

The NIST Johnson Noise Thermometry System for the Determination of the Boltzmann Constant

Nathan E. Flowers-Jacobs¹, Alessio Pollarolo¹, Kevin J. Coakley¹, Adam C. Weis¹, Anna E. Fox¹, Horst Rogalla¹, Weston L. Tew², and Samuel P. Benz¹

¹National Institute of Standards and Technology,
Boulder, CO 80305, USA

²National Institute of Standards and Technology,
Gaithersburg, MD 20899, USA

nathan.flowers-jacobs@nist.gov
alessio.pollarolo@nist.gov
kevin.coakley@nist.gov
adam.weis@nist.gov
anna.fox@nist.gov
horst.rogalla@nist.gov
weston.tew@nist.gov
samuel.benz@nist.gov

In preparation for the redefinition of the International System of Units (SI), five different electronic measurements of the Boltzmann constant have been performed using different Johnson noise thermometry (JNT) systems over the past seven years. In this paper, we describe in detail the JNT system and uncertainty components associated with the most recent National Institute of Standards and Technology (NIST) determination of the Boltzmann constant: $k = 1.380642\ 9(69) \times 10^{-23}$ J/K, with a relative standard uncertainty of 5.0×10^{-6} and relative offset of -4.05×10^{-6} from the Committee on Data for Science and Technology (CODATA) 2014 recommended value. We discuss the input circuits and the approach we used to match the frequency response of two noise sources. We present new measurements of the correlated noise of the 4 K on-chip resistors in the quantum-accurate, pseudorandom, voltage-noise source, which we used to estimate the correlated, frequency-dependent, nonthermal noise in our system. Finally, we contrast our system with those used in other measurements and speculate on future improvements.

Key words: Boltzmann equation; correlation; digital-analog conversion; Josephson arrays; measurement units; noise measurement; quantization; signal synthesis; standards; temperature.

Accepted: October 2017

Published: December 29, 2017

<https://doi.org/10.6028/jres.122.046>

1. Introduction

For the forthcoming redefinition of the kelvin in the International System of Units (SI), an exact numerical value for the Boltzmann constant k in units of joules per kelvin will be selected to relate the thermodynamic temperature T to energy E using the relation $E = kT$. The fixed value of the Boltzmann constant will be established from low-uncertainty measurements of k relative to the current SI definition of the kelvin based on the intrinsic definition of the triple point of water. The Consultative Committee for Thermometry (CCT) of the International Committee for Weights and Measures (CIPM) required that the value of k should be fixed only if the Boltzmann constant has been determined by at least one measurement technique with a relative uncertainty less than 1×10^{-6} , and at least one additional measurement technique with uncertainty less than 3×10^{-6} [1]. The most precise measurements of the Boltzmann constant k are

based on the acoustic properties of gases with less than 1×10^{-6} relative uncertainty [2, 3]. An alternative technique using dielectric constant gas thermometry [4, 5] recently produced a value for k with a relative uncertainty of 1.9×10^{-6} [6]. The combination of these two techniques meets the CCT requirement for redefinition.

In contrast to techniques based on atomic gases, Johnson noise thermometry (JNT) is based on the properties of a free electron gas in a conductor [7, 8, 9]. As predicted by the fluctuation-dissipation theorem [10, 11], the thermal motion of electrons results in voltage and current noise as described by Nyquist's equation as mean-square noise voltage $\langle V_R^2 \rangle$, which is proportional to resistance R for temperature T and bandwidth Δf :

$$\langle V_R^2 \rangle = 4kTR\Delta f. \quad (1)$$

This approximation to the Nyquist equation is accurate to better than 5×10^{-9} at frequencies below 100 MHz and $T > 25$ K, where quantum corrections are small. In this regime, the Johnson noise is “white”; that is, the power spectral density is independent of frequency for fixed temperature, resistance, and bandwidth. The generality of the link between voltage fluctuations and resistance allows Johnson noise to be used as a primary thermometer.

Johnson noise thermometry has an extensive history [12–15], but precision applications were made possible only more recently by NIST's development of a superconducting, quantum-accurate, pseudorandom, voltage-noise source (QVNS) as an exact synthetic noise reference. The QVNS generates an accurate synthetic pseudorandom noise signal with mean-square voltage $\langle V_Q^2 \rangle = |v_{JJ}|^2 / K_J^2$, where JJ indicates the Josephson junction, $|v_{JJ}|$ is a noise power spectral density that is calculable from QVNS implementation details, and K_J is the conventional value for the Josephson constant. When the resistance of the thermal noise source is expressed in terms of the conventional value of the von Klitzing constant R_{K-90} , the measured ratio $\langle V_R^2 \rangle / \langle V_{JJ}^2 \rangle$ will be proportional to k/h , the ratio of the Boltzmann constant to the Planck constant. The relative uncertainty in the value of the Planck constant, 1.2×10^{-8} , is less than 2 % of the smallest relative uncertainties in the Boltzmann constant, 5.7×10^{-7} [16]; therefore, we can assume a fixed Planck constant of $h = 6.626070040 \times 10^{-34}$ J s with negligible effect on the uncertainty of our measured value of k .

The small magnitude of Johnson noise, only 0.123 nV $\Omega^{-1/2}$ Hz $^{-1/2}$ at 273.16 K, poses a challenge for accurate measurement. However, JNT has progressively improved through the implementation of techniques such as switched cross-correlators [13], accurate QVNS references [17], and others [18–33]. The National Institute of Standards and Technology (NIST) has been developing precision QVNS-JNT systems since 1999 in cooperation with the Measurement Standards Laboratory of New Zealand [18–30], and since 2009 in collaboration with the National Institute of Metrology (NIM) of China [31–33]. Different iterations of QVNS-JNT systems at NIST and NIM have produced independent measurements of k in 2011–2017 [34–36], with the most recent NIM measurement determining k to a relative uncertainty of 2.7×10^{-6} , meeting CCT's 3.0×10^{-6} threshold [1]. The most recent NIST determination of k was reported with a relative uncertainty of 5.0×10^{-6} [37]. Finally, the National Metrology Institute of Japan (NMIJ) developed a similar system in which the room-temperature pulse generator used in other QVNSs is replaced by an on-chip, superconducting, integrated pseudorandom number generator [38]. NMIJ determined a value for k with a relative uncertainty of 1×10^{-5} [39].

This article discusses in detail the implementation of the NIST JNT system, and it is intended as a companion to the recent NIST report [37] focused on determining a value for k . Necessarily, this more detailed document will revisit some of the same material, but it will also provide deeper information and discussion on the measurement circuits and analysis techniques used in Ref. [37]. Emphasis will be placed on modifications incorporated since our earlier 2015 determination of k [34]. Major differences include an increase in the thermal noise resistance from 100 Ω to 200 Ω , attempts to match the frequency response of the measurement channels using trim resistors and capacitors, separate shielding of the twisted pairs of wires to reduce interchannel cross-talk, and selection of noise ratio spectra models with an automated cross-validation method developed by Coakley et al. [40].

The instrumentation used in Ref. [37] will be discussed in Sec. 2, covering the two noise sources (Sec. 2.1–2), measurement electronics (Sec. 2.3–4), and shielding (Sec. 2.5). Section 3 describes the use of the

complex ratio of the measured noise sources in the context of a measurement of the undesired correlated noise in the JNT system, such as possible current noise from the amplifiers that could cause frequency-dependence in the measured Johnson noise signal. The details of the measurement of the Boltzmann constant in Ref. [37] are discussed in Sec. 4, including the matching of measurement channels with trimming capacitors (Sec. 4.1). Section 5 similarly explores the uncertainty budget in greater detail than in Ref. [37] and includes some discussion of the ways in which the statistical uncertainty may be estimated from experimental parameters (Sec. 5.2). Section 6 compares measurement hardware among several JNT experiments at NIST and other laboratories.

2. QVNS-JNT Apparatus at NIST

In this section, we describe the components of the JNT system at NIST (Fig. 1) used to make the Boltzmann constant determination in Ref. [37]. Two voltage sources are measured alternately and compared: (1) the sense resistor is thermally coupled to a realization of the triple point of water (TPW), and generates the thermal mean-square noise voltage $\langle V_R^2 \rangle$, described in Sec. 2.1, and (2) the QVNS uses Josephson junctions to generate synthetic pseudorandom noise voltage with a known magnitude $\langle V_Q^2 \rangle$, described in Sec. 2.2 [29, 41, 42].

To accurately determine the ratio $\langle V_R^2 \rangle / \langle V_Q^2 \rangle \propto kT$, the two voltages V_R and V_Q must be measured using nearly identical data acquisition hardware. In Sec. 2.3, we describe how this is accomplished by measuring both voltages with the same chain of amplifiers and the same digitizers by switching between voltage sources every 100 s using mechanical relays on a custom circuit board.

Since the magnitude of the Johnson noise is similar to that of typical amplifiers, two separate channels of amplifiers, low-pass anti-aliasing filters, and digitizers are used (discussed in Sec. 2.4), and a cross-correlation measurement is performed to reduce the effect of amplifier noise [43, 44]. Any cross-talk between the channels or electromagnetic interference (EMI) picked up by both channels is indistinguishable in the cross-correlated data from Johnson noise. Therefore, we have made a large effort both to reduce cross-talk between the channels, as discussed throughout Sec. 2, and to implement careful shielding and grounding, discussed in Sec. 2.5. Earlier JNT measurements at NIST were limited by EMI; the EMI in Ref. [37] was drastically reduced by operating the system in a shielded room.

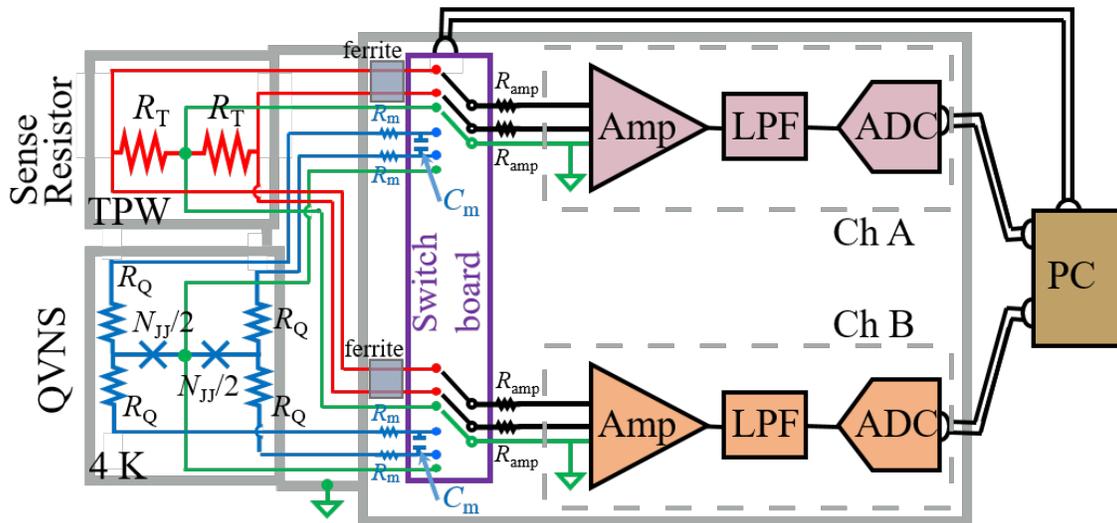


Fig. 1. Johnson noise thermometer schematic diagram. The 200 Ω sense resistor is composed of two series resistors $R_T = 100 \Omega$. The cryogenic QVNS is composed of two pairs of resistors $R_Q \approx 100 \Omega$ and two arrays with $N_{JJ} = 20$ Josephson junctions, 10 per array. The QVNS lines have additional trim resistors $R_m \approx 1 \Omega$ and trim capacitors $C_m = 5 \text{ pF}$ to 60 pF used to match the response of the measurement channels. A custom switch board is used to determine which noise source is connected to the two signal-processing channels (Ch A and Ch B). The amplifiers (Amp) are stabilized using a combination of common-mode ferrite chokes (light gray boxes) and resistors R_{amp} . The center of each source is grounded (green lines) at the amplifier input, and that ground is used to shield

each differential twisted pair. After the custom high-gain low-noise amplifier, there is a steep 11th-order 850 kHz low-pass filter (LPF) and an ac-coupled, 16 bit, 2.083×10^6 samples/s digitizer (ADC). There is an additional grounded shield surrounding the entire system (gray boxes), and the entire system is also in an electromagnetically shielded room. The switchboard and ADCs are optically controlled by a personal computer (PC) outside of the shielded room.

2.1 Thermal Noise Source

In the recent NIST measurement [37], the Johnson-noise-generating sense resistance coupled to the TPW (top left of Fig. 1) was chosen to be 200Ω , which increased the signal-to-noise ratio relative to earlier 100Ω JNT experiments [34]. This increase in resistance did not significantly change the measurement bandwidth of our system because the effective RC cutoff frequency associated with the measurement circuit is larger than the sampling frequency.

The sense resistor is formed by a series of two closely matched 100Ω custom-made nickel-chrome-alloy foil resistors bonded onto alumina chip substrates. Nickel-chrome foil resistors are commonly used in resistance metrology and are chosen here for their stability and low temperature coefficient of resistance. The foil resistors are approximately $25 \mu\text{m}$ thick serpentine patterns within a 1.2 mm square boundary and are part of a set of four such resistors mounted inside a hermetically sealed leadless chip carrier (LCC) that is approximately $8.9 \text{ mm} \times 8.9 \text{ mm}$ in size. The temperature coefficients of resistance (TCR) of the resistors are within the limits of $\pm 5 \mu\Omega/\Omega \text{ K}$ and are laser-trimmed to a relative match of 0.01% and an absolute tolerance of 0.1% of nominal. The LCC package (Fig. 2) has 10 external connection points to internal gold wire interconnections that provide the two-terminal-pair definition points at internal junctions within on-chip wire-bonding pads. A common connection point is located along the centerline of the package and provides the circuit ground point between all the resistors (see Fig. 2, showing two of the four resistors with five connection points). This “two-resistor five-point” configuration provides a symmetric thermal noise source for the two amplifier channels [27] in the absolute-measurement mode. The LCC package also allows for a separate connection configuration using all four of the foil resistors connected together as a four-wire short in the null-measurement mode for evaluation of EMI [45] (see Sec. 5.3).

