual dies are shown in Fig. 4. Each SRM 3461 device (die) has a unique three component identifier (W2 XX Y) describing the wafer number (e.g., W2 for this SRM), stepper cell number (XX from 01 through 52) within the wafer, and die letter (Y from A through L) within each stepper pattern. The complexity of the microfabrication process is such that only 94 of a possible 624 devices survived the process perfectly and yielded complete, accurate, and uniform devices. Of these, 90 are being sold under this SRM.

[1] "Certain commercial equipment, instruments, or materials are identified in this paper to adequately specify the experimental procedure. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the materials or equipment identified are necessarily the best available for the purpose."

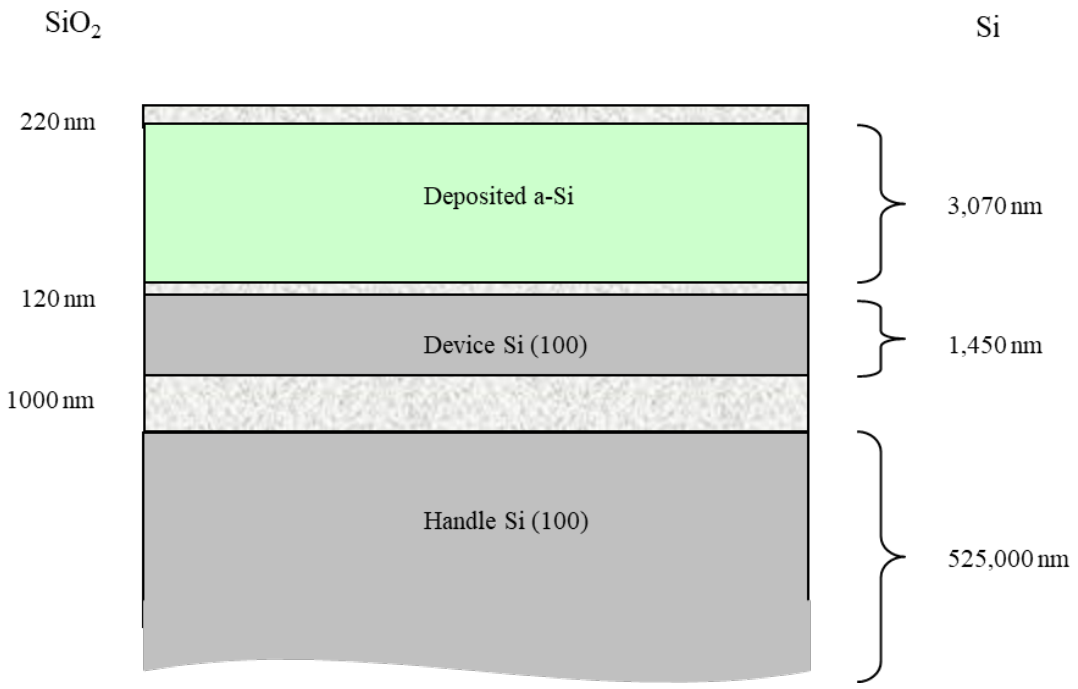

Fig. 3 Schematic cross section of the production wafer prior to lithography showing the silicon handle, device, and clamping (green) layers interspaced with silicon oxide layers.

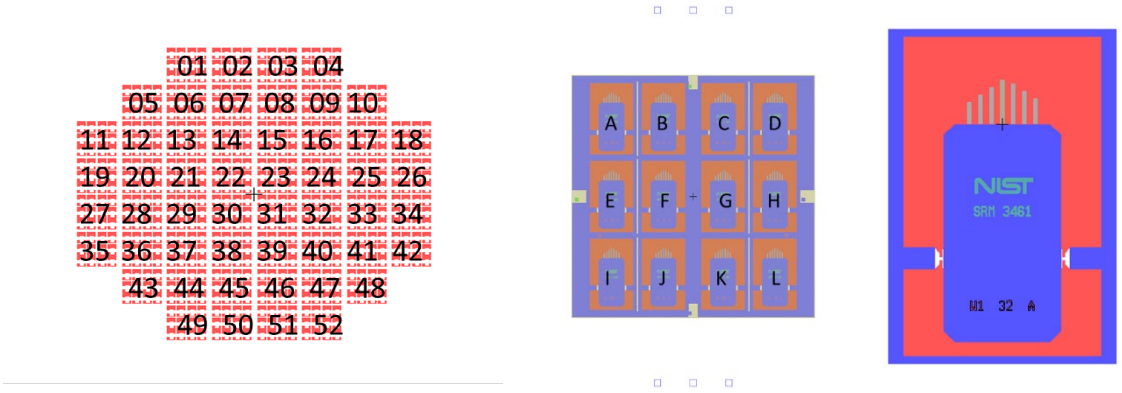

Fig. 4 SRM 3461 die location identification (left to right) from wafer scale, stepper scale, and individual die. The wafer contains 52 stepper patterns. Each stepper pattern contains 12 dies (devices).

4. Laser Doppler Vibrometry (LDV) Thermal Calibration Method

The thermal calibration method for measuring the stiffness of a cantilever is based on the equipartition theorem which states that the average energy of a system (e.g., cantilever) in thermal equilibrium with its environment is ½ $k_B T$ per degree of freedom. As a vibrational mode (e.g., flexural) is a degree of freedom for a cantilever, the average thermal, kinetic, and elastic strain energies are all equal. This average energy can be expressed mathematically as:

$$\frac{1}{2}k_BT = \frac{1}{2}k\langle z^2\rangle \quad (2)$$

where k_B is the Boltzmann constant, T is absolute temperature, k is the cantilever flexural stiffness (potential energy term) and $\langle z^2\rangle$ is the mean squared flexural displacement of thermal fluctuations in the cantilever (i.e., all orders of the same degree of freedom). In the case of ideal uniform rectangular cantilevers this can be tailored to measurements of a single (usually first) flexural resonance frequency as:

$$k = \frac{12k_BT}{\beta_i^4\langle z_i^2\rangle} \quad (3)$$

where βi is the eigenvalue for the i^{th} resonance mode (1.8751 for the first flexural mode of an ideal rectangular cantilever) and $\langle z_i^2\rangle$ is the mean squared displacement for that mode. For an LDV measurement of the first resonance peak in a spectrum this simplifies to:

$$k = \frac{0.9707*k_B*T}{Area} \quad (4)$$

Where Area is the total energy area under the first resonance peak of the cantilever. In a practical sense this requires an LDV measurement of the frequency spectrum on a Power Spectral Density (PSD) scale and fitting the data to a Lorentzian model to extract the area under the curve. A detailed summary of the specific Lorentzian model used in our Mathematica analysis of each spectrum is provided in Appendix A along with an explanation of how the key parameters of resonance frequency (f_0), Quality factor (Q) and *Area* under the resonance peak are obtained from the fit parameters for this model.

The particular Polytec MSA 500 instrument used [5] to calibrate the cantilevers in this SRM is incorporated into a microscope with variable magnification capabilities. It is dual beam (having both sample and reference laser spots) and can electronically steer the sample laser spot on the sample. This, coupled with the ability to tap into the Doppler signal strength indicator in the instrument, allowed the development of a Visual Basic macro for automating the calibration routine by consistently locating the sample laser spot tangent to the end of each cantilever and acquiring the spectra. A small cubic correction was applied to each measurement to account for the fact that the center of the laser spot was actually 3.2 μm (half the laser spot diameter) back from the end of the cantilever. An example of this "off-end correction" (OEC) is shown below for the LDV laser spot placement on a 600 μm cantilever:

$$OEC_{600} = \left(1 - \frac{3.2}{600}\right)^3 = 0.984 \; or \; 98.4\,\% \quad (5)$$

Which indicates that the stiffness at the very end of the 600 μm cantilever is 98.4 % of the value measured by the LDV (i.e., a 1.6 % reduction). The other cantilevers are handled in a similar manner. The off-end spatial correction of 3.2 μm is constant for all cantilever lengths, but the off-end stiffness correction varies from 98.4 % to 96.8 % depending on cantilever length.

All values reported in the SRM are for the exact end of each cantilever since it serves as a suitable fiducial for users to compare to where their comparison stiffness measurements are made. The device calibration routine utilized six replicate measurements on each of the seven different cantilevers on each device which provided sufficient statistical integrity.

A critical assumption in the thermal calibration method is that the sole driving force for vibrating the cantilevers comes from thermal energy in the ambient environment, therefore great lengths were taken to eliminate other possible driving source interferences such as acoustic and mechanical vibration from reaching the samples during calibration. The NIST LDV system used is housed 12 m underground in a special vibration isolated laboratory within the NIST Advanced Measurement Laboratory (AML). In addition to flooring that is vibrationally isolated from adjacent labs and hallways, the LDV system is placed on top of a dedicated vibration isolation air table. The cantilevers have first flexural resonance frequencies within the human audible range so even speech could be detrimental to calibration measurements. A custom small acoustic isolation shroud was constructed around the sample stage to reduce possible acoustic interferences from within the laboratory. Finally one last potential source of vibrational contamination (the researcher) was removed during each data acquisition routine which lasted approximately 15 minutes and ran autonomously. The laboratory is also a class 1000 clean room environment with robust humidity (± 5% RH) and temperature (± 0.01 °C) control.

The practical aspects of using LDV to measure the stiffness involved placing an SRM 3461 cantilever array on the sample stage at a magnification sufficient to encompass all 7 cantilevers on the device as well as the handle chip base. A macro automation program was written for the system to steer the sample laser spot to the end of each cantilever while the stationary reference beam was placed on the handle chip. This allowed for a repeatable placement of the sample spot so that a series of six replicate measurements could be made on each of the seven cantilevers. LDV measurements of each chip took place autonomously after the researcher had left the laboratory and required less than 15 minutes.

The acquired raw data files were then analyzed using a custom Mathematica program to fit each spectrum to a Lorentzian model and use the fit parameters to calculate the desired stiffness (k), resonance frequency (f_0) and quality factor (Q) values provided in this SRM. Details of the Lorentzian model used and the screening procedure to ensure that no acquisition anomalies occurred during the calibration measurements are provided in Appendix A.

A more rigorous statistical analysis of the measurement expanded uncertainties for f_0, Q, and k, based on the entire population of devices and an additional repeatability side-study on a single device is provided below.