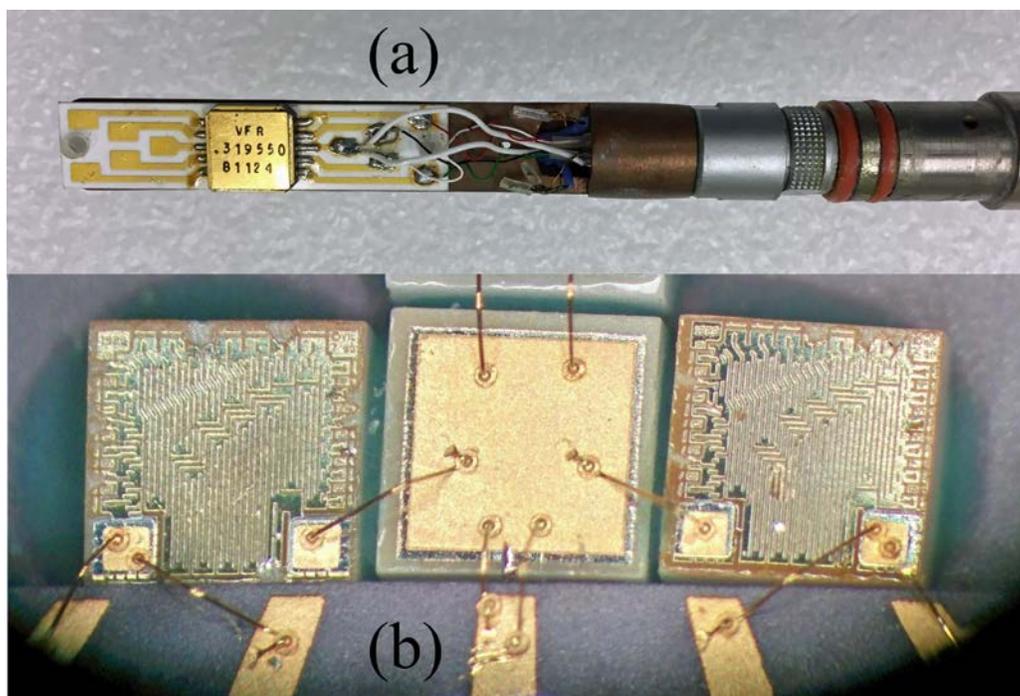


Fig. 2. Photographs of the resistor package used as Johnson noise source: (a) mounted in an 8.9 mm LCC at the end of the R-probe, and (b) inside the LCC, two 100Ω serpentine foil resistors (200Ω total) and a central grounding pad. Each resistor structure is contained in a $1.2 \text{ mm} \times 1.2 \text{ mm}$ square region.

The LCC is soldered onto an alumina printed circuit board (PCB) with gold-printed solder pads to provide the connections to Teflon-insulated twisted-pair leads.¹ This alumina PCB is fastened onto the flat surface of an internal copper heat sink (a 64 mm long, 10 mm diameter split cylinder) inside of the resistance probe (R-probe). The R-probe is made from a 400 mm long, thin-wall, 6.35 mm diameter stainless-steel sheath that contains a pair of 1.4 mm diameter stainless-steel shielding tubes. These internal shielding tubes are isolated from the external sheath and are connected to the resistor's circuit ground point at the bottom of the probe, near the LCC package. The twisted-pair lead wiring is contained within the internal shielding tubes to reduce differential cross-talk to less than -80 dB by decoupling the signal path along the length of the R-probe and keeping the wiring mechanically stable. Earlier versions of this system used unshielded twisted pairs, while the recent NIM system used a pair of coaxial lines per channel. The bottommost 10 cm section of the R-probe is formed from a 12 mm diameter removable copper sheath that allows access to the internal copper heat sink, the alumina PCB, and wiring connections. The copper sheath is mechanically fastened to the R-probe body and sealed with a pair of silicone O-rings. To keep the internal volume of the probe free of excess humidity, the R-probe is internally pressurized with dry nitrogen gas to approximately 7 kPa above ambient pressure.

The R-probe used here, designated "JT4B," is of a similar design to those used at NIST in the past [26] and was rebuilt in 2014 to provide better immersion characteristics. The probe is immersed in the thermowell of a "B-type" borosilicate-glass TPW cell. While other approaches to stabilizing the sense resistor temperature would have been possible, such as an electronically temperature controlled environment, we chose to use triple-point cells for their convenience, since the sense resistor was easily conformed to an immersion probe configuration. The TPW cell's thermowell is 13 mm in diameter and 270 mm in total immersion depth and is filled with ethanol to serve as a heat-transfer fluid. The TPW cell is maintained inside a thermoelectrically cooled dry-well maintenance system, which was specifically designed to maintain ice mantles for this type of cell. We refer to this TPW cell, serial number 1302, as the "working cell," which has always remained within NIST, Boulder, Colorado. The working cell has been directly compared with other TPW transfer cells that have been transported between NIST's Boulder, Colorado, and Gaithersburg, Maryland, campuses. We describe these comparisons as well as the R-probe immersion tests in Sec. 4.4 and Sec. 5.8.

The 200 Ω sense resistor is calibrated every 3 to 7 d versus a Tegam SR102 100 Ω standard resistor, identification number 1202T, which also serves as a portable transfer standard between NIST's Boulder and Gaithersburg campuses and provides traceability to the NIST ohm, maintained with the quantum Hall effect (QHE) [46]. The standard resistor has a history of QHE calibrations dating back to 2007 and exhibited an average drift rate of $+0.068 \mu\Omega/\Omega$ per year between 2007 and 2011. More recent measurements made in 2015 indicate a lower drift rate of $+0.020(7) \mu\Omega/\Omega$ per year. The sense resistor is much less stable than this and requires relatively frequent recalibration. When maintained at a constant temperature inside of the TPW cell, the sense resistor will usually drift on the order of $-0.1 \mu\Omega/\Omega$ per day. Transient drift can occur at higher levels when the resistor is thermally cycled or otherwise perturbed; we avoided making noise measurements under these conditions (see discussion in Sec. 4.5).

The sense resistor calibrations are performed with either of two resistance bridges. The first resistance bridge is an automatic direct current comparator (DCC) with a specified uncertainty of $0.1 \mu\Omega/\Omega$ [47]. The second bridge is an automatic ac digital substitution bridge (DSB), also with a specified uncertainty of $0.1 \mu\Omega/\Omega$ [48]. Both bridges are used with the 1202T 100 Ω standard as the reference and hence cannot be operated simultaneously. The DCC bridge is operated at reversal time of 60 s, whereas the DSB operates with a 6 Hz square wave excitation. In both cases, the sense resistor is calibrated at an excitation current of 0.25 mA, which is a compromise between sensitivity of the measurement and self-heating effects. We describe the resistance measurements in greater detail in Sec. 4.5.

¹ Certain commercial materials, devices, and instruments are identified to specify the experimental study adequately. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the materials, devices, and instruments are necessarily the best available for the purpose.

2.2 Quantized Voltage Noise Source

The QVNS superconducting-integrated circuit was fabricated at NIST (see Fig. 3) and consists of a total of 20 Josephson junctions. The Josephson junctions use Nb electrodes with Nb-Si barriers and behave like superconductor-normal metal-superconductor (SNS) junctions [49, 50]. The critical current of the Josephson junctions is about $I_c = 10$ mA, and the normal resistance is about $R_n = 4$ m Ω , corresponding to a characteristic frequency of 20 GHz. As described in Sec. 5.7, precise QVNS output is possible because this characteristic frequency is much higher than the highest frequencies generated in the pseudorandom noise signal. The chip is indium-soldered to a copper block that both serves as the heat sink for the chip and fixes the location of the chip relative to a low-loss circuit board, which is bolted to the copper block [51]. The copper block is surrounded by a 3.8 cm diameter Cryoperm magnetic shield to protect the JJ arrays from stray magnetic fields. The entire package is immersed in liquid helium at the bottom of a 100 L storage Dewar.

The 20 JJs are divided into two arrays of 10 Josephson junctions each. Each array is separately pulse-biased via an AC-coupled coplanar waveguide (CPW) by current pulses from a fast pulse generator (inputs to the chip in the top-right and bottom-left of Fig. 3). We attempted to match the transfer function of the low-frequency voltage connection leads of the QVNS circuit (top-left and bottom-right of Fig. 3) with that of the connection leads and sense resistors of the Johnson noise source. Since the output impedance of the Johnson noise source at dc is equal to the 200 Ω resistance of the sense resistor, while the inherent output impedance of the QVNS JJ arrays at dc is 0 Ω , we achieve matching at dc by adding a 100 Ω on-chip resistor to each lead of the QVNS, for a total of 200 Ω per channel [52]. This necessarily adds additional Johnson noise to the measurement, but we use separate resistors on each pair of output leads so that the 4 K Johnson noise measured by each amplifier is uncorrelated and does not contribute to the cross-correlation.

The detailed configurations of the QVNS on-chip ground and output leads were also designed to match the configuration of the sense resistor probe (as shown in Fig. 1). To reduce cross-talk, the twisted-pair wiring in the QVNS probe is shielded in the same way as the wiring in the resistor probe. As described in Sec. 4.1, detailed tuning of the QVNS transfer function is accomplished by use of tuning capacitors and resistors on the relay board.

The QVNS uses the arrays of JJs to create a repeating sequence of voltage pulses. The time integral of each voltage pulse from each JJ is quantized and equal to $K_J^{-1} = h/2e$ per JJ [29]. These pulses are fast with a width < 50 ps; therefore, the low-frequency, < 10 MHz components of the waveform are precisely calculable. The sequence of pulses in the repeating pattern is determined using a delta-sigma modulator algorithm, and is a repeating pseudorandom noise-like time-dependent voltage waveform that, in the frequency domain, is a frequency comb of odd harmonics $f_0, 3f_0, 5f_0, \dots$, with equal amplitude and random phase. In Ref. [37], the frequency comb extended up to 4 MHz and used $f_0 = 79$ Hz. The amplitude of the comb is designed so that the average power spectral density of the comb, over the < 2 MHz measurement system bandwidth, matches the 1.74 nV/ $\sqrt{\text{Hz}}$ power spectral density of the thermal noise of the 200 Ω sense resistor at the TPW.

However, the delta-sigma modulator algorithm is not guaranteed to produce the designed waveform with precision to 10^{-6} . We therefore directly calculate the spectrum of the QVNS using the pulse sequence $p[m]$, where at each time step m , there is either a positive pulse $p = 1$, a negative pulse $p = -1$, or no pulse $p = 0$. We take the fast Fourier transform (FFT) of the actual pulse sequence to determine the magnitude of each tooth of the frequency comb $|v_{JJ}(f)| = N_{JJ} f_S \mathcal{F}[p[m]]$, where N_{JJ} is the number of Josephson junctions, and f_S is the number of steps in the pulse sequence per second. We choose these units, where $v_{JJ} K_J^{-1}$ has units of volts, to simplify the relationship between the JNT measurement and fundamental constants.

We use a Sympuls pulse generator to apply fast current bias pulses to each JJ array. Because current passing through the JJs at the comb frequencies can create an inductive voltage error, we pass the bias pulses through three analog high-pass filters and also employ a “zero-compensation” scheme to further reduce the current at frequencies < 10 MHz [33]. The component chain from the Sympuls generator to the chip starts with a K251 bias tee from Anritsu, followed by a 1 dB attenuator, a DCB-3510-MF-SMA-02 inner-only dc block, a 2 dB attenuator, a DCB-3511-MF-SMA-02 inner/outer dc block, a 1 dB attenuator, a feedthrough into the shielded room, another DCB-3511-MF-SMA-02 inner/outer dc block, a 1 dB

attenuator, and a feedthrough into the QVNS probe, which is inserted into a 100 L liquid-helium Cryomag CMSH-100 Dewar. Inside the QVNS probe, we use CobraFlex cables to reach the chip package, which is immersed in liquid helium. An SMA adapter connects the cable to a coplanar transmission line on the circuit board. The transmission line is wire-bonded to the chip, followed by an on-chip inside-outside block, the JJ array, and finally the large on-chip ground plane.

This is the first time we have employed on-chip dc blocks on the QVNS chip. They serve two purposes. First, the combination of off-chip attenuators and dc blocks forms three single-pole high-pass filters, each with a 6 dB point of around 5 MHz (assuming a 50 Ω environment). The on-chip block contributes another single-pole high-pass filter with a 6 dB point near 20 MHz (again assuming a 50 Ω environment). This filtering reduces the inductive error from pulse-bias signals at 2 MHz by 50 dB. Second, the inner/outer dc blocks separate the pulse generator ground from the shielded room and from the system ground, avoiding ground loops. The grounding configuration will be discussed in more detail in Sec. 2.5.

To further reduce the inductive error, we also use a zero-compensation pulse scheme [33] in which each current-bias pulse from the pulse generator is surrounded by two half-magnitude pulses of the opposite polarity. Because of the nonlinearity of the JJs, and with appropriate amplitudes, this composite bias will still induce each JJ to create a single quantized output voltage pulse with the same sign as the central current pulse. Most importantly, the low-frequency current created by this composite bias is reduced because the average current of the composite bias is zero. In practice, we have measured the magnitude of the low-frequency bias by digitizing the output of the low-frequency port of the bias-tee, and we have observed a power reduction of > 30 dB. However, the zero-compensation waveforms created by the Sympuls pulse generator are unipolar; that is, they only create output pulses of a single polarity (it produces only two voltage levels). The QVNS output voltages for these patterns therefore have a calculable dc voltage offset of about 6 μ V.



Fig. 3. QVNS chip and cryogenic package (composite photo courtesy of Dan Schmidt, NIST). For scale, the silicon chip (blue) is 1 cm square.

2.3 Relays Connect Voltage Sources to Amplifiers

We use a set of latching relays on two identical custom circuit boards to switch between the two voltage sources (Fig. 1). The circuit board is designed to minimize the capacitive load, cross-talk, and leakage/loss through the circuit board by keeping the traces short and physically separated. An ATMEGA-8535L-8AU-1551A microprocessor on each circuit board controls the six dual-coil latching signal relays. Each microprocessor is powered by its own battery and is linked to the main controller (a Linux computer) via an optical cable. During noise measurements, the microprocessor is put into a power-down sleep mode, where all clocks in the microprocessor are stopped, which prevents any EMI close to the sensitive amplifier inputs. When switching sources, the noise measurement is paused, and the microprocessor is activated by an optical signal.

In the recent NIST measurement [37], a signal relay circuit board (represented by black relays in Fig. 1) was used to switch the voltage source signal every 100 s. A second, identical ground relay circuit board (represented by green relays in Fig. 1) was used to switch the ground connection; earlier measurements did not use this second circuit-board and instead tied all the ground wires together in a static manner [34]. In Ref. [37], we used two different configurations for the ground connection. In the earlier configuration, the ground relay circuit board completely separated the two voltage sources, so that the voltage source that was not being measured was completely floating. In the later configuration, the ground relay circuit board still sent the grounds of the source being measured to the amplifiers, but it also shorted the grounds of the other source to the shielding enclosure, which was connected to the system ground. In both configurations, we waited 5 s after switching for transients to settle before starting to collect noise data from the digitizers.

The QVNS input to the signal relay circuit board also includes trimming capacitors and resistors to match the QVNS-to-amplifier connections to the R-probe-to-amplifier connections. Each QVNS input has its own trim resistors in series with the input leads and both a ~ 100 pF fixed surface mount capacitor and a tunable capacitor (5 to 57 pF) in parallel. See Sec. 4.1 for a description of the procedure used in Ref. [37] to match the voltage sources.

Finally, in Ref. [37] we used common-mode ferrite chokes and resistors to stabilize the amplifiers. The ferrites are cylindrical rings with damping above ~ 1 MHz (Fair-Rite part number 2643021801). Common-mode coupling between the field-effect transistor (FET) differential amplifiers on each channel can cause Colpitts oscillations and amplifier saturation. This saturation does not occur when the amplifiers are connected to the QVNS for two reasons. First, the connection between the positive terminal of one amplifier and the positive terminal of the other amplifier has an impedance of 200Ω at dc (Fig. 1) and similarly for the connection between the negative terminals. This resistance is sufficient to damp the oscillations. Second, the JJ array acts as a low-inductance short, which further decouples the amplifiers.

However, when the amplifiers are connected to the Johnson noise source, there is a low-ohmic connection between the terminals of the two amplifiers (Fig. 1), and in that configuration, amplifier saturation occurs. We therefore use two different configurations of ferrite chokes and resistors to damp the Colpitts oscillation and avoid amplifier saturation. This change in configuration was made at the same time as the change in the ground relay circuit board discussed above.

In the earlier configuration, we added a common-mode ferrite choke around each twisted pair input from the Johnson noise voltage source immediately before the signal relay circuit board and added $R_{\text{amp}} = 10 \Omega$ between the circuit board and each terminal of the amplifier. In the later configuration, we removed the common-mode ferrite chokes and increased the resistance on each amplifier lead to $R_{\text{amp}} = 21.1 \Omega$. We experimentally determined that, without ferrite-core inductors, the amplifiers saturated with a resistance of 10Ω per lead, but not with 20Ω or 21.1Ω per lead.

This change in configuration simplified the electrical circuit and addressed concerns that the presence of ferrites could introduce nonlinearity that would increase at lower frequencies. Moreover, the differential-mode leakage inductance for the choke may also introduce frequency dependence. This change also slightly increases the effective amplifier noise.

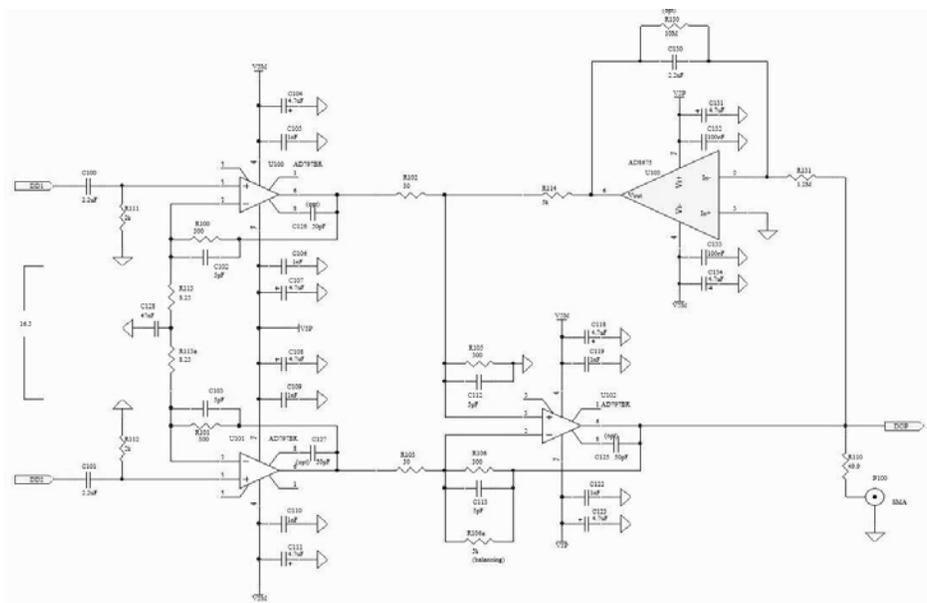
Other measurements used similar common-mode ferrite chokes between the amplifier and relay circuit boards [34, 53]. Some earlier experiments also used high-resistance nichrome wire in series with the sense resistor twisted-pair cable.

As the input stage, a differential junction field effect transistor (JFET) amplifier was designed with JFET constant-current source and cascaded outputs into $200\ \Omega$. The amplifier allows adjustment of the sum of the drain-source currents of the JFETs to about 10 mA. The JFET gates can be dc-coupled to the sense resistor and the QVNS. A JFET-pair, the SNJ903L42, specially selected and packaged by InterFET Corporation, was used and houses two matched IF9030 JFETs in one package. Each FET has an equivalent short circuit input noise of $0.5\ \text{nV Hz}^{-1/2}$ at 1 kHz, an input capacitance of 60 pF at 1 MHz, and a forward transconductance of 80 mS. The gate resistance is of the order of $10^{10}\ \Omega$. The bandwidth of the input stage with the input being dc-coupled is 10 MHz at $-0.5\ \text{dB}$ and 30 MHz at $-3\ \text{dB}$. For high-linearity operation, the peak-to-peak output voltage should not exceed 2 V.

Since no feedback to the gates of the JFETs is allowed, a cascode arrangement of the input stage is used that keeps the voltage at the JFET drain constant and thus prevents feedback via the drain-gate capacitance [54, 55]. Any feedback into the input circuit would result in current flow through the sense resistor. Without feedback, temperature affects the operating point of the JFETs, resulting in small changes to the gain. Since the power ratio between two noise-voltage sources is measured, long-period gain drift can be compensated by frequent switching between the two sources, such that their outputs are measured in similar amplifier conditions. However, sudden temperature changes need to be prevented, since short time drifts can be only partially compensated.

To avoid changes in the operating points of the JFETs caused by changes in the supply voltage, and to minimize addition of noise from the power supply, a voltage of $\pm 10\ \text{V}$ is supplied to the input stage from a precision and very-low-noise cascaded regulator, receiving its input from four lithium-ion batteries. This precision power supply is based on a circuit suggestion of Analog Devices for a single 10 V power supply, based on an AD587KR voltage reference. This circuit was modified and extended to a $\pm 10\ \text{V}$ power supply. It is capable of supplying more than 20 mA to the input stage at a noise voltage of about $1\ \text{nV Hz}^{-1/2}$. AD8675 operational amplifiers were used because of the possibility of rail-to-rail operation and their low noise and excellent stability.

The differential JFET input stage with a $200\ \Omega$ differential output resistance is ac-coupled to a low-noise, low-distortion, wide-bandwidth amplifier. The coupling capacitor is voltage overrated to reduce nonlinearities due to voltage-dependent changes in the dielectric constant. The amplifier design is similar to a classic instrumentation amplifier based on AD797 operational amplifiers of Analog Devices (see Fig. 5). Its 0.5 dB bandwidth is 1 MHz, and the 3 dB bandwidth is 2 MHz. For stability and gain flatness, some phase compensation is added. The gain of this amplifier is typically 51 dB.



To allow dc-coupling without dc offset voltages at the output of the instrumentation amplifier, an automatic offset compensation has been added. The dc component of the output is amplified in an AD8675 operational amplifier used as an integrator and fed back to the noninverting input of the last stage of the differential amplifier. To drive subsequent 50 Ω filters, a dc-coupled high-linearity buffer was added to the output of the amplifier, typically used with a gain of one into a 50 Ω load.

The full amplifier has a gain of about 70 dB, a common-mode rejection of up to 100 dB at 100 kHz, low differential input voltage noise of around 1.2 nV Hz^{-1/2}, high linearity, high input resistance, and low dc-output offset.

Earlier measurements used an amplifier design with an input ac-coupling via a 0.1 μ F ceramic capacitor and 20 M Ω resistance to ground. Compared to the earlier ac-coupled design, the amplifier described here allows for dc-coupling, thus preventing possible nonlinear effects due to coupling capacitors and dramatically increasing the input resistance.

The amplifier is followed by a low-pass filter, which determines the actual bandwidth of each channel. The filters are 11-pole passive Butterworth filters with a passband between dc and an upper corner frequency with a steep roll-off. Filters with different upper corner frequencies are available; for these measurements, a corner frequency (3 dB point) of 850 kHz is used, which significantly reduces aliasing in the subsequent analog-to-digital converter. However, for an accurate measurement, signals aliased from the second Nyquist zone are non-negligible, even with aggressive filtering. The contribution of aliased signals is characterized and minimized following a procedure described in Sec. 4.2.

For further data processing, the signals at the output of the two amplifier channels are digitized, transferred to a main computer, Fourier transformed, and then auto- or cross-correlated. The analog-digital converter (ADC) must convert the analog input with a sufficiently high sampling rate and resolution to keep up with the linearity and signal-to-noise ratio of the analog channels. The NIST-designed ADC is built around AD7625 by Analog Devices. This is an ADC with 16 bit resolution and a maximum sampling rate of 5×10^6 samples/s (5 MSPS), thus effectively converting the equivalent of 0.3×10^{12} 1 bit samples/s. In this application, a sampling rate of only 2 MSPS is used, since the filter-limited bandwidth is less than 1 MHz. The input of the analog-digital circuit board is single-ended 50 Ω and needs to be adapted to the differential input of the ADC via an operational amplifier, which also does some additional signal filtering. The signals for the start of the conversion and the conversion clock are generated in an external field-programmable gate array (FPGA) and transmitted to the ADC via optical cable. The internal oscillator in the FPGA is locked to a 10 MHz standard rubidium source. After the conversion, each 16 bit data point is read serially. For simplicity, the main computer supplies an additional optical readout clock that converts the serial 16 bit data back to 16 bit words, where we assume a constant delay in the data transfer between the ADC board and the main computer. The readout clock in the main computer can be shifted relative to the data stream so that the data content can be sensed in the middle of a transferred signal.

The 16 bit data words received by the main computer are stored in its memory via direct memory access (DMA). For each measurement, data are collected for 1 s. Each 1 s sequence of data, sampled at 2 MSPS, is Fourier transformed using a FFT algorithm and stored on a hard disk drive for further processing and for documentation. Groups of 100 consecutive transforms (100 s total data), known as "chops," are averaged from a single noise source and form the basis for further analysis.

The switchboards, the amplifiers, and the ADC boards are mounted in different compartments machined from a single block of aluminum for shielding and to supply a well-grounded environment. In addition, the whole block can be inserted in a high-permeability shield to suppress low-frequency EMI. This additional shielding was not used in the actual measurements in the shielded room.

The whole system in the electronics block is battery operated. Well-shielded battery boxes, each with two sets of lithium-ion batteries, are used to guarantee continuous operation of the electronics. While one battery set is charged from an ac-power supply or from an automobile battery, the second set is connected to the measurement electronics. The charging and switching are organized by a microprocessor in the battery box that is controlled by the main computer via an optical link. Like the microprocessor on the switchboard, this microprocessor remains in "sleep mode" during the measurement, and it is "woken up" for the switching and housekeeping operations.

2.5 Grounding and Shielding

Any signal within the analysis bandwidth of the system that appears in both channels or at the output because of the finite common-mode rejection of the amplifier chain is interpreted as correlated noise energy and thus influences the apparent temperature of the sense resistor. Shielding and elaborate grounding are very important to achieve the required accuracy for the determination of k . For a long time, only limited shielding was available for the project in the NIST laboratories in Boulder. Beginning in 2017, the system was placed in a magnetically and electrically shielded room (similar to Vacuumschmelze type AK 3a/b), which was originally designed and dedicated to biomagnetic measurements in the 10^{-15} T magnetic field regime. It shields static and low-frequency magnetic fields by use of layers of high-permeability Mu-metal with a relative permeability of up to 2.5×10^5 . Higher frequencies are shielded with thick Al layers and are attenuated by 100 dB or better at frequencies of 1 MHz and above (Fig. 6).