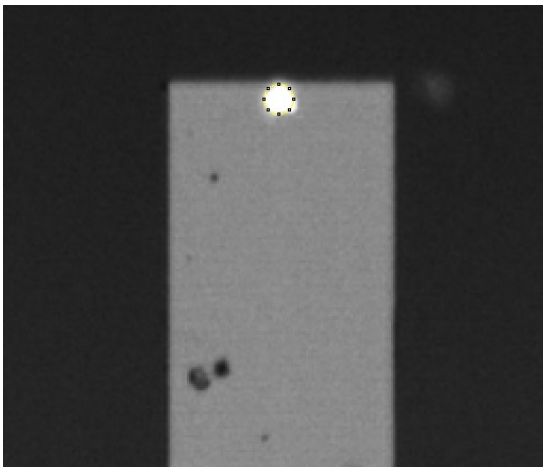

Fig. 5 LDV sample laser spot placed at the end of a cantilever

5. Statistical Analysis of Calibration Measurement Uncertainty

A rigorous statistical analysis was performed of the entire set of data acquired on the 94 devices comprising SRM 3461 to estimate the combined expanded uncertainties of the measured responses of stiffness (k), resonance frequency (f_0), and quality factor (Q). This utilized the base data set of 94 devices, with 7 different cantilevers per device and 6 repeat tests per cantilever for a total of 3,948 observations for each of the three responses. The main emphasis was on the stiffness values since they are NIST certified. The informational values of resonance frequency and Q do not warrant a critical uncertainty statement, so the combined expanded uncertainty analysis was used only to approximate an upper bound on the typical uncertainty of those measurements for all cantilever lengths.

Details of the statistical analysis methodology are available in Appendix B which itself represents a condensed summary of a comprehensive 200-page report of analysis performed on this SRM by statistician James J. Filliben of the NIST Statistical Engineering Division. The relevant conclusions of the analysis are summarized here.

The first question to address was whether the stiffness values for SRM 3461 devices should be certified "globally" (one value for each length cantilever on all devices) or "locally" (individually for each cantilever and device).

In the case of the certified values of stiffness, it was statistically determined that certified values and uncertainties were poorly served by assigning them globally, failing mainly for homogeneity criteria across the 94 devices. This was somewhat expected because the devices are microfabricated through a complex process with hundreds of steps and there may be small variations in the dimensions of the cantilevers (from plasma etch variations across the wafer for example) which may introduce small but real variations in stiffness. The LDV thermal calibration method has the sensitivity and precision to measure these differences.

Statistical analysis of the uncertainties applied to "local" (individual) certification focused on applying adequate statistical rigor to the main uncertainty components of replication and repeatability. In the case of LDV stiffness calibrations, replication is typically represented by the

n=6 measurement observations that are made during the calibration of each cantilever on a device where each cantilever is measured six times and the mean and standard deviation are obtained.

A second important component of uncertainty is the repeatability which represents the uncertainty in repeating the same measurement sequence (typically at other times or days). Initially this was introduced using a small side-study where the device LDV calibration was performed on a single SRM 3461 device 8 times (repeats). While these replication and repeatability components provided a good logical approach for a combined expanded uncertainty it was hampered by the variability associated with the small number of observations (n=6 and n=8) in each case. This deficiency was addressed by utilizing global statistical uncertainty values from the larger n=94 data set as more consistent representations for the replication and repeatability components. The case study outlined in Appendix B describes the systematic exploration of six methodologies to find an optimal combined expanded uncertainty for use in the stiffness certification of the SRM 3461 devices. The final combined expanded uncertainty method selected (method 5 of the 6 investigated) had reasonable and consistent uncertainties across all cantilever lengths (in a range of ± 2.5% to ± 3.0%). These uncertainties were then applied to the individual stiffness means obtained from the original (n=6) replicate series of tests to deliver the values and uncertainties provided in the certificate for each device.

As a result of these considerations, each unit of SRM 3461 has a unique certified stiffness. The serial number on the certificate and etched into the chip need to match.

6. SI Traceability – LDV Comparison to EFB

Historically, the thermal calibration stiffness measurements made by the NIST LDV instrument used in this study have agreed very well with measurements made using the NIST Electrostatic Force Balance (EFB) which has traceability to the SI. These comparisons have covered a relatively large range of stiffness values (0.03 N/m through 14 N/m) and a variety of cantilevers including rectangular and triangular shaped, tipped, and tipless as well as colloidal probe, and made from both silicon and silicon nitride [5-7] materials. Agreement was typically within 1 %.

For this SRM, a narrower direct comparison was made between LDV and EFB stiffness calibration measurements encompassing the more limited stiffness range of about 0.03 N/m to 0.2 N/m, from the longest (600 μm) to the shortest (300 μm) cantilevers. Four different cantilevers from two dies were independently measured by both LDV and EFB. One cantilever (W2 46 A 300) was measured twice by the EFB as an additional spot check of the repeatability of that measurement. Each of the cantilevers had EFB-LDV discrepancies of less than 0.5 %.

A graph of the LDV (Thermal) stiffness versus EFB (Static) stiffness measurements for the four different cantilevers is provided in Fig. 6 showing the strong correlation (slope 1.000) and goodness of fit (R^2 =1.000) between the two sets of data.

Table 1 Comparison of LDV and EFB cantilever calibrations on the same cantilevers

Die	Cantilever	EFB	EFB Unc.*	EFB Unc.*	LDV	LDV Unc.*	LDV Unc.*	EFB – LDV Discrepancy
W2 46 A		k, N/m	SD, N/m	RSD, %	k, N/m	SD, N/m	RSD, %	Δ, %
	300**	0.2273	0.0046	1.9	0.2271	0.00160	0.70	0.09
	400	0.09854	0.0013	1.2	0.0982	0.00121	1.23	0.35
W2 30 D								
	300	0.2242	0.0032	1.3	0.2244	0.00222	0.99	-0.08
	600	0.02868	0.00024	0.82	0.02877	0.000363	1.26	-0.32
							Average:	0.01

*Uncertainties for both EFB and LDV were summed in quadrature combining Type A (statistical uncertainty) + B (placement uncertainty) components.
**Average of two repeat tests using EFB which had k of about 0.2262 N/m & 0.2282 N/m

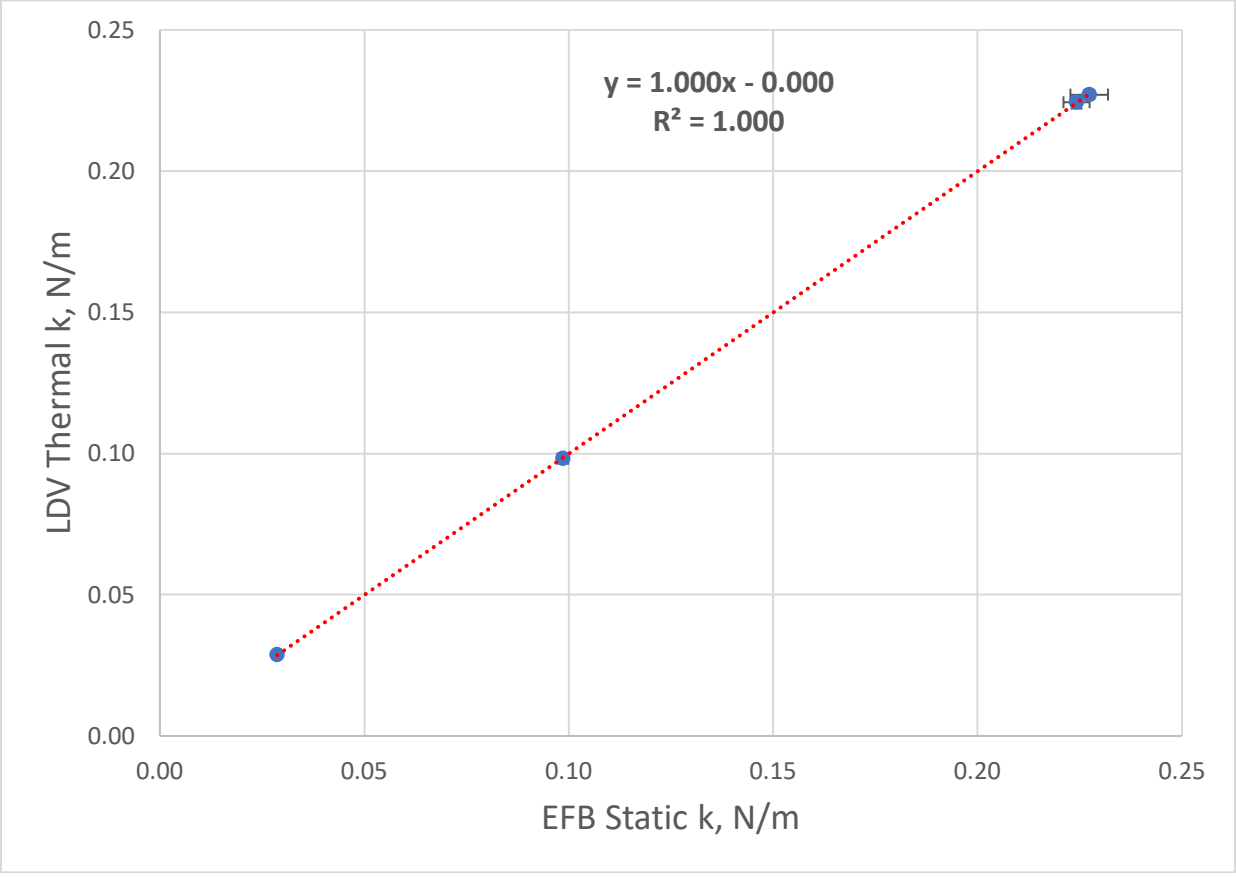

Fig. 6 Direct comparison of LDV versus EFB stiffness calibrations on four of the SRM 3461 cantilevers.

On average, the LDV stiffness measurements are smaller than EFB stiffness measurements by only 0.01 %. A Student's t-test conducted on the comparison data indicated that the results were not statistically different at the usual 0.05 level. Notwithstanding this observation, the contribution of a 0.01% error bias was estimated by adding it in Quadrature with the previous combined expanded uncertainty values for stiffness and it increased the uncertainties by less than 0.001%. This was considered insignificant considering the already conservative combined expanded uncertainties of 2.5% to 3.0% prescribed and no further adjustment in the uncertainties was made.

7. Test AFM Calibration using the Reference Cantilever Method

The reference cantilever method is a direct and useful method for calibrating individual AFM test cantilevers. Cantilevers are springs and coupling springs in series and applying forces (monitoring the length of each spring) - where one spring stiffness is known (the reference) and one is unknown (the test), allows calculation of the stiffness of the test spring. In the case of cantilevers, the AFM test cantilever (which is inclined in its holder in the instrument) is placed above a horizontal reference cantilever and pressed down on it as the photodiode output of the laser lever registers the tilt of the end of the cantilever. A second measurement on an infinitely stiff surface (e.g., the silicon chip off the base of the cantilever) essentially calibrates the tilt vs displacement of the cantilever and the stiffness of the test cantilever can then be estimated. The details of the procedure for conducting the reference calibration method are provided in Appendix D.