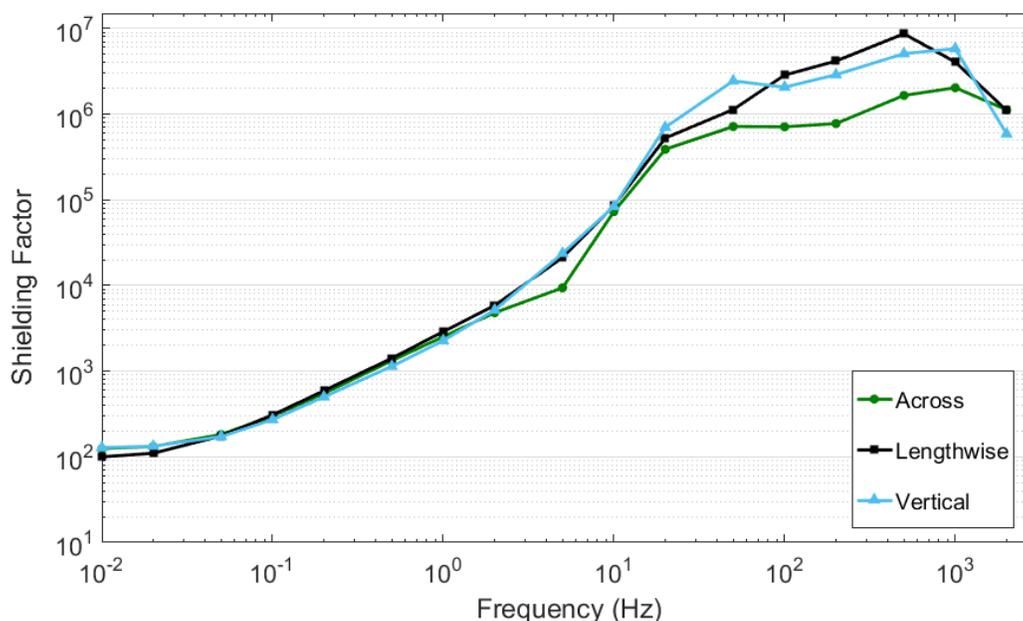


Fig. 6. Shielding factor of the shielded room as a function of frequency measured by Vacuumschmelze.

To prevent coupling of EMI into the room via electrical connections, the measurement system is battery operated and connected to the external environment primarily via optical links. Only the microwave inputs to the QVNS and the Peltier cooler of the TPW cell are directly connected via cables to electronic devices outside of the room (see Fig. 7). This arrangement is not optimal, but, in practice, no additional EMI was detected.

All electronic devices that were not needed inside the shielded room were located close to the outside walls of the shielded room to minimize connection length and coupling to outside noise sources. The Linux computer used for control and data collection, the rubidium clock, and the FPGA used to generate clock signals for the ADC are powered from a single ac-power subcircuit. These instruments are optically connected to the inside of the shielded room.

The Sympuls pulse generator, the 10 GHz clock generator that synchronizes the Sympuls pulse generator, the digitizer used to optimize the zero-compensation, and the TPW maintenance system controller are powered from a single ac-power subcircuit. A separate computer connected to a different ac-power subcircuit controls the pulse generator and the digitizer via optical interconnections.

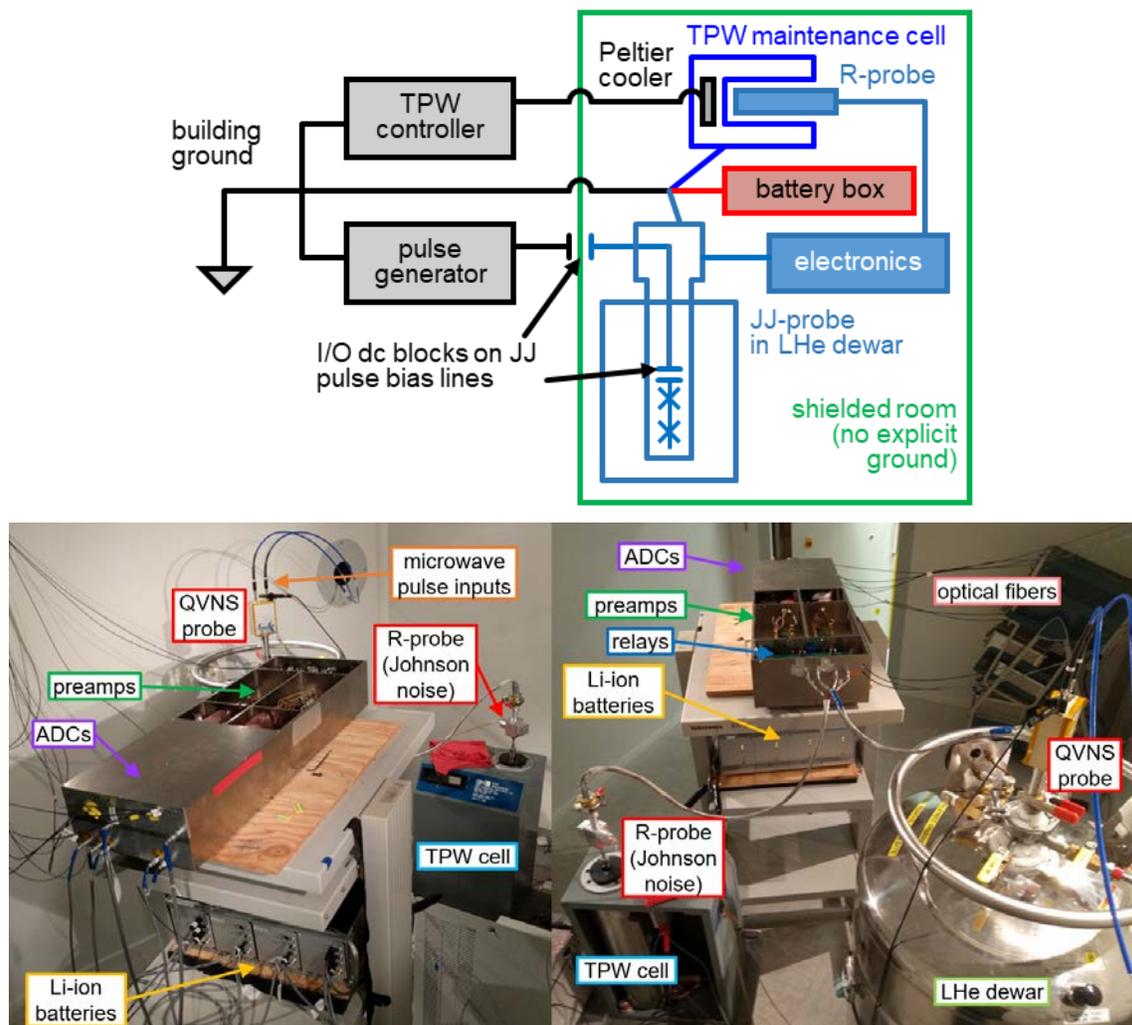


Fig. 7. (Top) Illustration of the star configuration used to ground and shield the experiment. (Bottom) JNT system inside a shielded room, including shielded electronics box, liquid helium Dewar with QVNS, and TPW maintenance cell with TPW and resistance probe. The electronics box cover is partly removed; during normal operation, the electronics box is fully enclosed.

The TPW maintenance system uses a superinsulated vacuum Dewar and a Peltier cooler that actively cools the dry-well chamber for the TPW cell by controlling a core heat exchanger to a set point temperature of approximately -0.25 °C. This maintenance system is designed to maintain a single TPW mantle for more than a month, but the digital control circuitry can add additional EMI. We therefore rebuilt the system with all of the electronics except for the actual Peltier element and thermometer outside of the shielded room. These connections are all differential and require a total of three twisted pairs of wires formed into a shielded, multiple-conductor cable, which is passed multiple times through a ferrite ring. The vacuum Dewar that holds the TPW is floating relative to the control electronics, though we chose to ground the portion of the cable shield that was inside the shielded room to the vacuum Dewar. The Peltier cooler also required a small fan that we powered from an isolated battery inside the battery box.

The ac-power subcircuit that powers the pulse generator also provides the ground for the electronics inside the shielded room. We use a star-configuration for the grounding (see Fig. 7), with three ground lines radiating from that central location: one to the TPW maintenance system vacuum Dewar, one to the battery box, and one to the top of the QVNS cryoprobe. From the top of the QVNS cryoprobe, there is a connection to the liquid helium Dewar. There is also a connection through the shield of the signal cables to

the electronics box, and from the electronics chassis (containing the relay board, amplifiers, and ADCs see Sec. 2.4) to the R-probe, again through the shield of the signal cables. This creates a continuous ground that entirely surrounds all of the voltage-noise signal sources and measurement electronics.

The cabling between the voltage sources and the electronics chassis continues the approach used within each probe of individually shielding each twisted pair. Each twisted pair and shield are fed through the top of the probe and connected using a three-pin LEMO connector (two connectors per probe). On the outside of the probe, the shield is connected to a metallic braid that surrounds the twisted pair. Both sets of shielded twisted pairs are contained within another ground shield, but isolated from the ground shield and each other. This outer ground shield is connected to the outside of the probe through the outer shield of the LEMO connectors. This process is reversed using similar three-pin LEMO connectors at the electronics chassis, and each twisted pair is connected to one relay board inside the chassis, while the shields are connected to the other relay board (see Sec. 2.3). In contrast, earlier versions of the measurement used a single 12-wire connector, which could potentially allow for cross-talk between the measurement channels.

The outer ground shield is connected to the measurement chain at the input to each amplifier. There is an explicit connection between the center pin of each amplifier and the surrounding, grounded electronics chassis. This connection, using a relay board, grounds the center of the measured voltage source, while the unmeasured voltage source remains floating, or is connected to the electronics chassis, depending on configuration (as discussed in Sec. 2.3).

The shielded room itself is left floating relative to our system. There did appear to be a $< 100 \text{ k}\Omega$ leakage path between the shielded room and the system ground. We also measured large EMI when the shielded room was grounded to the center of the star ground discussed earlier, independent of whether the center of the star ground was connected to the ground of the ac-power subcircuit used by the pulse generator. This EMI took ~ 10 min to damp when the shielded room was floated. We hypothesize that this is because of the magnetic nature of the Mu-metal shielding.

During the measurement in Ref. [37], the temperature of the shielded room also was not stable. Before this problem was understood, when we stopped collecting noise data and opened the door of the shielded room to measure the resistance of the sense resistor, change batteries, and change liquid helium Dewars, the temperature inside the room would increase, and the gain of the amplifiers would drop by $> 1 \%$. Then, when we closed the door and started taking data, the temperature would slowly drop, and the gain would slowly rise to the previous equilibrium, which caused our measurement to predominantly include data where the gain was increasing over time. After this problem was understood, we were able to equilibrate the inside and outside of the room so that opening the door had minimal effect. In this state, the room temperature drifted by $\sim 0.5 \text{ }^\circ\text{C}$, corresponding to gain drift of $\sim 0.1 \%$. In the end, we took noise data at temperatures between $16.4 \text{ }^\circ\text{C}$ and $24.4 \text{ }^\circ\text{C}$, we ignored data with a gain drift $> 0.3\%$, we corrected the measurement of the resistance of the sense resistor based on the room temperature (see Sec. 4.5), and we averaged the remaining data to remove small variations caused by asymmetric gain drift (see Sec. 4.3).

3. JNT System Model for Measuring Temperature and Correlated Noise

Recent precision JNT measurements using a QVNS reference have concentrated on determining the value of the Boltzmann constant by measuring the voltage-noise power spectral density of a resistor at the TPW. This has led to a data analysis procedure that concentrates on comparing (1) the ratio of the real part of the measured cross-correlation of the sense resistor $\Re[C_R] = S_R$ to the real part of the measured cross-correlation of the QVNS source $\Re[C_Q] = S_Q$ to (2) the ratio of the calculated power spectral densities

$$S_{R,\text{calc}} = 4kTR \text{ and } S_{Q,\text{calc}} = |v_{\text{JJ}}(f)|^2 K_1^{-2} / \Delta f .$$

In the typical JNT analysis, an even-order polynomial as a function of frequency $a_0 + a_2 f^2 + \dots$ is fit to the measured ratio S_R/S_Q , where this general model is based on the lumped element nature of the circuit elements and assumptions about circuit symmetry. The estimate of a_0 is compared to the calculated value $a_{0,\text{calc}} = S_{R,\text{calc}}/S_{Q,\text{calc}}$ to determine the consistency between the assumed value of the Boltzmann constant k and the measured value; that is, when $a_0 - a_{0,\text{calc}} = 0$, then the measured value of the Boltzmann constant is equal to the assumed value.

In this section, we present a mildly modified version of this analysis, which concentrates on transferring the magnitude and phase of the voltage sources through the entire measurement chain, which is assumed to be linear. First, we motivate the use of a complex ratio and the fitting of $\Re[C_R/C_Q]$ to an even-order polynomial, though in our most recent measurement [37], there is less than a 0.4 % difference in the relative offset and uncertainty using the different types of ratios. Second, we express the frequency-dependent ratio as the quantum-calibrated noise temperature of the resistor, with units of Kelvin. We then apply this approach to measure the temperature of a sense resistor that is collocated with the JJ arrays on the QVNS chip and use the symmetry of the measurement channels to directly determine the undesired correlated noise in our JNT system, which is not related to Johnson noise.

3.1 Measurement Model

To restate the JNT data analysis procedure, we start by defining complex, frequency-dependent voltage transfer functions $H_{A,R}(f)$ and $H_{B,R}(f)$ between the Johnson noise voltage source V_R and the input to the amplifier on channel A and channel B, respectively (see Fig. 1), and we similarly define voltage transfer functions $H_{A,Q}(f)$ and $H_{B,Q}(f)$ between the QVNS voltage source V_Q and the input to the amplifiers. We also define another pair of complex, frequency-dependent voltage transfer functions $G_A(f)$ and $G_B(f)$ between the input to the amplifiers and the data stored on the computer in the form of a time-stream of 16 bit integers proportional to voltage; because the same amplifiers and subsequent measurement hardware are used for both the QVNS and Johnson noise measurements, these transfer functions are the same for both voltage sources. The cross-correlation is performed on the computer after fast Fourier transforming about 1 s of data from each channel, resulting in measured cross-correlations between channels A and B for the two voltage sources:

$$C_R(f) = (V_R H_{A,R} + V_{A,n,R}) G_A \overline{(V_R H_{B,R} + V_{B,n,R}) G_B} = (|V_R(f)|^2 H_{A,R} H_{B,R}^* + C_{n,R}) G_A G_B^*, \quad (2)$$

and

$$C_Q(f) = V_Q H_{A,Q} G_A \overline{V_Q H_{B,Q} G_B} = (|V_Q(f)|^2 H_{A,Q} H_{B,Q}^*) G_A G_B^*, \quad (3)$$

where in Eq. (2), we have also added the possibility of undesired correlated noise, $C_{n,R} = V_{A,n,R} V_{B,n,R}^*$, due to some voltage noise source at the input to the amplifiers that is not related to, or correlated with, the Johnson noise. $C_{n,R}$ can be both complex and frequency-dependent. For example, this term would include current noise from one amplifier creating voltage noise in the other amplifier, but it would not include cross-coupling between the digitizers. Cross-coupling and other sources of correlated noise that occur between the input to the amplifiers and the data being stored on the computer are tested as part of the EMI test (Sec. 5.3) and are negligible. It is also possible to include a similar term, $C_{n,Q}$, in the QVNS cross-correlation measurement, but we assume $C_{n,Q} = 0$ because the JJ array acts as a low-inductance short that isolates the two measurement channels; removing this assumption results in more complicated equations but does not significantly change the conclusions herein.

It is also implicitly assumed that the evaluation frequencies for these equations are limited to those of the frequency comb ($f_0, 3f_0, 5f_0, \dots$). In the case of the QVNS source, there is no information about the QVNS tones in the gaps between the tones, and therefore the bandwidth of evaluation is a single frequency bin. Nonlinearities can introduce pattern-dependent effects at even multiples of f_0 , but this analysis assumes linearity, and this assumption is explicitly checked in Sec. 5.4. For the Johnson noise source, we sum the measured cross-correlation across a bandwidth Δf surrounding the frequency combs. It is our observation that Δf is small enough that the transfer functions are frequency-independent, within statistical error, over that bandwidth. In practice, we choose $\Delta f \approx 2f_0$ to maximize the bandwidth while keeping the measurement at each frequency independent.

Using these quantities, we take the complex ratio and assume that the extra correlated noise terms are small,

$$\frac{C_R}{C_Q} \approx \frac{|V_R(f)|^2 H_{A,R} H_{B,R}^*}{|V_Q(f)|^2 H_{A,Q} H_{B,Q}^*} + \frac{C_{n,R}}{|V_Q(f)|^2 H_{A,Q} H_{B,Q}^*}. \quad (4)$$

The typical JNT analysis assumes an approximately ideal case where channels A and B are the same between the sources and the input to the amplifiers; under this assumption, $H_{A,R} = H_{B,R} = H_R$ and $H_{A,Q} = H_{B,Q} = H_Q$, and the complex ratio simplifies to

$$\Re\left[\frac{C_R}{C_Q}\right] \approx \frac{|V_R(f)|^2 |H_R|^2}{|V_Q(f)|^2 |H_Q|^2} + \frac{\Re[C_{n,R}]}{|V_Q(f)|^2 |H_Q|^2} \quad \text{and} \quad \Im\left[\frac{C_R}{C_Q}\right] \approx \frac{\Im[C_{n,R}]}{|V_Q(f)|^2 |H_Q|^2}. \quad (5)$$

For comparison, if we take the real part of each measured cross-correlation (as is done in the typical JNT analysis) and additionally assume that the two measurement channels are the same between the input to the amplifier and the computer (including the amplifiers, filters, and digitizers), where $G_A = G_B$, then we end up with the same result:

$$\frac{\Re[C_R]}{\Re[C_Q]} \approx \frac{|V_R(f)|^2 \Re[H_{A,R} H_{B,R}^* G_A G_B^*]}{|V_Q(f)|^2 \Re[H_{A,Q} H_{B,Q}^* G_A G_B^*]} + \frac{\Re[C_{n,R} G_A G_B^*]}{|V_Q(f)|^2 \Re[H_{A,Q} H_{B,Q}^* G_A G_B^*]} \approx \frac{|V_R(f)|^2 |H_R|^2}{|V_Q(f)|^2 |H_Q|^2} + \frac{\Re[C_{n,R}]}{|V_Q(f)|^2 |H_Q|^2}. \quad (6)$$

A circuit composed of linear lumped elements has $|H|^2 = h_0 + h_2 f^2 + h_4 f^4 + \dots$, and most sources of correlated noise can also be modeled as having an even-order frequency dependence with negligible magnitude at dc for similar reasons. Therefore, the typical JNT analysis fits the real ratio to an even-order polynomial $a_0 + a_2 f^2 + \dots$, where a_0 is the term of interest that is proportional to temperature and the Boltzmann constant, while the higher-order terms are caused by correlated noise and imperfect matching between the transfer functions of the voltage sources H_R and H_Q .

However, in a real system, no two channels are perfectly matched. A lack of phase-matching can introduce errors where the two voltage sources have different phase shifts and where the imaginary part of the cross-correlated noise affects the real part of the measured ratio. This error is typically frequency-dependent because phase-matching is excellent at low frequencies but gets progressively more difficult as the frequency increases. Notably, the complex ratio does not depend on differences between the amplifiers, multi-pole alias filters, and digitizers; these active and complicated elements are particularly difficult to phase-match. On the other hand, if these components are not phase-matched, then taking the ratio of the real parts of the cross-correlators could allow some of the imaginary $C_{n,R}$ to leak into the ratio.

In our recent measurement [37], there is negligible difference between the $\Re[C_R/C_Q]$ approach and the $\Re[C_R]/\Re[C_Q]$ approach. When the same cross-validation model selection and bandwidth selection procedures [40] are applied to both cases, the resulting relative offsets agree to better than 0.4 %, and their associated uncertainties agree to better than 0.2 %. This is expected, since the imaginary parts of the both C_R and C_Q are small compared to the real parts, but this choice of ratio technique may become more important at higher frequencies, where phase matching is more difficult. Looking forward, calculating the ratio of real parts (Eq. 6) has the disadvantage that the mismatch between G_A and G_B can cause the imaginary component of the undesired correlated noise to impact the ratio of the real parts of the cross-correlators, whereas calculating the real part of the complex ratio (Eq. 5) is insensitive to this effect. Rather than emphasizing one specific approach, in the future it will be important to apply both techniques and understand the source of any differences in the results.

3.2 Measuring Correlated Noise and Temperature at 4 K

In Eq. (4), either an imperfect match, where $H_{A,R} H_{B,R}^* \neq H_{A,Q} H_{B,Q}^*$, or correlated noise $C_{n,R}$ can lead to the ratio C_R/C_Q having frequency-dependence or an imaginary component. This ambiguity made it difficult to experimentally determine the quality of the match and the magnitude of the correlated noise during our

recent Boltzmann constant measurement [37]. Correlated noise due to amplifiers has been considered in studies such as Refs. [56, 57].

To directly measure the presence of correlated noise, we modified the NIST JNT system to measure Johnson noise at 4 K from resistors collocated with the QVNS source (see Fig. 8). We simultaneously measure the QVNS frequency tones and the cryogenic Johnson noise using the same physical hardware, so $H_{A,R} = H_{A,Q} = H_A$ and $H_{B,R} = H_{B,Q} = H_B$, though $H_A \neq H_B$. We also expect that, because of this system modification, these transfer functions differ from the transfer functions during the Boltzmann constant measurement. Specifically, the stray capacitance increases relative to both the previous sense resistor and QVNS measurements, because the sense resistor is now in the Dewar with the longer lines, and because the QVNS no longer shorts the effect of one measurement channel relative to the other, so the stray capacitance approximately doubles.

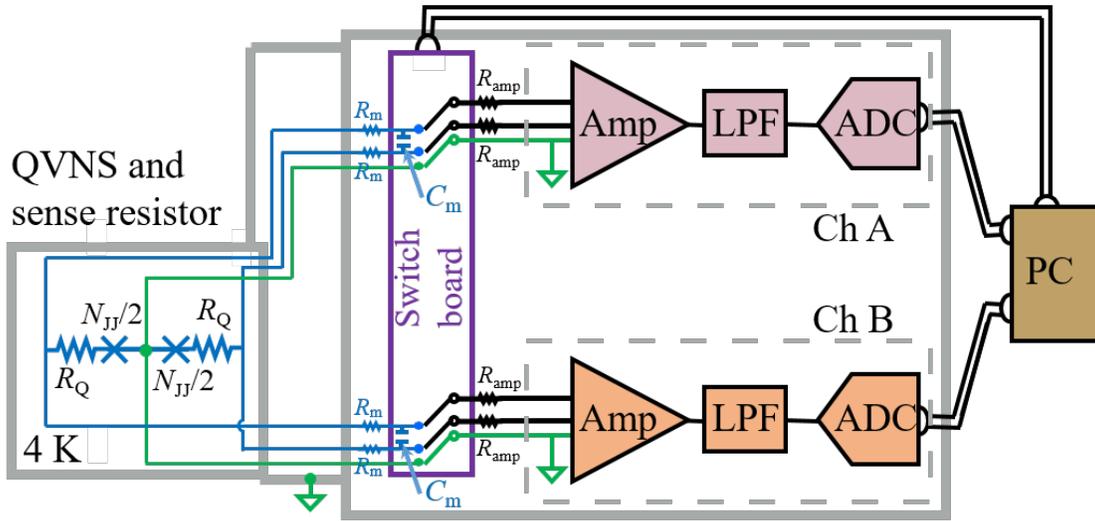


Fig. 8. Johnson noise thermometer schematic diagram with cryogenic Johnson noise sense resistors R_Q located on the QVNS chip.

Because the two voltage sources are simultaneously present in the measured cross-correlator, the complex ratio that is equivalent to the typical JNT ratio, that is, Eq. (4), requires subtraction of the measured Johnson noise C_R from the measured cross-correlator at the QVNS tone frequencies C_Q before taking the ratio

$$\frac{C_R}{C_Q - C_R} \approx \frac{|V_R(f)|^2}{|V_Q(f)|^2} + \frac{C_{n,R}}{|V_Q(f)|^2 H_A H_B^*}, \quad (7)$$

where the term that includes the ratio of voltage sources is not affected by any transfer functions and should therefore be frequency-independent. This has two important effects: First, any frequency dependence in the real part of the ratio is due to correlated noise; second, any imaginary component to the ratio is also due to correlated noise. However, since the correlated noise is scaled by the transfer functions $H_A H_B^*$, we still need additional assumptions to directly extract the real and imaginary parts of the correlated noise.

To use more relevant units and easily extract the temperature of the sense resistor, we also define a calibration constant $T_{\text{cal}}(f)$ as a function of frequency so that the quantity $T_{\text{cal}} C_R / (C_Q - C_R)$ has units of kelvin.

$$T_{\text{cal}} = |v_{JJ}(f)|^2 K_J^{-2} / (4kR \Delta f). \quad (8)$$

This calibration constant is closely related to $a_{0,\text{calc}}$, with $a_{0,\text{calc}} \approx T_{\text{cal}}/T$.

In Fig. 9, we plot the real and imaginary parts of the measured ratio $T_{\text{cal}}C_R/(C_Q - C_R)$ as functions of frequency. We observe correlated noise, where $\Re[C_{n,R}/H_A H_B^*]$ has approximately f^2 dependence, and $\Im[C_{n,R}/H_A H_B^*]$ has approximately f -linear dependence on frequency up to 600 kHz. The magnitude of the noise is significant; we expect that this is a lower bound on the correlated noise present during the Boltzmann constant measurements, because the additional capacitance in this measurement configuration would tend to reduce the magnitude of the correlated noise. Therefore, at 1 MHz, we would expect at least a 3 % effect in the real part of the ratio in the TPW measurement. We also measure an average temperature of $T_0 = 4.15$ K in a pressurized liquid helium Dewar by fitting to $T_0 + T_2 f^2$ from 5 kHz to 600 kHz, though it is expected that the temperature drifts as a function of pressure and liquid helium level.