Three aspects of a test cantilever calibration measurement are important to the successful application of the reference cantilever method.

First, it is assumed that the reference cantilever array chip is firmly mounted and the force application on the end of the cantilever displaces the end of the cantilever only and no chip motion takes place. The reference cantilever method can apply μN of force; therefore, it is advisable to re-mount the NIST reference cantilever array artifact onto a stiffer surface in a more secure way to avoid chip motion that would corrupt the calibration measurement. A suggested procedure for re-mounting is provided in Appendix D.

Second, because the reference calibration method utilizes physical contact between the test cantilever tip and the reference cantilever, there is the possibility of damage occurring to the tip. It is therefore suggested that the test cantilever calibration procedure be conducted AFTER any experimental measurements in case the calibration procedure damages the tip.

Thirdly, physical contact between the test cantilever and the reference cantilever must necessarily occur on a known location on the reference cantilever away from the exact end of the cantilever and an "off-end correction" (OEC) applied to transfer the calibrated stiffness of the reference cantilever (at the end) to the known contact location. This small correction is cubic with length as described in Appendix D and since they are all inside the end of the cantilever they will all be small increases in reference stiffness at the new measurement contact point.

The basic reference cantilever calibration method requires collecting approach and retract force curves on both an infinitely stiff surface and the compliant reference cantilever surface to obtain the stiffness of the test cantilever. The SRM 3461 devices have small 50 μm square recesses

near the base of each cantilever that are specifically designed to provide an optimal "rigid" surface to use. They were made from etching away the same clamping layer originally above the cantilever, exposing the top of the device layer in the same way as is found on the top of the cantilever so specific surface chemistries and textures should be identical.

In addition to single reference cantilever calibrations which have uncertainties of near ± 10%, multipoint reference cantilever methods have been demonstrated with much lower uncertainty, as low as ± 2%, that can be used in cases where a more rigorous calibration is desired. These methods include using a cantilever array similar to this SRM [8] as well as "beam walking" methods where several calibrations performed along the length of a standard cantilever artifact are used [9].

Finally, since the reference cantilever method relies on measured deflections of both the test and reference cantilever it will be optimized for calibrations where the stiffness of each cantilever is the same. Uncertainties will increase as the two cantilever stiffness values deviate from one another due to the increasing difficulty in estimating the smaller deflection. In general, good results are obtained when the cantilevers are within a factor of three of each other, but often reasonable results can be obtained within an order of magnitude of the reference cantilever stiffness but with a larger uncertainty penalty.

References

[1] Ohler B (2007) Application Note 94: Practical Advice on the Determination of Cantilever Spring Constants.

[2] Pratt JR, Smith DT, Newell DB, Kramar JA, Whitenton E (2004) Progress toward Systeme International d'Unites traceable force metrology for nanomechanics. *J Mater Res* 19(1):366-379. https://doi.org/10.1557/jmr.2004.19.1.366

[3] Kim MS, Pratt JR, Brand U, Jones CW (2012) Report on the first international comparison of small force facilities: a pilot study at the micronewton level. *Metrologia* 49(1):70-81. https://doi.org/10.1088/0026-1394/49/1/011

[4] Timoshenko S , Goodier JN (1969) *Theory of elasticity* (McGraw-Hill, New York,), 3d Ed., pp xxiv, 567 p.

[5] Gates RS , Pratt JR (2012) Accurate and precise calibration of AFM cantilever spring constants using laser Doppler vibrometry. *Nanotechnology* 23(37):375702. https://doi.org/10.1088/0957-4484/23/37/375702

[6] Gates RS, Osborn WA, Pratt JR (2013) Experimental determination of mode correction factors for thermal method spring constant calibration of AFM cantilevers using laser Doppler vibrometry. *Nanotechnology* 24(25):255706. https://doi.org/10.1088/0957-4484/24/25/255706

[7] Gates RS, Osborn WA, Shaw GA (2015) Accurate flexural spring constant calibration of colloid probe cantilevers using scanning laser Doppler vibrometry. *Nanotechnology* 26(23). https://doi.org/10.1088/0957-4484/26/23/235704

[8] Gates RS , Reitsma MG (2007) Precise atomic force microscope cantilever spring constant calibration using a reference cantilever array. *Rev Sci Instrum* 78(8). https://doi.org/10.1063/1.2764372

[9] Ying ZC, Reitsma MG, Gates RS (2007) Direct measurement of cantilever spring constants and correction for cantilever irregularities using an instrumented indenter. *Rev Sci Instrum* 78(6):-. https://doi.org/10.1063/1.2747095

[10] Shaw GA, Stirling J, Kramar JA, Moses A, Abbott P, Steiner R, Koffman A, Pratt JR, Kubarych ZJ (2016) Milligram mass metrology using an electrostatic force balance. *Metrologia* 53(5):A86-A94. https://doi.org/10.1088/0026-1394/53/5/A86

[11] Stirling J , Shaw GA (2017) Realising traceable electrostatic forces despite non-linear balance motion. *Measurement Science and Technology* 28(5). https://doi.org/10.1088/1361-6501/aa5e15

[12] Upstill C (1988) Numerical Recipes in C. *Nature* 333(6174):613-614. https://doi.org/10.1038/333613a0

[13] Pratt JR, Shaw GA, Kumanchik L, Burnham NA (2010) Quantitative assessment of sample stiffness and sliding friction from force curves in atomic force microscopy. *J Appl Phys* 107(4). https://doi.org/10.1063/1.3284957

[14] Gates RS, Pratt JR (2006) Prototype cantilevers for SI-traceable nanonewton force calibration. *Meas Sci Technol* 17(10) 2852 https://doi.org/10.1088/0957-0233/17/10/041

Appendix A. LORENTZIAN PEAK FITTING MODEL USED FOR THE LDV THERMAL CALIBRATION METHOD

Certification of SRM 3461 required collecting and processing over 100 GB of LDV data, including calculating Fast Fourier Transforms (FFTs) and fitting Lorentzian peak shapes millions of times. To facilitate analyzing this quantity of data, the process was automated using Mathematica. The commented code is provided in the SRM 3461 data publication, (https://doi.org/10.18434/T4/1503158) but this section describes the functions of this code in a narrative format.

LDV data was collected using a custom Sax Basic macro within the Polytec data acquisition software, which provided 42 (.pvd format) files (6 replicates × 7 cantilevers). The acquisition macro uses the Doppler signal intensity to find the end of the cantilever by walking the sample laser spot off the end of the cantilever and then pulling back a prescribed amount (approximately 3.2 μm). The macro does this 6 times per cantilever with small lateral steps in location across the cantilever as it goes to generate 6 replicate measurements per cantilever. The macro saves one .pvd file for each measurement.

Each .pvd file contains a displacement time series that is 32,768,000 data points long ($2^{15} \times 100$). The Mathematica code works through .NET to the Polytec File Access API (freely available from Polytec) to import the displacement time series. The time series is partitioned into record lengths of 65,536 (2^{16}) and then 50 records are FFT'ed and the results are converted from displacement versus frequency into power spectral density (PSD) versus frequency. The average of 50 records is used as the spectrum to pass to the fitting routine.

The fitting routine is a multistep process that includes initial guesses, linear pre-fits to refine the guesses, and then non-linear curve fitting that uses the refined guesses. The non-linear curve fitting uses uniform weighting (i.e., unweighted) and the Levenberg-Marquardt method to fit the Lorentzian function

$$y = bkgdm\, x^{-2} + bkgdb + \frac{a}{1+\left(\frac{x-c}{b}\right)^2} \quad \text{(A1)}$$

to the averaged FFT of each replicate measurement. All five fit parameters report a best fit value and a standard error (1σ).

The fit parameters relate to general spectral features where:

bkgdm = the slope of the Brownian ($1/f^2$) noise component of the baseline (m^2 Hz)
bkgdb = the noise floor component of the baseline (m^2/Hz)
a = the height of the resonant peak (m^2/Hz)
b = the half-width half-max of the resonant peak (Hz)
c = the resonance frequency of the peak (Hz)

Using these fitting parameters one can then define useful spectral peak terms such as quality factor (Q) and Area under the resonance peak.

$$Q = \frac{c}{2\,b} \tag{A2}$$

$$\text{Area} = a\,b\,\pi \qquad \text{(units are m}^2\text{)} \tag{A3}$$

If the thermal data is plotted as power spectral density (m^2/Hz vs Hz) then the mean square amplitude of the resonance peak is equal to the area under the fitted peak:

$$k = \frac{\chi_f \cdot k_B \cdot T}{\text{Area}} \tag{A4}$$

where χ_f is the mode correction factor for the specific resonance peak (0.9707 for the first flexural resonance peak for uniform rectangular cantilevers); k_B is the Boltzman constant (1.381×10^{-23} J/K); T is the temperature in degrees Kelvin (294 K); and Area is the area under the first flexural mode resonance peak (m^2). The resulting stiffness, k, is in N/m.

Note that the uncertainty in k lies mainly in the uncertainty of the area estimation from fitting the data. This can be obtained by adding the relative uncertainties of the a and b fitting parameters in quadrature.

Although the fitting routine described generally above (and in detail in the code, available in the SRM 3461 data publication) is robust, non-linear curve fitting is best performed under a watchful eye using a thorough "human check." This was performed by generating a graphics output summary page for each chip's 42 non-linear curve fit results and inspecting the results.

An example Mathematica screening analysis of a set of 6 repeat thermal calibration acquisitions of the 7 cantilevers of a single device is presented below in Fig. A 1. Raw spectra data are plotted (each of the 6 replicates represented by a different color), offset for clarity, as a visual check that each first flexural resonance peak obtained is clean and artifact free. Each peak was then fit to the 5 component Lorentzian model and the residuals plotted, overlapping, shown underneath. Fit values and the derived stiffness values are plotted as box plots with means (black line) and standard deviations (gray shaded areas) for visual reference. Summary text of the device identifier and key data (f_0, Q, k) are also provided along with the standard deviation of the mean uncertainties for the six measurements.