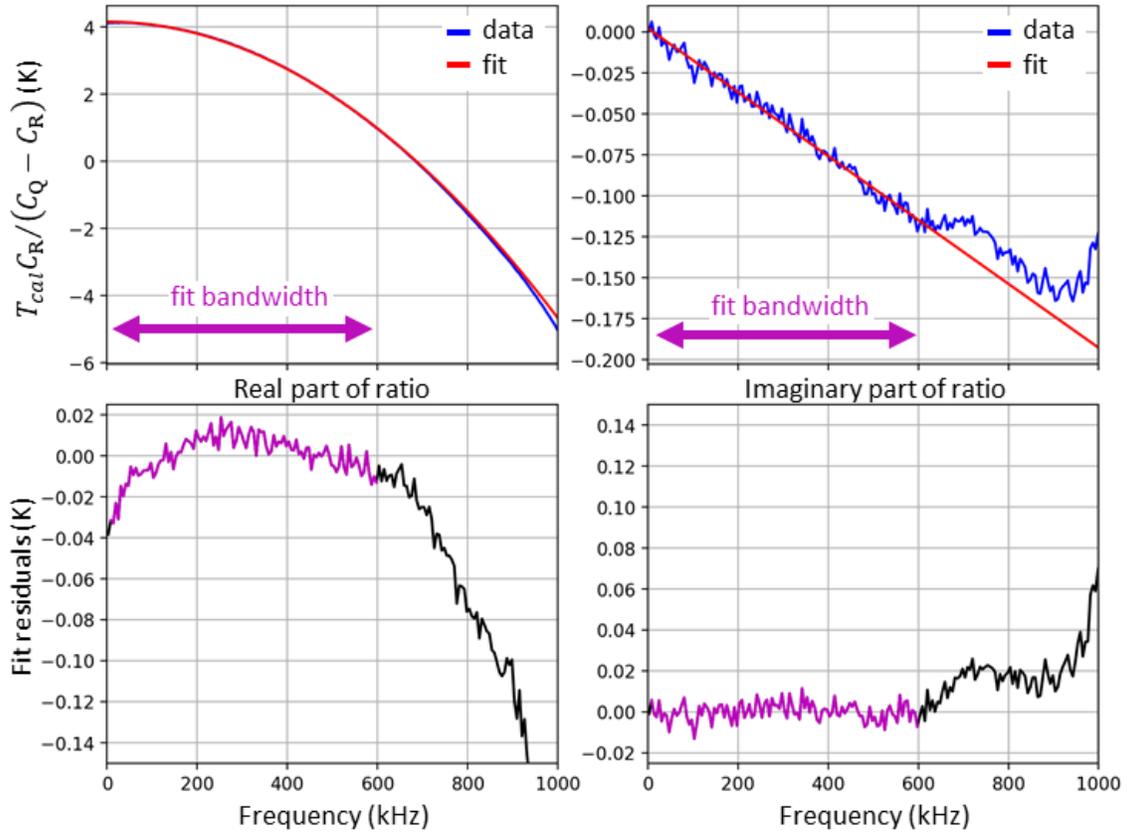


Fig. 9. Measured ratio of cryogenic Johnson noise to QVNS in units of kelvin (blue) with fit (red); the real part of the ratio is plotted on the left, and the imaginary part is on the right, with residuals of the fit shown below (magenta line over the fit bandwidth).

One possible source for this correlated noise component, with a real part quadratic in frequency and an imaginary part linear in frequency, is current noise in the input stage of the FET amplifier [55] (shown in Fig. 4). In this type of FET amplifier, the primary source of current noise is thermal fluctuations in the FET channel and any other resistive loads on the drain. This causes a white-noise current I_{nd} to flow through the drain of each FET on our two differential amplifiers. I_{nd} causes equivalent voltage noise I_{nd}/G on the gate electrode of each FET with transconductance G . It also causes a fluctuating voltage $I_{\text{nd}}R_d$ on the drain with resistive load R_d . This drain voltage couples to the gate through the gate-drain capacitance C_{gd} and causes noise current $I(\omega) = j\omega C_{\text{gd}}R_d I_{\text{nd}}$ to flow out of the gate and into the sense resistor.

In the JNT system, the measured voltage noise source is connected to the gate electrode of a total of four FETs (see Fig. 4 and Fig. 8). Each senses its own transconductance-coupled voltage noise, and the noise due to all four FET gate currents flows through the JJ arrays and the thermal noise source resistance

R_T . The impedance of the JJ arrays is negligible compared to R_T ; therefore, the voltage on differential amplifier #1 from FET noise currents will be

$$V_1 = (I_{nd1+} - I_{nd1-})/G + j\omega C_{gd}R_d(I_{nd1+} - I_{nd1-} + I_{nd2+} - I_{nd2-}), \quad (9)$$

and likewise for amplifier #2. The cross-correlation of the two voltage channels will then be

$$\begin{aligned} \langle V_1 V_2^\dagger \rangle = & -\omega^2 C_{gd}^2 R_d^2 R_T^2 (\langle I_{nd1+}^2 \rangle + \langle I_{nd1-}^2 \rangle + \langle I_{nd2+}^2 \rangle + \langle I_{nd2-}^2 \rangle) \\ & + j(\omega C_{gd} R_d R_T / G) (-\langle I_{nd1+}^2 \rangle - \langle I_{nd1-}^2 \rangle + \langle I_{nd2+}^2 \rangle + \langle I_{nd2-}^2 \rangle). \end{aligned} \quad (10)$$

The real part of this contribution to the correlated noise is negative and quadratic in frequency, whereas the imaginary part is linear in frequency, which is consistent with the results of Fig. 9. If each distribution of drain current noise $I_{nd(1.2)\pm}$ is equal, then the imaginary component should vanish. The observed linear dependence of the imaginary component on frequency is therefore caused by a combination of unequal current noise powers $\langle I_{nd(1.2)\pm}^2 \rangle$ or by imperfect matching of transfer functions between the two channels $\Re[H_A H_B^*] \neq 0$. Note that the previous section described the possibility that the imaginary part of the undesired correlated noise could modify the ratio $\Re[C_R]/\Re[C_Q]$ in a frequency-dependent way. When the imaginary correlated noise is small, as in this case, the leakage into the real ratio is negligible, and $\Re[C_R]/\Re[C_Q] \cong \Re[C_R/C_Q]$. Assuming equal drain noise currents on each FET, then the real part of Eq. (10) becomes

$$\Re[\langle V_1 V_2^\dagger \rangle] = -4\omega^2 C_{gd}^2 R_d^2 R_T^2 \langle I_{nd}^2 \rangle = -4I(\omega)^2 R_T^2. \quad (11)$$

Applied to the measured quadratic dependence on frequency in Fig. 9, this equation reveals that $I(2\pi f) = 7.8 \times 10^{-19} \text{ A/Hz}^{3/2} \times f \text{ Hz}$. At $f = 600 \text{ kHz}$, for example, the current noise of each FET driven through the sense resistor will be $0.47 \text{ pA}/\sqrt{\text{Hz}}$; this small value is typical of JFETs. In summary, the real component of the measured correlated noise scales negatively with frequency squared, and its magnitude implies plausible FET gate currents, whereas the imaginary part is linear in frequency and small, as expected for well-matched transistors with approximately equal drain current noise. We conclude that the correlated noise is consistent with FET noise currents.

Future work may include direct measurements of JFET current noise, to more explicitly confirm that this unintended correlated noise is indeed current noise through the input transistor. If current noise does appear to be a dominant source of unintended correlated noise, future preamplifier designs may seek to lower current noise through the sense resistor by using FETs with lower gate-drain capacitance. Typically, FETs with low noise voltage have high capacitance, and vice versa, so this would be a non-trivial optimization problem.

4. Boltzmann Constant Measurement

We recently reported [37] a measurement of the Boltzmann constant $k = 1.380642 \text{ 9(69)} \times 10^{-23} \text{ J/K}$, with a relative standard uncertainty of 5.9×10^{-6} and relative offset of -5.47×10^{-6} from the Committee on Data for Science and Technology (CODATA) 2014 recommended value [16]. In this section, we present details of the measurement that were not included in [37], and in the next section, we describe certain parts of the uncertainty budget in more detail.

We start by describing the procedure we used to attempt to optimize the matching between the measurement channels. We then briefly discuss the effects of aliasing on the measurement and how we compensated for some of those effects, as well as detailing the averaging procedure used in processing the data. Finally, we describe our measurements of the triple point of water temperature realization and the resistance of the Johnson noise sense resistor.

4.1 Tuning the Electrical Circuit

Ideally, all the voltage transfer functions $H_{A,R}$, $H_{B,R}$, $H_{A,Q}$, and $H_{B,Q}$ would be exactly equal. In practice, the inherent differences in the output impedance of the JJ arrays in the QVNS (which are purely inductive) and the Johnson noise source (which is purely resistive and $200\ \Omega$ in this experiment) and the required physical locations at cryogenic temperatures and in a TPW cell, respectively, make matching the transfer functions difficult.

The first step in our matching procedure, as used in Ref. [37], is to fabricate resistors on the QVNS chip to partly match the output impedance of the $200\ \Omega$ Johnson noise resistor. However, exactly matching the output impedance of the noise resistor by adding a $200\ \Omega$ resistor directly in series with the JJ array also adds 4 K Johnson noise to the QVNS signal. We addressed this by fabricating separate $100\ \Omega$ resistors on each output lead, for a total of $200\ \Omega$ per channel. Since different resistors are used on each channel, the 4 K Johnson noise measured by each channel is uncorrelated and does not contribute to the measured QVNS cross-correlation C_Q .

The resistance of the on-chip QVNS resistors is slightly smaller than the Johnson noise resistor. In Ref. [37], we measured the dc resistance at the input to the relay board between the differential inputs of each channel of both the QVNS and Johnson noise and added $2\ \Omega$ series tuning resistors at the input of the relay board to all the QVNS leads. After matching, the dc resistance of the two QVNS channels was $201.80\ \Omega$ and $201.60\ \Omega$, while the resistance of the two Johnson noise channels was $201.75\ \Omega$ and $201.79\ \Omega$.

We then tuned the parallel matching capacitors on the QVNS leads. Each QVNS lead has a ~ 100 pF fixed surface-mount capacitor between the differential leads as well as a 5–57 pF trimmer capacitor (part number BFC280908003 from Vishay BC Components), which uses a polytetrafluoroethylene dielectric and has excellent long-term stability. We must then add extra parallel capacitance to the QVNS circuit, even though the individual QVNS cables must reach the bottom of a liquid helium Dewar and are therefore longer than the individual Johnson noise cables, because the JJ array on the QVNS acts as a low-inductance short. This isolates the two sides of the QVNS circuit, so that the parallel capacitance of the two channels is independent. However, for the Johnson noise circuit, the parallel capacitance of the cable and amplifier for channel 2 also capacitively loads channel 1 (and vice versa). Therefore, we must add additional capacitance to each QVNS circuit to match these additional loads.

In the recent Boltzmann constant measurement [37], we averaged data across multiple tuning capacitor configurations and modified the tuning capacitors based on the shape of the measured ratio C_R/C_Q as a function of frequency. Ideally, the imaginary part of the ratio $\Im[C_R/C_Q]$ would be equal to zero, and the real part $\Re[C_R/C_Q]$ would be independent of frequency. In practice, the imaginary part of the ratio is linearly dependent on frequency, and the slope can be tuned by making differential changes to the tuning capacitors (for example, increasing the value of one trim capacitor while decreasing the value of the other trim capacitor). This can be thought of as matching one arm of the QVNS to the other arm of the QVNS.

In practice, the real part of the ratio $\Re[C_R/C_Q]$ is largely dependent on common changes to the tuning capacitors (for example, increasing both trim capacitors), which can be thought of as matching the QVNS arms to the Johnson noise arms. The frequency dependence of the real part of the ratio is more complicated than that of the imaginary part; we have observed both monotonically increasing and monotonically decreasing data, and when we try to minimize the frequency dependence, then we typically see nonmonotonic behavior. We assume that this is due to the detailed differences between the different circuits and the presence of undesired correlated noise $C_{n,R}$. This behavior can be roughly reproduced by circuit models.

During the measurement in [37], we tried to optimize the trim capacitors to meet the conflicting goals of making the frequency dependence of the ratio both small and monotonic. Having ratio data that are monotonic in frequency is favorable because they have a greater likelihood of being fit by a low-order polynomial. However, this approach assumes that any undesired correlated noise $C_{n,R}$ that is not associated with Johnson noise is small; as shown in Sec. 3.2, this is not necessarily the case. Therefore, during this measurement [37], we were likely tuning the trim capacitors away from the best match to compensate for the undesired correlated noise, which resulted in data best fit by a higher-order polynomial.

In the end, we averaged over a range of different trim capacitor settings, as well as two different configurations for the amplifier input and grounding (with and without common-mode ferrite chokes, as described in Sec. 2.3). The stray capacitance also changed whenever we thermally cycled the QVNS probe after consuming most of the liquid helium in the storage Dewar. In Fig. 10, we plot the average of the real and imaginary parts of the ratio, separated into traces with different stray capacitance (colors) and with different configuration for the amplifier input and grounding (left and right plots). For more details on how the use of different transfer functions impacts the data analysis, see Sec. 5.1.

The recent paper from NIM [36] used a slightly different approach, where the dc resistances of the QVNS channels were chosen to be twice the Johnson noise resistor channels. This design was intended to compensate for the difference in parallel capacitance and provide the same effective RC time constant to both noise sources. NIM did not add any additional parallel capacitance to tune the match and instead aimed for identical cable lengths.