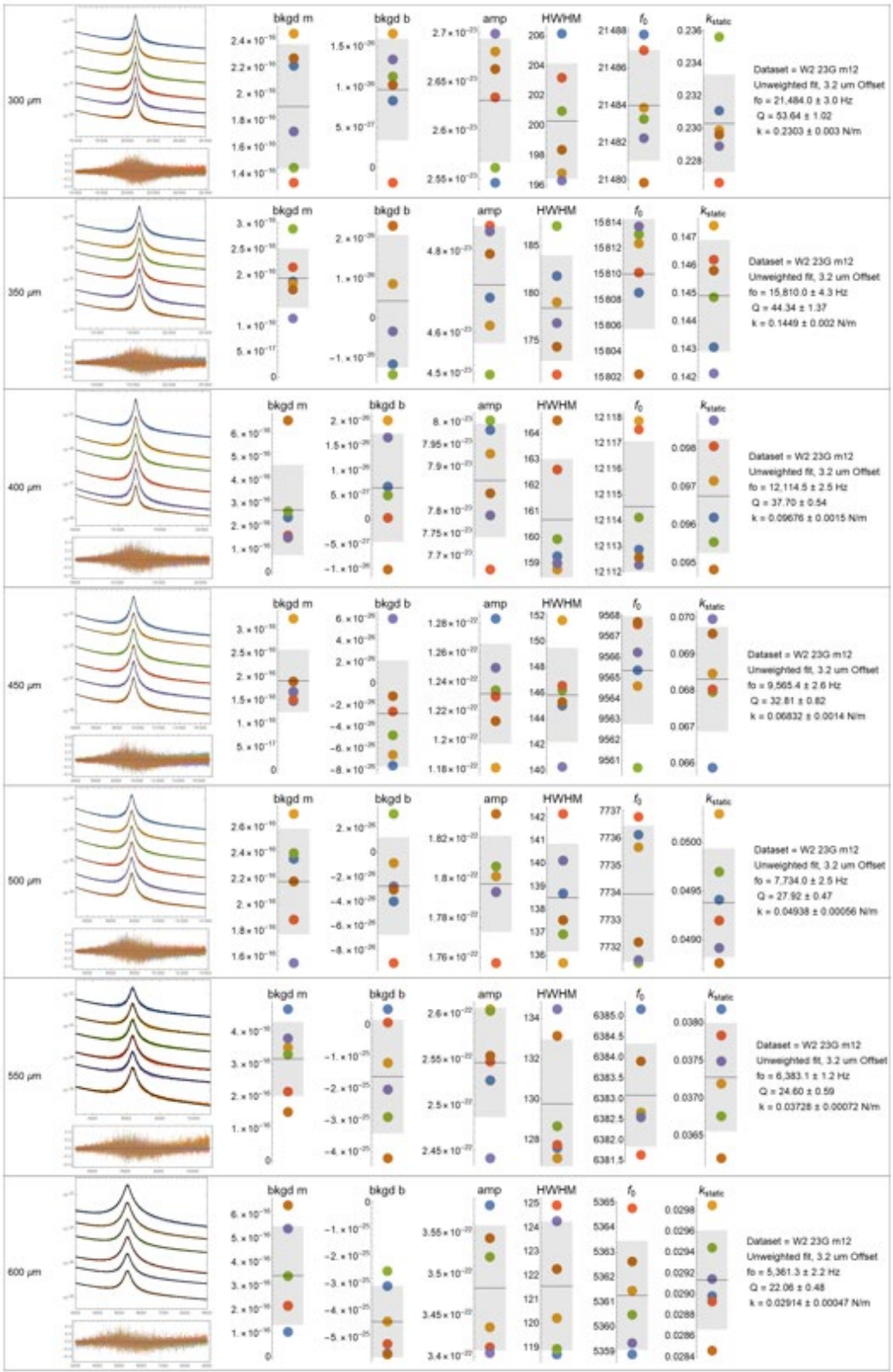

Fig. A 1 Human check summary screening page for a device.

Appendix B. STATISTICAL ANALYSIS

Two options were available for certification of the SRM 3461 devices. The first was to pool all the data and certify the devices "Globally" such that each length of cantilever would have the same calibrated stiffness value and uncertainties across all devices. The second, "local", option involved using the six repeat test data for each of the seven cantilevers to generate the stiffness data for that particular device.

Initially, a "Global" study of the data was performed using a variety of both general (3) and specific (10) statistical methodology tools to probe the key reliability aspects of the measurement process including:

- Homogeneity
- Stability
- Outliers
- Statistical Control
- Consensus Reference Value
- Uncertainty of the Consensus Reference Value

Examples of some of these analyses are shown below for the 300 µm long cantilevers using all of the raw data (n = 6 × 94 = 564 observations).

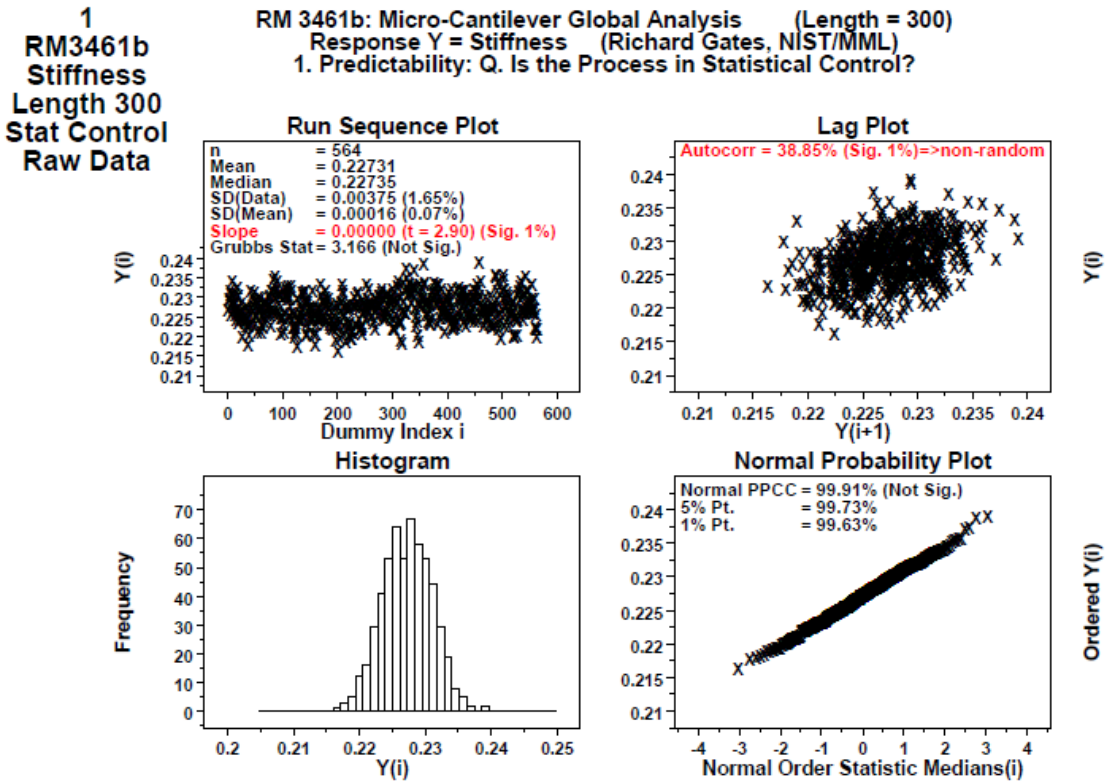

Fig. B 1 Global (n=564) stiffness statistical control example for 300 µm long cantilevers.

and also for just the mean stiffness data (n=94),

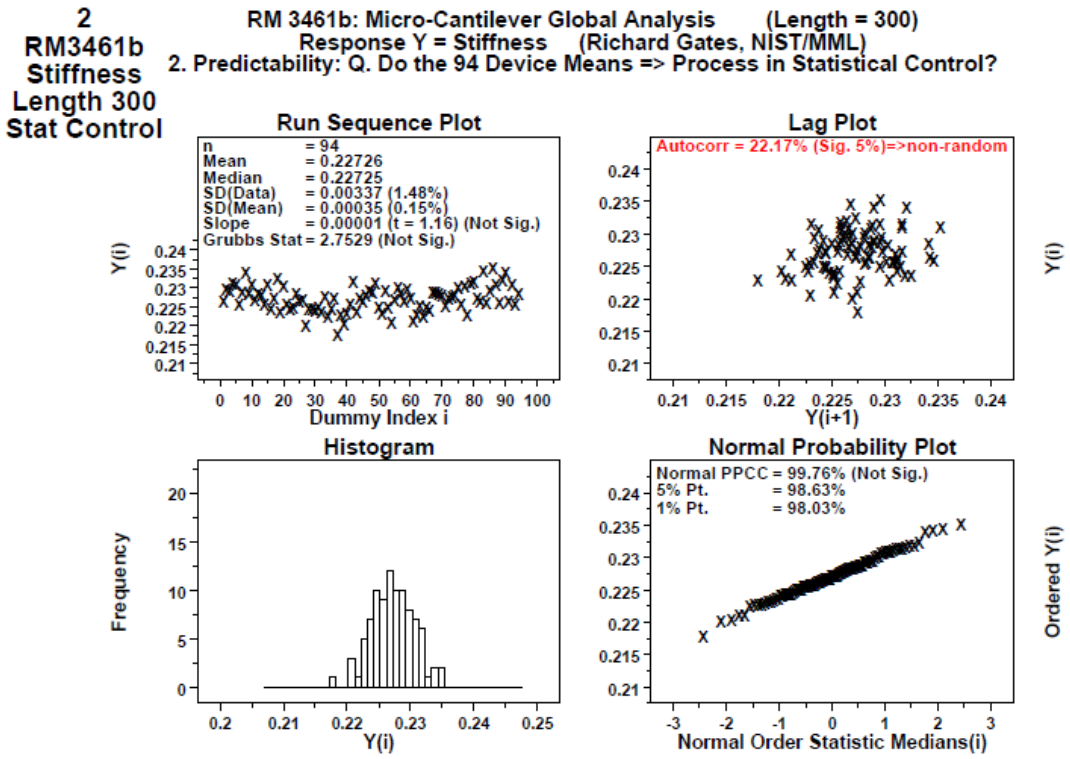

Fig. B 2 Global mean (n=94) stiffness statistical control example for 300 µm long cantilevers.

as well as consensus value plots for all devices:

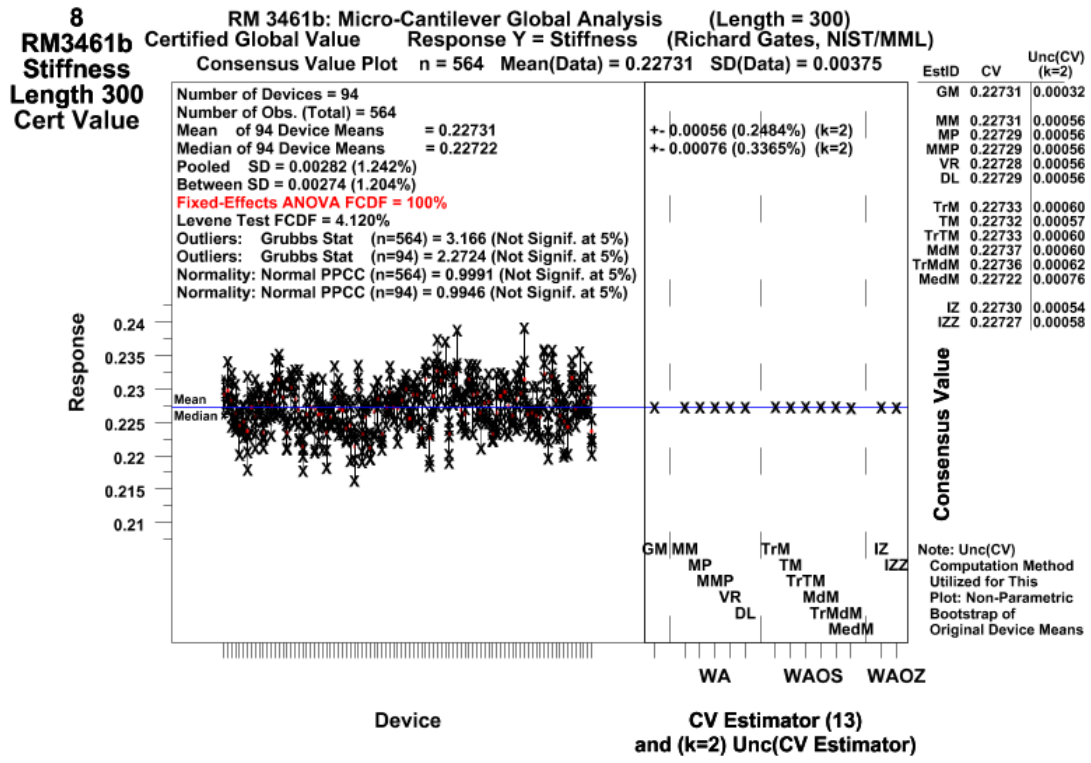

Fig. B 3 Global mean (GM) stiffness consensus value (n=94) example for 300 µm long cantilevers.

These analyses were also repeated for the other six cantilever lengths. The conclusion of the Global study was that because of the presence of significant outliers and stability test issues, the 94 devices were not statistically homogeneous enough across all devices to justify a Global SRM certification with a reasonably low uncertainty and thus local (individual) certifications would be warranted.

Exploration of the local (Individual) certification option for the SRM 3461 devices involved use of the six repeat measurements to establish the mean (stiffness) value for each cantilever of the seven cantilevers on each device. An example for a single device, shown below where each row represents the cantilever length (300 µm, 350 µm, …, 600 µm) and the first five columns represent scatter plots (first one in order of replication and the rest sorted low to high) with overlaid uncertainty limits (standard deviation of the data, Confidence Limits (CL), Prediction Limits (PL), and Tolerance Limits (TL)). The text in the last column provides numerical values for the mean (M) and accompanying limit uncertainties as both stiffness values and percent. This was performed on all 94 devices comprising SRM 3461.

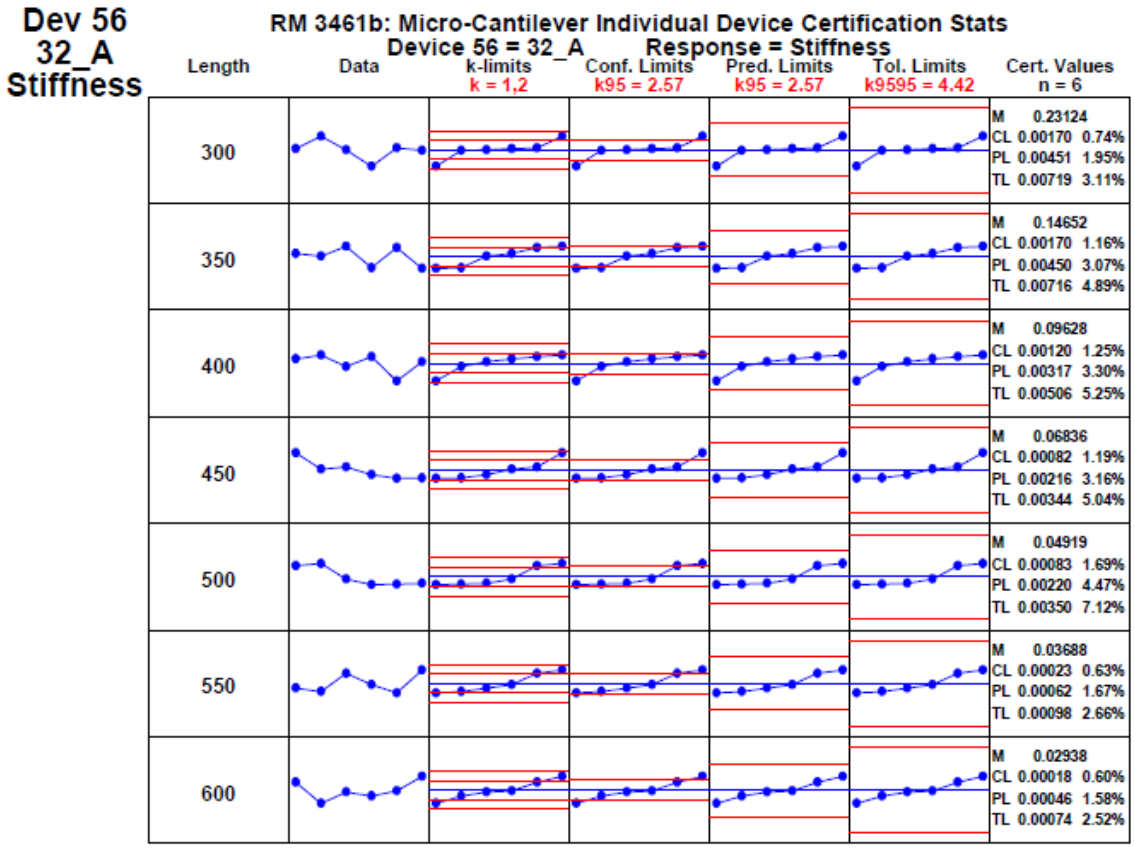

Fig. B 4 Individual device stiffness certification example (n=6) for device 32_A.

While this approach provided a reasonable value for the mean stiffness, the consistency of the uncertainty limits varied considerably from cantilever to cantilever and device to device because of the relatively small (n=6) number of observations. In the example of the one device shown above, the statistical uncertainty can be seen to vary by almost a factor of three (e.g., from 0.60%

to 1.69% for the confidence limits uncertainties). This was observed across all 94 devices examined. Attempts were made to provide a more consistent replicate uncertainty estimate by utilizing the larger (n=94*6=564) observations from the Global analysis as well as adding a repeatability analysis contribution to the overall uncertainty estimates. The repeatability contribution was first attempted using a small repeat study performed on a single device in which the LDV calibration protocol (n=6) was repeated eight times (n=8) as well as a more Global repeatability estimate from the whole dataset (n=94). This was individualized for each of the 7 cantilever lengths. The resulting exploration of options for defining a realistic, reliable, consistent, and conservative estimate of the uncertainty of the stiffness measurements based on the data acquired during the calibration of the devices became a case study in finding an optimal method.

The six methods explored were:

1. Method 1: Replication (n=6) for both mean & standard deviation
2. Method 2: Replication (n=6) mean + Global Pooled SD (n=94)
3. Method 3: Replication (n=6) mean + Global Pooled SD (n=94) + Repeatability (n=8)
4. Method 4: Replication (n=6) mean + Global Max SD (n=94) + Repeatability (n=8)
5. Method 5: Replication (n=6) mean + Global Pooled SD (n=94) + Global Between (n=94)
6. Method 6: Replication (n=6) mean + Global Max SD (n=94) + Global Between (n=94)

Where multiple uncertainty contributions were utilized in a method, the contributions were summed in quadrature. Example summary data for each of the methods are provided below for a single cantilever (300 μm) from a representative device (32A) but it should be noted that the methodologies were tested on different length cantilevers and devices to observe the consistency of the method uncertainties for these different cases:

Method 1	0.23124 ± 0.00133 (0.574%) (k=2)
Method 2	0.23124 ± 0.00234 (1.014%) (k=2)
Method 3	0 23124 ± 0.00380 (1.644%) (k=2)
Method 4	0 23124 ± 0.00518 (2.241%) (k=2)
Method 5	0 23124 ± 0.00604 (2.613%) (k=2)
Method 6	0 23124 ± 0.00670 (3.025%) (k=2)

The results demonstrate that the methods become progressively more conservative with each version. The methods also become more consistent (less “noisy”) for different length cantilevers as larger data sets (e.g., n=94 versus n=6 and n=8) were utilized.

In all methods, the reference value itself is the local mean (0.23124) of the 6 runs. As for the k=2 uncertainty, a discussion of the issues, merits, and comparison of the 6 methods is provided below:

Method 1 has the advantage of being completely local but the disadvantage of using only n = 6 observations for the uncertainty resulting in significant variations in uncertainty from case to case. The net effect is the k=2 uncertainty value (0.574%) is optimistically small & unrealistic (and also did not use a repeatability component), and noisy (varies with device and cantilever since it only used a single calibration run, n=6).

Method 2 improves on this by not using the local standard deviation, but by using pooled within-device variation information "borrowed" from the full collection of 94 devices. This will make the final k=2 uncertainties more consistent, but latent information about mean variation (i.e., test set repeatability) is still being unused. The net effect is that the final k=2 uncertainty value (1.014%) is still optimistically small and unrealistic.

Method 3 builds on Method 2 in using the pooled standard deviation and deriving a SD(mean), but improves on Method 2 by making use of mean-to-mean variation (i.e., test set repeatability component) from a side-experiment involving 8 repeat runs (of 6 observations each) for a particular device (23G). The 2 estimates are then combined in quadrature. The advantage of this method is that a more informed statistical analysis (both test uncertainty as well as test-to-test repeatability) is being used to arrive at the final answer. The net effect for this method is that the final k=2 uncertainty value (1.644%) is as expected larger and more realistic.

Method 4 builds on Method 3 but replaces the pooled standard deviation across the 94 devices with the maximum standard deviation across the 94 devices. It then uses this max SD to derive a SD(mean), and then (like Method 3) makes use of the mean-to-mean variation from the side-experiment involving 8 repeat runs (of 6 observations each) for a particular device (23G). As with Method 3, the 2 estimates are then combined in quadrature. The advantage of this method is that more data is being used to arrive at the final answer. The disadvantage is that using the max SD tends to make this method overly conservative for the SD component values. The net effect is that the final k=2 uncertainty value (2.241%) is larger and more realistic. It serves as a good reference point for the 2 upcoming Methods 5 and 6.

Method 5 mimics Method 3 in using the pooled standard deviation and deriving a SD(mean), but differs from Method 3 by replacing Method 3's use of mean-to-mean variation from the 8 23G side-experiments with the use of mean-to-mean variation from the 94 experiments involving all 94 devices. The advantage of this approach is that SD(mean) is based on more devices (94 versus 1 (= 23G)) and more runs (94 v 8), thus more extensive data is being used to arrive at the final answer. The disadvantage is that device-to-device differences are known to exist, so the mean-to-mean differences will reflect not only natural mean variation, but also device-to-device variation. As before, the pooled standard deviation estimate for SD(mean) and the mean-to-mean SD from the 94 devices are combined in quadrature. The net effect is that the final k=2 uncertainty value (2.241%) is larger--but not too large--and more realistic.

Method 6 mimics Method 4 in replacing the pooled SD with the very conservative maximum within-device SD to form a local SD(mean). Method 6 also mimics Method 5 in using the mean-to-mean variation from the 94 experiments involving all 94 devices. The advantage of this approach is that more data is being used to arrive at the final answer. The disadvantage is that the use of the max SD is ultra-conservative. The net effect is that the final k=2 uncertainty value (3.025%) is large--the largest of all 6 methods.

Conclusion: After weighing the pros and cons of each method from a consistent statistical point of view it was deemed that Method 5 was the best and most realistic choice for arriving at the final uncertainty estimates. A summary of the virtues of Method 5 includes:

1. it is conservative
2. it utilizes both within-device replication information and between-device mean-to-mean differences.
3. it makes use of the larger data set--the 94-device set as opposed to the

8 G23 sets.

4. it will yield uncertainty values that are more consistent and less-perturbed by outliers.

A summary table of the results of using Method 5 for each of the cantilevers in device 32A is shown below.

1 Device	2 Length	3 Stiffness Mean (n=6) N/m	4 Stiffness Unc (k=2) N/m	5 Stiffness RelUnc (k=2) %
32A	300	0.23124	0.00604	2.613%
32A	350	0.14652	0.00391	2.666%
32A	400	0.09628	0.00255	2.645%
32A	450	0.06836	0.00173	2.537%
32A	500	0.04919	0.00133	2.702%
32A	550	0.03688	0.00097	2.635%
32A	600	0.02938	0.00087	2.962%

The consistency of the uncertainty is apparent as all values lie between 2.5% and 3.0 %. The relative uncertainty values (column 5) will be applied to all devices in SRM 3461 and multiplied against the local mean stiffness values for each cantilever (e.g., column 3) to produce the stiffness uncertainty values (e.g., column 4) listed in each SRM 3461 device certificate.

It should be noted that the same statistical rigor of using Method 5 was applied to both frequency and Quality Factor (Q) but since these two measured values are supplied as informational only the statistical uncertainty is provided in the footnotes of the certification table merely as guidance as “typically better than 0.3%” (for frequency), and “typically better than 5 % (for Q).

Appendix C. SI TRACEABILITY THROUGH COMPARISON WITH ELECTROSTATIC FORCE BALANCE

The NIST Electrostatic Force Balance (EFB) was used to validate the LDV cantilever spring constant measurements by measuring a subset of SRM 3461 reference artifacts. The EFB is an electromechanical balance using electrostatic force from a capacitor to provide a reference force and laser interferometry to provide displacement traceable to the SI through physical constants [10].

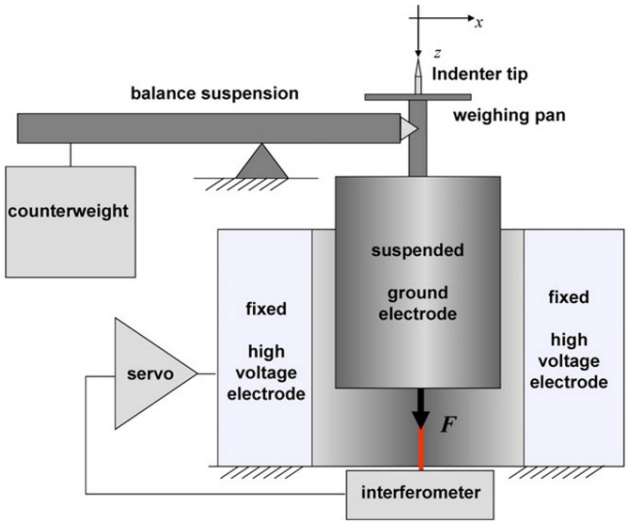

Fig. C 1 Electrostatic Force Balance Schematic, after [14].

The EFB is housed in a vacuum chamber, 12 m underground, in a ± 0.01°C controlled laboratory in NIST's Advanced Materials Laboratory (AML). It sits on a vibrationally isolated metal pad and a local gravitational constant measurement taken within 2 m of the instrument ensures proper calibration accuracy during balance validation experiments.

Electrostatic force is determined using

$$F = \frac{{}^{dC}/_{dz}(V-V_s)^2}{2}. \qquad \text{(C1)}$$

Here dC/dz is the gradient in capacitance, C, with displacement, z, measured as the inner capacitor cylinder is translated within the outer cylinder along a trajectory defined by the balance mechanism [11]. Parameter V is the applied electrical potential during force measurement, and V_s is a surface potential on the electrodes which can be accounted for by averaging voltage measurements at opposite polarity. Capacitance is traceable to an AC/DC transfer of quantum Hall impedance though the NIST Calculable Capacitor, position is traceable to the stabilized HeNe laser frequency used for interferometry, and voltage is traceable to the NIST Josephson voltage standard.

To obtain a Hook's law spring constant, force and displacement, d, of a reference spring must be measured such that

$$k = F/d \tag{C2}$$

In the case of the EFB, both F and d can be measured traceably, as described above. Because both these measurements are traceable, the reference spring constant is traceable as well. By attaching a sharp stylus to the balance, bringing the stylus into contact with the reference cantilever at a known location, and measuring a series of force and displacement values, a combined stiffness, k_m, is measured that is a combination of the balance and cantilever properties.

The cantilever spring constant can be determined from these measurements using

$$k_c = [(k_m - k_b)^{-1} - k_l^{-1}]^{-1} OEC, \tag{C3}$$

where k_c and k_b are cantilever and balance spring constants, respectively. The load frame stiffness of the balance, k_l, has previously been measured to be 7×10^3 N/m [2], and its effect is negligible for the testing of the current reference cantilevers. The value of k_b is measured in a separate experiment when the balance is not in contact with the cantilever.

To position the indenter tip on the reference cantilever, a long stand-off objective lens was used to focus an image of the cantilever and indenter tip onto a video camera operating in the balance enclosure. The cantilever array was attached to a cantilever holder on an x-y-z translation stage, which was in turn mounted rigidly to the balance body. The translation stages used calibrated resistive encoders to provide a measurement of position. The specified repeatability of the stage position is 2 μm, which can be treated as a tolerance such that position uncertainty is 0.67 μm. To locate the test point at the end of the cantilever, the indenter tip is brought just into contact with the cantilever, and the location of the two corners of the rectangular cantilever profile are determined multiple times. A test point on the center of the cantilever's long axis and approximately 5 μm from the end is then used for testing. The distance of the test point from the end of the cantilever is then used to calculate the stiffness off-end correction (OEC), as described previously.

A representative data set for EFB spring constant measurement is shown in Fig. C 2. The data are collected by cycling through 5 consecutive displacements in both the increasing and decreasing directions. No significant dependence on the direction is observed. By taking the average force of the 10 data points in each cycle, and fitting a 10^{th} order polynomial to the timeseries of the averages, a drift correction is determined and subtracted from the original force measurements. A linear fit of the resulting force-displacement curves yields k_m. Quadratic and cubic fits were performed and showed higher order coefficients to be sufficiently small to ignore, as expected for the case where displacement of the cantilever is much less than its length. The mean of between 4 and 10 separate measurements of k_m was determined for each cantilever.

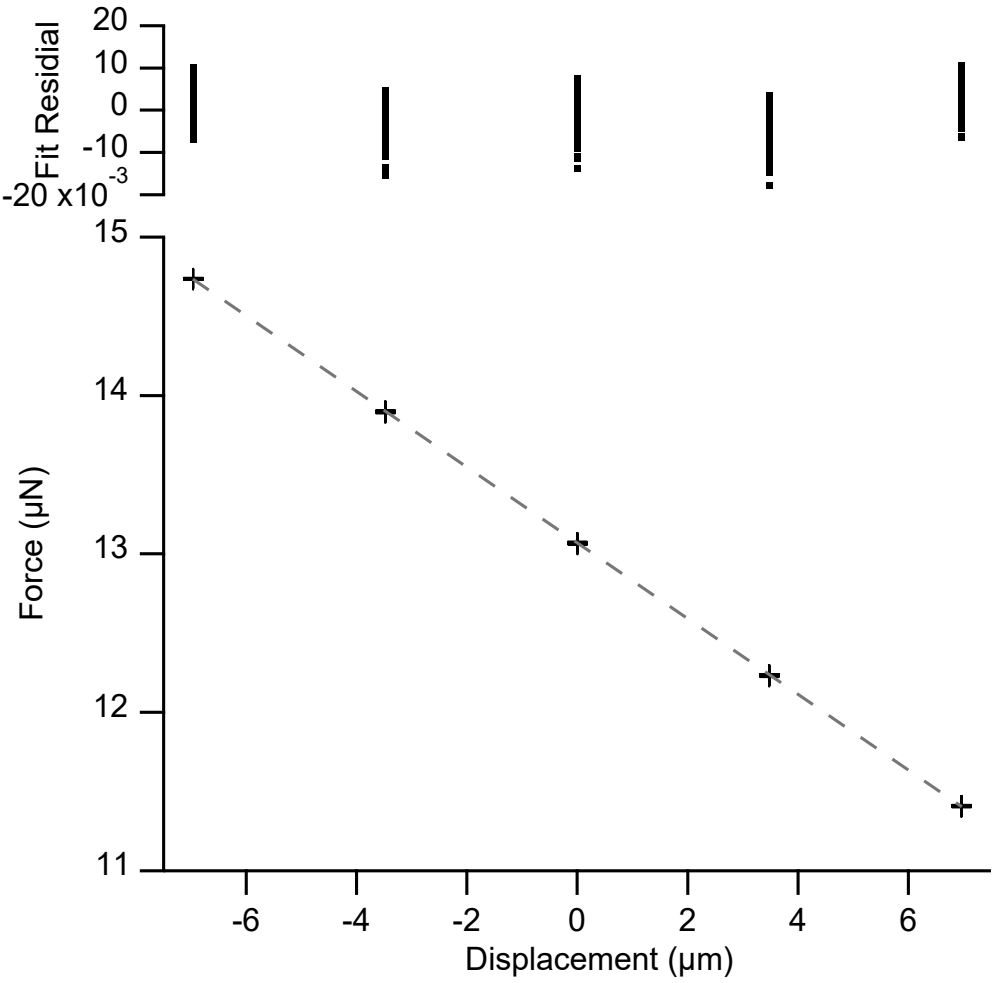

Fig. C 2 Force vs. displacement and linear fit residuals from the EFB test of the reference cantilever array W 23 D, 300 um long cantilever. Note, the slope of the fitted line (and determined stiffness) is negative because of the sign convention chosen for the force and displacement.