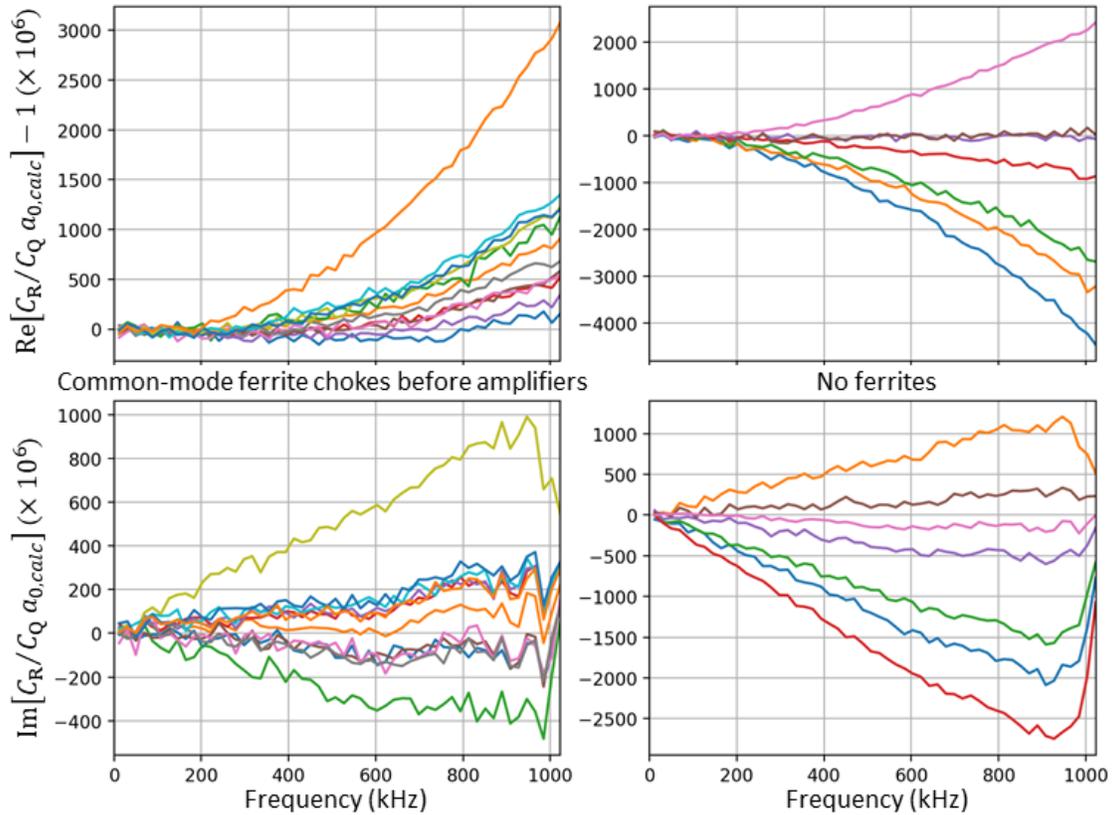


Fig. 10. Real (top) and imaginary (bottom) parts of the complex ratio with the two different amplifier configurations. The earlier configuration with common-mode ferrite chokes is on the left, and the later configuration without ferrite chokes but increased resistance is on the right.

4.2 Compensation for Aliasing

One additional complication in evaluating the frequency dependence of the ratio C_R/C_Q in our recent data set is the frequency of the anti-alias filter relative to the ADC sampling frequency. We use a passive 800 kHz 11-pole Butterworth anti-alias filter, while the Nyquist frequency f_{nyq} of the ADC is 1.04 MHz. Because the filter only imperfectly rejects signals at frequencies $> f_{nyq}$, the resistor noise as measured by the ADC C'_R has a small component from the second Nyquist zone:

$$C'_R(f) = C_R(f) + C_R^*(2f_{\text{nyq}} - f). \quad (12)$$

However, the peaks in the frequency comb of the QVNS at frequencies above f_{nyq} do not overlap with the peaks below f_{nyq} ; therefore, unlike with the Johnson noise, we separately measure the contributions from the first Nyquist zone and second Nyquist zone. To compensate for the aliasing in the Johnson noise measurement, we modify the QVNS term used in the ratio to $C'_Q(f)$, where

$$\frac{C'_Q(f)}{|v_{\text{JJ}}(f)|^2} = \frac{C_Q(f)}{|v_{\text{JJ}}(f)|^2} + \frac{C_Q^*(2f_{\text{nyq}} - f)}{|v_{\text{JJ}}(2f_{\text{nyq}} - f)|^2}. \quad (13)$$

The frequencies $2f_{\text{nyq}} - f$ are not quite the same as the QVNS tone frequencies, so we use two different interpolation methods to determine $C_Q^*(2f_{\text{nyq}} - f)$. At frequencies near the Nyquist frequency, where the signal-to-noise ratio is very high, we use a linear interpolation between the measured data points. At above 1200 kHz, that is, more than 200 kHz above the Nyquist frequency, we interpolate the effect of the low-pass anti-aliasing filter using a logarithmic fit.

Ideally, we would include this aliasing in the fit to the even-order polynomials as a function of frequency. However, in practice, we restrict our analysis to frequencies where the aliasing is small. In Fig. 11, we plot $|C_Q(2f_{\text{nyq}} - f)/C'_Q(f)|$ as a function of frequency f for 2 d of typical data and see that at frequencies f below 700 kHz, the fractional contribution of the second Nyquist zone to the measured data is less than 4×10^{-6} .

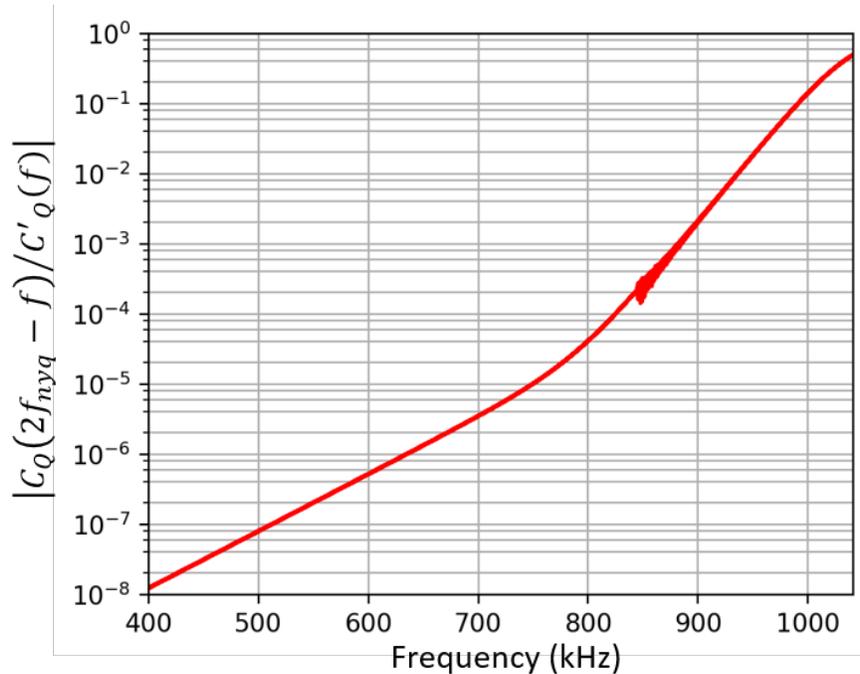


Fig. 11. Typical fractional contribution of noise in the second Nyquist zone to the total measured noise. The transition in interpolation techniques occurs at 850 kHz.

4.3 Details of the Averaging Procedure

Our raw cross-correlation data are saved as an average of 100 complex spectra with a frequency resolution of ~ 1 Hz at a sampling rate of 2.083×10^6 samples/s. We further average the data in both frequency and time before calculating the ratio C_R/C_Q . First, the Johnson noise data are averaged into

maximally nonoverlapping bins centered on the frequencies of the QVNS comb. If the number of frequency points between QVNS comb frequencies is odd, then all the data are included; otherwise, a single frequency point between combs is dropped. In our recent measurement [37], a QVNS tone spacing of 159 Hz with 160 frequency points between QVNS tones results in a Johnson noise measurement bandwidth of 158 Hz per bin, that is, per QVNS tone.

After this first round of averaging, we further divide the data into sets that are N chops (N 100 s units) long in time and P comb frequency points long in frequency. The Johnson noise data C_R are averaged over the full $N \times P$, but the QVNS data C_Q are only averaged over the $(N - 1) \times P$ set centered in time on the Johnson noise data. Since the QVNS data and the Johnson noise data are not taken simultaneously, this procedure will remove the effect of net gain drift over the N chops used in the average. In our earlier analysis [37], we used $N = 11$ and $P = 35$, so that each ratio value was derived from an average over ~ 1100 s and a bandwidth of 2.52 kHz. These averaging procedures were applied before the data were sent to the cross-validation routine (Sec. 5.1).

This averaging has two goals: First, we want to compensate explicitly for drift over time as early in the analysis process as possible. Second, we want to increase the signal-to-noise and reduce the size of the data set without losing relevant information about changes in the data as a function of frequency or time. Reducing the size of the stored data by a factor of approximately $N \times P = 385$ satisfies this goal, and we did not detect any structure in the data at frequency scales of 5 kHz or time scales of 30 min.

4.4 Triple Point of Water Realization

The working cell used in [37] provided a direct realization of the SI kelvin, but checks are required to verify the estimated realization uncertainties (see Sec. 5). This section describes the auxiliary testing and cell comparisons used to support the uncertainties.

The R-probe was designed to achieve adequate immersion into a standard type B TPW cell with a 13 mm bore. This was verified by using a diagnostic thermistor sensor, which was embedded in a copper heat sink near the position of the LCC package inside of the probe. The thermistor has a sensitivity of 1.66 Ω /mK and, when used with a digital substitution bridge, can resolve 20 μ K changes. An immersion profile for the probe may be generated by determining the change in equilibrium temperature with varying immersion depth. At the uppermost portion of the immersion coordinate, an exponential warming trend normally occurs, and in the lowermost portions, a linear trend should be consistent with the hydrostatic head effect, exhibiting a slope of -0.0073 mK/cm, where the sign indicates colder temperatures with increasing immersion depth. The immersion profile for the JT4B probe as used inside the working cell is shown in Fig. 12. These results are consistent with adequate immersion for immersion coordinates below 8 cm.

The results of TPW cell comparisons carried out in 2015 and 2017 are shown in Table 1. The comparisons were performed with a 25.5 Ω standard platinum resistance thermometer (SPRT) and the DSB in Boulder, Colorado, and with other bridges of similar specifications in Gaithersburg, Maryland. Cells with serial numbers starting with “Q” are made from fused quartz glass. The observed temperature difference between cells 1302 and Q1454 was larger than typical cell comparisons, but is still not a large contribution to the combined error analysis presented in Sec. 5.

Table 1. Summary of TPW cell comparison results involving the working cell, 1302, and other transfer cells as used in both NIST facilities in Gaithersburg and Boulder.

| Date | Location | Cell 1 | Cell 2 | $T(1) - T(2)$ /mK | $u(DT)$ /mK |
|------|--------------|--------|--------|-------------------|-------------|
| 2015 | Gaithersburg | Q1454 | Q5009 | 0.010 | 0.03 |
| 2015 | Boulder | 1302 | Q1454 | -0.061 | 0.027 |
| 2017 | Boulder | 1302 | Q1454 | -0.205 | 0.20 |
| 2017 | Gaithersburg | Q1454 | Q1034 | -0.024 | 0.023 |
| 2017 | Gaithersburg | 1302 | Q1454 | -0.15 | 0.05 |

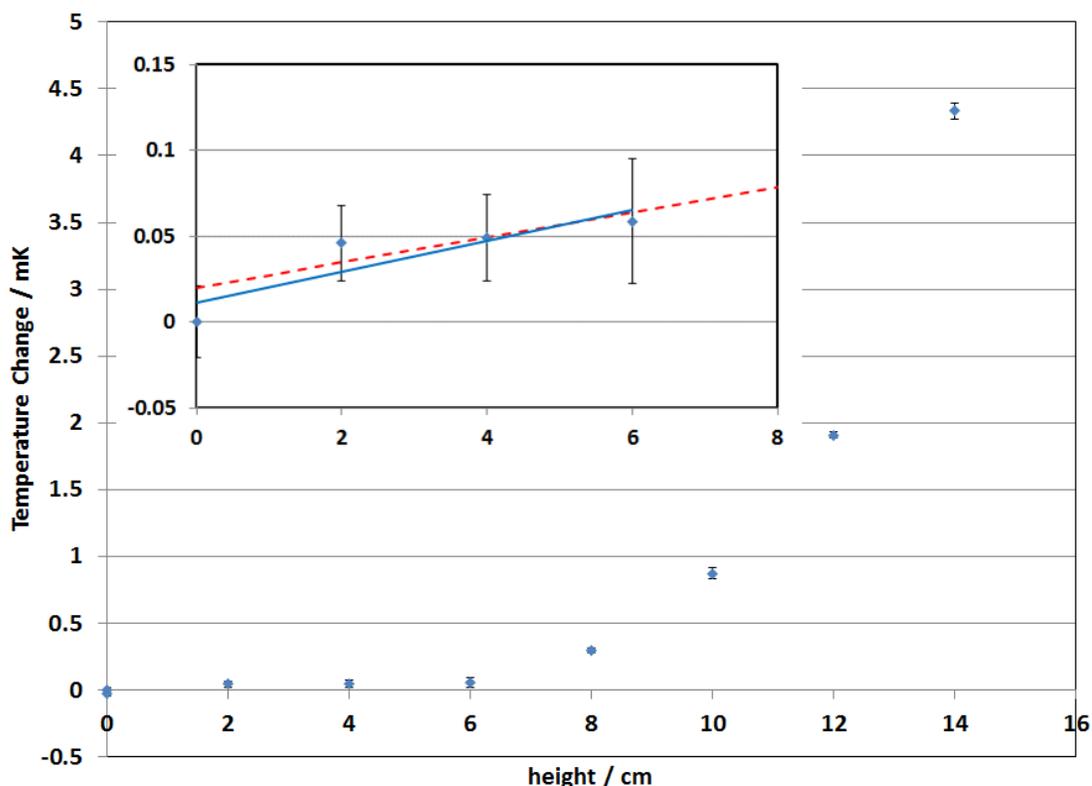


Fig. 12. The immersion profile for the R-probe JT4B inside the working cell. The immersion coordinate is a relative height for the probe, where 0 is taken when the probe is sitting at the bottom of the thermowell. An exponential warming characteristic is observed for the probe height of 8 cm and greater. Inset: For probe heights less than 8 cm, the observed temperature changes are linear (solid blue line is a linear fit) and consistent with the expected linear trend as predicted by the known hydrostatic head effect (red dashed line) of -7.3 nK/cm of depth.