The balance stiffness was determined repeatedly over the course of the measurements by cycling the balance position as above while out of contact with the cantilevers. The EFB uses a buckling spring to reduce the stiffness of the balance mechanism. This was adjusted to give the lowest possible stiffness over the displacement ranges used prior to beginning the measurements, and changed twice during the measurement campaign. After the tension spring was changed, it was observed that the balance stiffness would change as shown in Fig. C 3. A series of exponential fits to the timeseries of the measured stiffness were used to determine k_b at the time of testing for each reference cantilever.

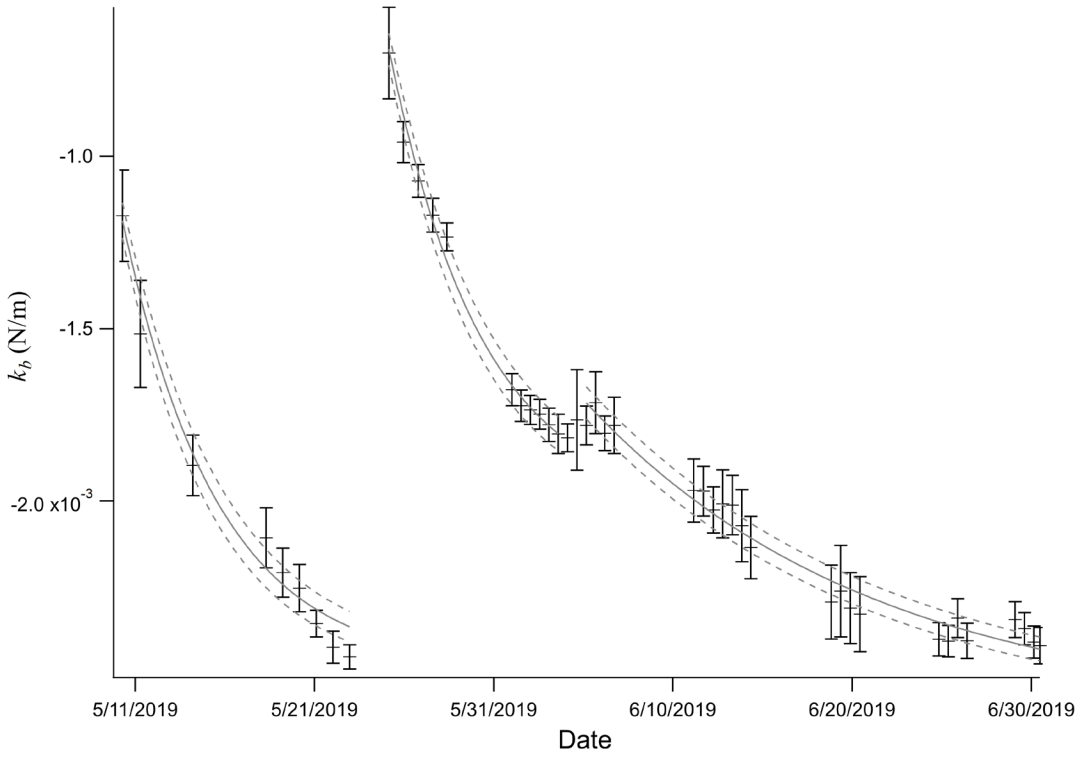

Fig. C 3 Time series data of balance stiffness. Markers and error bars show individual trials with the standard deviation of multiple measurements conducted during the daily trial. Solid lines show exponential fits to kb, and dashed lines show the fit uncertainty.

The uncertainty in the EFB force and displacement measurements has been described in detail elsewhere [10]. This shows the type B uncertainties in the EFB force and displacement measurement to be negligible for these measurements. In the current spring constant determinations, the uncertainty is dominated by the type A uncertainty in the measurements of k_m, k_b and type B uncertainty in OEC. The uncertainty in k_m is determined as the standard deviation of multiple daily measurements. Uncertainty in k_b is determined from the uncertainty in the exponential fit parameters used to track the balance stiffness over time. The fit parameter uncertainty is determined from the covariance matrix of the least squares fitting algorithm [12]. The expected uncertainty in k_b is calculated from the fit parameter uncertainty at the times the values of k_m are determined. The uncertainty in the OEC is dominated by the type B position uncertainty from the specified repeatability of the position sensors in the x-y stage used to locate the cantilever for testing, as described above, and is the largest uncertainty term. A smaller type A uncertainty from multiple measurements of the contact point location procedure is also included. The uncertainty reported in Table 1 within the LDV-EFB comparison section is the quadratic sum of these four uncertainties.

Appendix D. REFERENCE CANTILEVER CALIBRATION METHOD

A more detailed procedure for utilizing the reference cantilever method is provided here, starting with advice on remounting the SRM 3461 device on a more secure stage (e.g., a steel puck for inserting onto magnetic stages often used in AFM).

Reference Device Remounting Option

Since the reference cantilever calibration method may apply sufficient forces at the end of a cantilever to rock the device chip on the Gelpak, it is suggested to re-mount the device more securely to an appropriate surface (e.g., steel puck) that could be used in the AFM for this method.

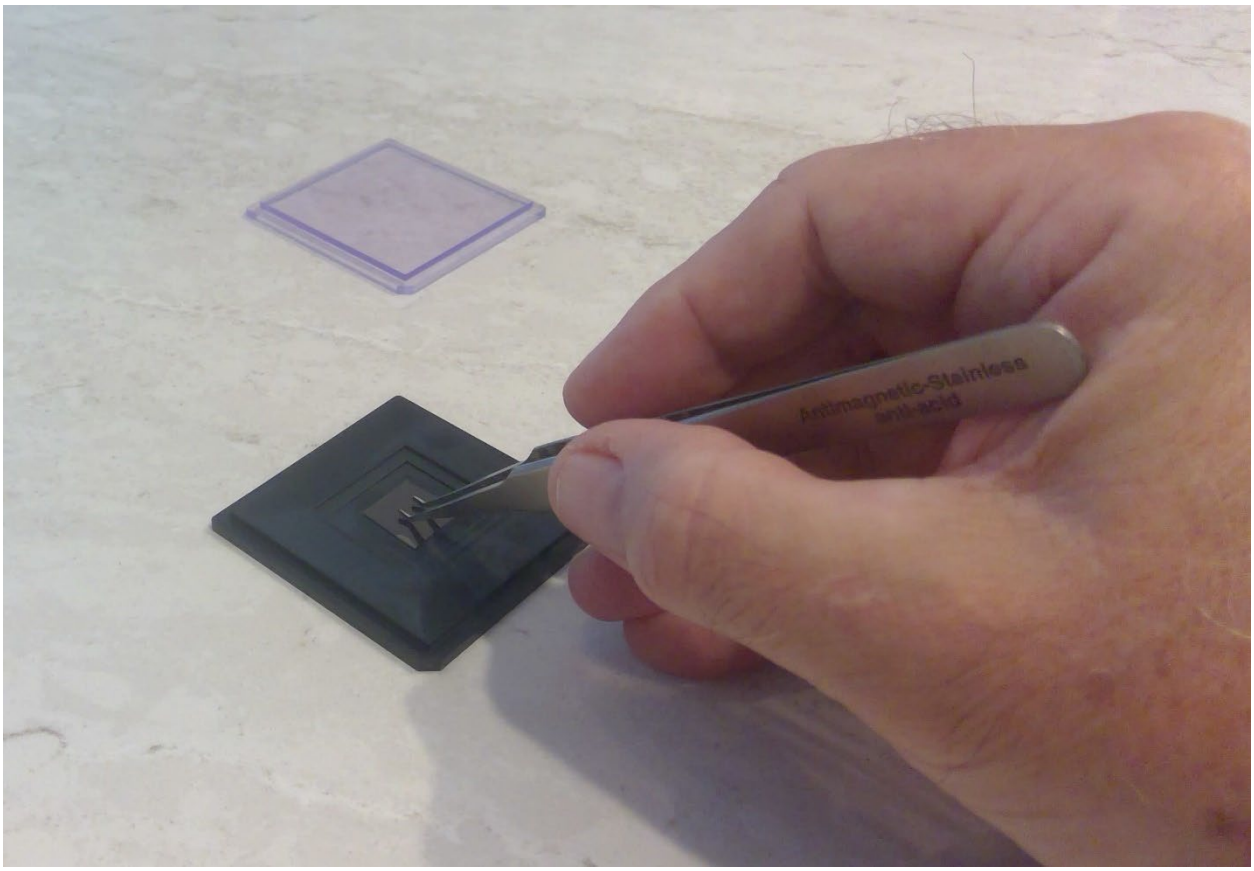

Fig. D 1 SRM 3461 device holder, with cover removed. Handling should occur only on the sides of the chip with the reference cantilevers pointing away from you (toward upper left in this picture). The heel of your palm should be braced against a firm, flat surface for stability.

Clean the puck and tweezers using a laboratory wipe moistened with a few drops of acetone and allow to dry thoroughly

Place the puck on a flat surface alongside the SRM 3461 holder/cover (already removed from its clip) under an appropriate magnifier with sufficient illumination. Apply a small drop of adhesive (e.g., cyanoacrylate) at the center of the puck where you want the SRM 3461 device to sit. Take the cover off the holder exposing the device and orient the cantilevers toward the top of your view, rotated slightly to accommodate the proper orientation of your grip holding a tweezer (either right or left-handed as appropriate). Holding a flat bladed tweezer, rest the lower part of your hand near the holder on the flat surface puck to steady your hand and carefully straddle the sides of the chip with the cantilevers away from you. Slowly close the tweezer on the sides of the chip until it lightly grips it. With your other hand, carefully rotate the Gelpak holder around the chip until it "releases" the chip from the gel (this may take more than a 20 degree rotation. Once you feel it disengage, maintain your slight pressure on the chip and using the base of your palm as a fulcrum, raise up the chip several mm. Slide the holder out from under your hand and substitute the steel puck. Rock your hand forward to lower the chip onto the puck off center slightly so that the back of the chip is nearest the droplet of glue. Slide the puck forward toward the chip until the droplet contacts the chip and capillary action pulls the glue under the chip. Release the chip by spreading the tweezer, trying not to get any glue on it. A gentle tap on the top of the chip using a corner of the closed tweezer will ensure that the chip is lying flat on the puck. Let the glue dry and secure it in a magnetic case for safe keeping.

One word of caution, Gelpak films are made of polydimethylsiloxane and residues from these films can act as release agents and prevent/reduce bonding between AFM chips and substrates to which they are glued. Care should be used when storing, handling, and even shipping these bonded materials since thermal expansion mismatch between the chip and substrate may initiate a catastrophic debonding of the chip from the substrate. This is especially true for shipping where transportation of items in airline cargo holds can undergo significant temperature fluctuations during flights.

Reference Cantilever Calibration Method

The reference cantilever calibration method is a direct way to calibrate the stiffness of an AFM test cantilever using a reference cantilever of known stiffness. The pressing of one (unknown stiffness) cantilever against another (known stiffness) is a (multiple) springs in series problem and allows determination of the unknown spring by measuring the deflections of each.

Practically, the method requires aligning a test cantilever above and near the end of the reference cantilever and performing an approach-retract curve. The point of contact should be along the long central axis of the reference cantilever and a known distance from the end (typically a few μm).

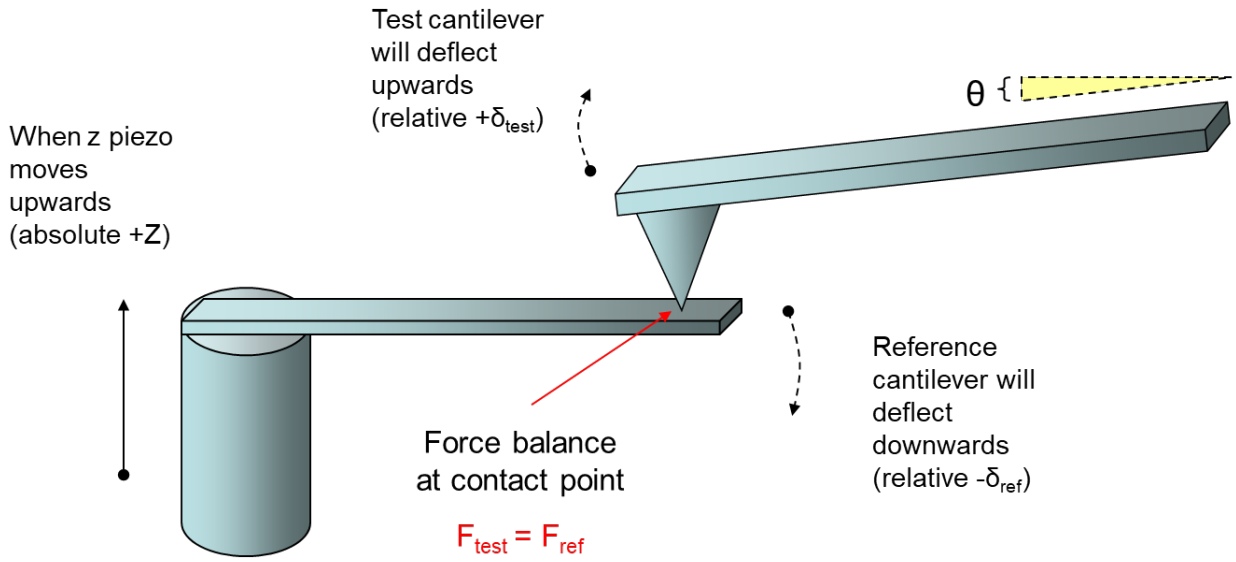

Fig. D 2 Schematic for contact between test and reference cantilevers.

A typical approach-retract curve for an AFM conducted on a surface has the appearance shown below. The x axis represents the piezo displacement in nm (calibrated independently using step height standards). This is the same as the z axis displacement of the z piezo shown in Fig. D 2. The y axis represents the cantilever deflection in volts which is actually interpreted from the change in tilt of the test cantilever as a laser spot is reflected off the back of the cantilever and onto the center of a quadrant optical detector. After a period of approach without contact (a), the cantilever tip "snaps in" and contacts the surface (b). With further approach, a linear region is observed where further test cantilever displacement (actually tilt) is proportional to the piezo displacement and is called the compliance region (c). After the piezo displacement setpoint is reached, the cycle reverses, becoming retraction. At some point the tip "snaps off" the surface and the depth of the snap indicates the degree of adhesion force between the tip and the surface. The non-contact retract portion is horizontal and overlaps the similar approach region.

The approach-retract curve depicted is for a stiff surface. When a similar curve is attempted against a reference cantilever, a similar curve is obtained but the slope of the compliance portion will lie in between the stiff version and a horizontal line (zero slope).

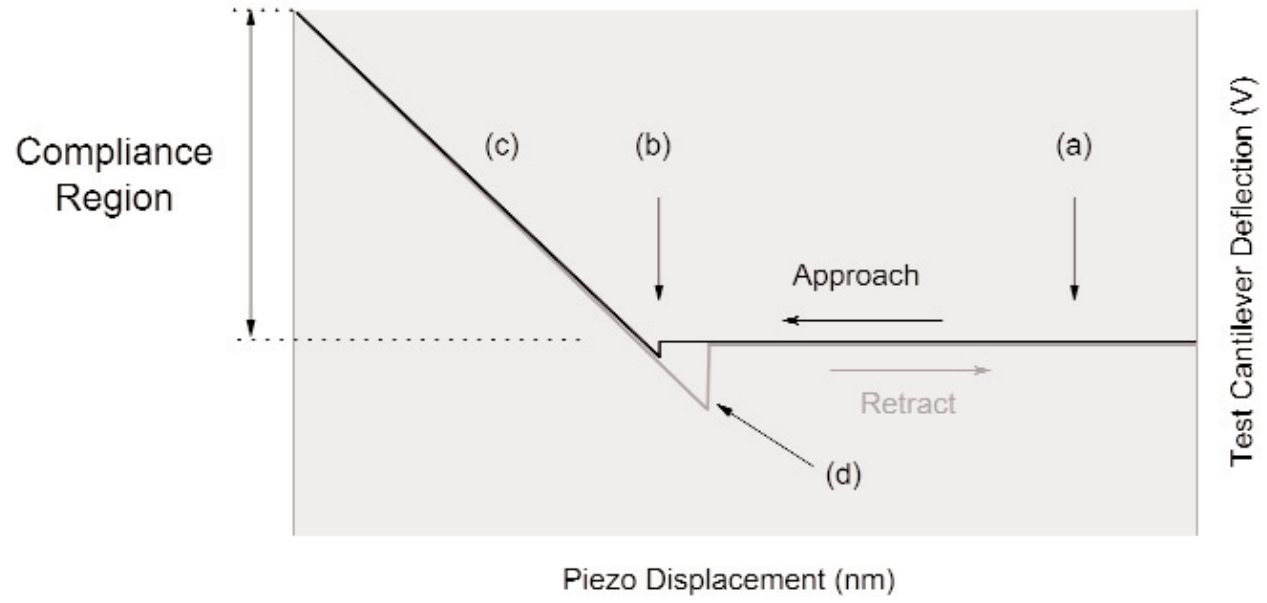

Fig. D 3 Example of AFM approach-retract curves.

The slopes (S) of both compliance curves (stiff surface and reference cantilever) are then used along with the equation below to estimate the intrinsic stiffness of the test cantilever.

$$k_{test} = k_{ref}\left(\frac{L}{L-d_{tip}}\right)^3\left(\frac{S_{stiff}}{S_{ref}} - 1\right)cos^2\,\theta \qquad \text{(D1)}$$

Where k_{test} is the intrinsic stiffness of the test AFM cantilever, k_{ref} is the stiffness of the reference cantilever, L is the length of the reference cantilever, and d_{tip} is the distance from the end of the reference cantilever to the contact point of the test cantilever. This assumes that the calibrated stiffness of the reference cantilever is provided for the end of the cantilever.

One point of order about cantilever stiffness is that there are two kinds of stiffness encountered in AFM: intrinsic and effective. The intrinsic stiffness (k_i) is the stiffness perpendicular to the long axis of the cantilever and represents the stiffness that is reported for commercial cantilevers when they are purchased. When used in an AFM the surface is actually exposed to the effective stiffness (k_e) of the cantilever which is larger than the intrinsic stiffness by a factor of $1/cos^2\,\Theta$ where theta represents the angle of repose of the cantilever toward the surface (approximately 11° but varies with AFM instruments). The relationship is:

$$k_e = \frac{k_i}{cos^2\,\theta} \qquad \text{(D2)}$$

Which for an 11° angle of repose represents about a 4% increase in stiffness.

It has been observed, particularly for low stiffness cantilevers with large tip heights, that hysteresis in the loading and unloading curves may occur [13]. In this case, it has been shown that the mean of the loading and unloading curve slope can be used for S_{stiff}.

Appendix E. CHANGE LOG

Corrections made in an errata update do not alter existing or introduce substantive technical information, but rather are intended to remove ambiguity and improve interpretation of the work.

Change 1, 10/16/2023, equation C3 on page 23.

There was a typographical formatting error in appendix E of the publication, equation C3 was:

$$k_c = [(k_m^{-1} - k_b^{-1})^{-1} - k_l^{-1}]OEC$$

The expression should be:

$$k_c = [(k_m - k_b)^{-1} - k_l^{-1}]^{-1}OEC$$

This is solely an error in the typing of the equation, all calculations in the paper were performed using the correct equation, and no other modifications are required.