4.5 Resistance Measurements

In Ref. [37], the two-terminal pair resistance measurements were made for the R-probe $200\ \Omega$ series combination resistors at intervals of 2 to 7 d using the DCC bridge from 1 March 2017 to 7 April 2017. Starting on 7 April, measurements were switched to exclusive use of the DSB at intervals of 1 to 3 d. In both cases, an excitation current of 0.25 mA was used to limit the power dissipation to $6.25\ \mu\text{W}$ in each of the two $100\ \Omega$ resistors and a total of $12.5\ \mu\text{W}$ inside of the LCC package. The effective power coefficient for the series combined $200\ \Omega$ resistors inside the LCC was $0.005(3)\ \mu\Omega/\Omega\cdot\mu\text{W}^{-1}$. Measurements made using both bridges on 7 April agreed to within $0.08\ \mu\Omega/\Omega$ difference.

The $100\ \Omega$ resistance standard, 1202T, is located inside the shielded room and passively equilibrates with that room's ambient temperature. The resistance value for the standard $R_{1202T}(T)$ has a mild quadratic temperature dependence, with a relative maximum near $T = 24.5\ ^\circ\text{C}$. Since all measurements are performed at ambient temperatures below that point, we apply corrections to the calibration value based on either direct or indirect temperature measurements. The direct temperature measurements are based on an SPRT (s/n 3368) as installed in the measurement thermowell of the 1202T standard starting from 19 April 2017. Those measurements have variable statistical uncertainties anywhere between 2 mK and 65 mK, depending on the ambient stability during the measurement. Prior to 19 April, the indirect temperature measurements are based on inferred values from thermostat controls, preamplifier gain, and the observed drift rate of the R-probe resistors. Those inferred values are highly uncertain and are assigned a 0.25 K standard uncertainty. Temperature-dependent corrections to the standard resistance may have been larger than optimal because the room's temperature was much colder than the standard resistor's ideal operating

temperature. The largest correction was $1.1 \mu\Omega/\Omega$, which was made when the room reached its coldest point of $16.4 \text{ }^\circ\text{C}$ on 26 April 2017.

The metal foil resistors used for sense resistors in this and our other previous works are susceptible to thermally induced stress relaxation effects, as are all resistors. These effects manifest as a small negative drift in the measured resistance, having variable rates with time. Once a stress is applied and removed, the subsequent drift will exhibit weak exponential or stretched-exponential characteristics. When monitored for longer time scales, the drift will appear approximately linear with a constant rate until some subsequent thermal perturbation is applied, usually by a finite temperature excursion. These perturbations are avoided by always keeping the R-probe near $0 \text{ }^\circ\text{C}$. The drift is shown in Fig. 13 and is fit by a simple exponential characteristic.

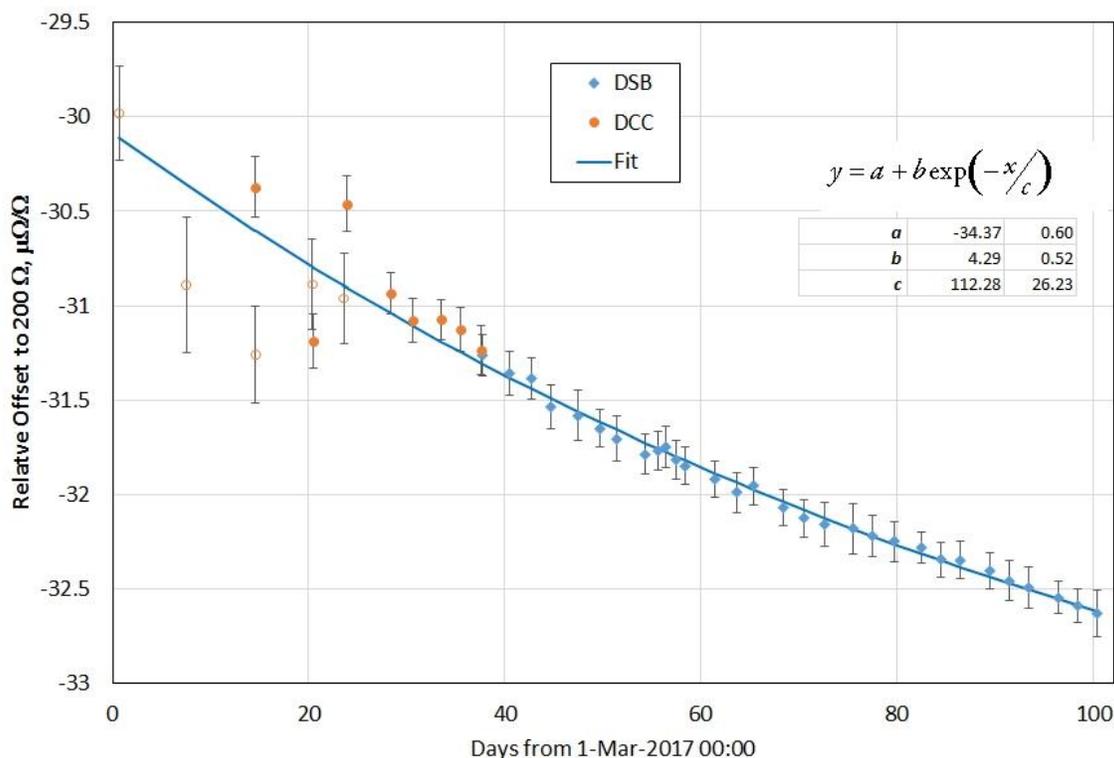


Fig. 13. The 100 d record of resistance measurements made on the R-probe combined 200Ω resistance using the DCC and DSB. The DCC data taken prior to 29 March have higher uncertainty. The weighted data are fit to a simple exponential relaxation. The fitted relative drift rates over the entire 100 d record vary between $-0.038 \times 10^{-6} \text{ d}^{-1}$ and $-0.016 \times 10^{-6} \text{ d}^{-1}$.

5. Uncertainty Analysis and Discussion

In this section, we add additional details to the description of the uncertainty budget from Ref. [37]. The uncertainties are organized into four primary categories of factors that determine the value of k . The first category concerns the determination of the parameter a_0 from the measured ratio of noise powers, and it is the dominant source of uncertainty. This category includes: the statistical uncertainty estimates and model selection uncertainty estimates, which are determined using a cross-validation analysis, EMI, nonlinearity, and spectral aberrations. We also supplement the cross-validation analysis by estimating the expected statistical uncertainty based on a simple frequency-independent noise model.

The second category is specific to the QVNS synthesis and noise waveform realization, and it has negligible effect on the total uncertainty. The third category is specific to the realization of the kelvin via the TPW, and the fourth category concerns the resistance measurements and traceability to the QHE

realization. The tabulated uncertainties are listed at the end of Sec. 5, in Table 3, expressed as relative uncertainties in parts per million.

5.1 Cross-Validation Method

In Ref. [37], we followed the typical JNT cross-validation analysis [35, 36, 40] and modeled the ratio of the real parts of the expected cross-correlators $\langle \Re[C_R] \rangle / \langle \Re[C_Q] \rangle$ as a d th order even polynomial function of frequency as follows:

$$\frac{\langle \Re[C_R] \rangle}{\langle \Re[C_Q] \rangle} = \sum_{i=0}^{i_{\max}} a_{2i} \left(\frac{f}{f_0} \right)^{2i}, \quad (14)$$

where $d = 2i_{\max}$, f_0 is a reference frequency, and $a_0 - a_{0,\text{calc}}$ gives the consistency between our measurement and the value of the Boltzmann constant used to calculate $a_{0,\text{calc}}$.

When determining a_0 from experimental data, the assumed order d (complexity) of the polynomial model and the maximum frequency analyzed (fitting bandwidth or f_{\max} for short) dramatically affect results. For each fitting bandwidth, we independently select the complexity of the model by a cross-validation method similar to that in Refs. [35, 36, 40].

In our analysis, for any frequency of interest, we acquire multiple measurements of C_R and C_Q . Data acquisition times are the same for each measurement. From these multiple measurements, we estimate $\langle \Re[C_R] \rangle / \langle \Re[C_Q] \rangle$ as the ratio of the average of all the C_R measurements to the average of all the C_Q measurements. In contrast, if we had calculated the average of the ratio C_R/C_Q from each observation, we would have introduced unnecessary systematic error due to the nonzero variance of the denominator term C_Q . It is difficult to estimate this variance accurately; instead, our method of averaging observations of C_Q makes both the variance of the denominator term and the associated systematic error negligible. For more discussion of this point, see Ref. [24].

In our recent measurement [37], we averaged over several different transfer functions $H_{A,Q}H_{B,Q}^*$ (see Fig. 10, which shows the data separated by approximate transfer function). This averaging approach is still reasonable because each of these transfer functions is an even-order polynomial, so the sum is also an even-order polynomial, and the deterministic change in the transfer functions will not affect the signal-to-noise ratio of the average. However, at higher frequencies, there will effectively be a slightly different weight applied to the data with different transfer functions because the expected value of the different transfer functions is different. In Ref. [37], the difference in weight was less than 0.5 % at 1 MHz. Even though this is a small effect, we still applied the cross-validation analyses described below to a subset of the total data using both $\Re[C_R]/\Re[C_Q]$ and its inverse, $\Re[C_Q]/\Re[C_R]$. The results in these two cases were similar, and, because the effective weighting has the opposite effect in these two cases, this confirms that the effect of the weighting is negligible.

In the cross-validation analysis, we consider candidate polynomial models with orders ranging from $d = 2$ to $d = 14$. We randomly split observed spectral data from 2025 subsets of experimental data into 45 equally sized subsets. Data from each run appear in just one of the 45 subsets. Each of these 45 subsets is then assigned to one of five subsets. From the five subsets, we construct training and validation data for a fivefold cross-validation study, where we select the model determined from training data that is most consistent with validation data according to a mean-square deviation criterion. Based on 20 000 random splits of the 45 subsets, we determine model selection fractions. Given that a d th order model is valid, and d is known, asymptotic theory predicts a sampling distribution for the estimate of a_0 . The standard deviation of this sampling distribution is the statistical uncertainty, σ_{stat} , of the estimate. The asymptotic distribution of parameter estimates determined by the method of ordinary least squares is well known and discussed in many textbooks, such as [58]. To account for the effect of imperfect knowledge of the model on results, we form a mixture of the sampling distributions from the candidate models weighted by their associated model selection fractions determined by cross-validation. We estimate the uncertainty of estimated a_0 as the total standard deviation of the mixture model distribution $\hat{\sigma}_{\text{tot}}$, where

$$\hat{\sigma}_{\text{tot}}^2 = \tilde{\sigma}_{\alpha}^2 + \tilde{\sigma}_{\beta}^2 \quad (15)$$

$$\tilde{\sigma}_{\alpha}^2 = \sum_d \hat{p}(d) \hat{\sigma}_{\hat{a}_0(d)}^2, \quad (16)$$

and

$$\tilde{\sigma}_{\beta}^2 = \sum_d \hat{p}(d) (\hat{a}_0(d) - \hat{a}_0)^2. \quad (17)$$

Above, $\hat{a}_0(d)$ is the estimate of a_0 associated with a d th order model, $\hat{\sigma}_{\hat{a}_0(d),\text{ran}}^2$ is the predicted variance of the estimate according to asymptotic theory, $\hat{p}(d)$ is the estimated model selection fraction for the d th order model, and $\hat{a}_0 = \sum_d \hat{a}_0(d) \hat{p}(d)$.

We stress that both $\tilde{\sigma}_{\alpha}$ and $\tilde{\sigma}_{\beta}$ are affected by imperfect knowledge of the ratio spectrum model. Instead of reporting $\tilde{\sigma}_{\alpha}$ and $\tilde{\sigma}_{\beta}$, we report $\hat{\sigma}_{\text{stat}}$, which is defined as the statistical uncertainty associated with the selected model, assuming the model is valid. Then, the additional uncertainty associated with model ambiguity, that is, imperfect knowledge of the polynomial order d of the ratio spectrum model, is defined as $\hat{\sigma}_{\text{model}} = \sqrt{\hat{\sigma}_{\text{tot}}^2 - \hat{\sigma}_{\text{stat}}^2}$, with the assumption $\sigma_{\text{tot}} \geq \sigma_{\text{stat}}$ [36].

We select f_{max} by minimizing $\hat{\sigma}_{\text{tot}}$ on a grid in frequency space from 20 kHz to 700 kHz with a resolution of 25 kHz. We restrict our search to f_{max} values no larger than 700 kHz to keep a sufficient distance from the corner frequency of the low-pass filters. Since the uncertainty estimates are realizations of random variables, the selected f_{max} is a realization of a random variable. Hence, following arguments in Refs. [36] and [40], we determine an additional component of uncertainty $\hat{\sigma}_{f_{\text{max}}}$ that accounts for uncertainty associated with imperfect performance of our selection method due to random effects, as well as possible systematic effects, including frequency-dependent physical effects. We set this component to the estimated standard deviation of the estimates that correspond to fitting bandwidths that yield the five lowest values of $\hat{\sigma}_{\text{tot}}$ for f_{max} no larger than 700 kHz.

We emphasize that, in our cross-validation analysis, we select the order of the ratio spectrum model (where d ranges from 2 to 14) and determine σ_{tot} (Eq. 15) for each of many candidate values of f_{max} on a grid ranging from 200 kHz to 700 kHz. We select the f_{max} that yields the lowest value of σ_{tot} on this grid. In Table 3, we report the statistical uncertainty corresponding to the selected model at the selected fitting bandwidth, and components of uncertainty due to model ambiguity and bandwidth ambiguity. We stress that the selected values of f_{max} , d , and the above-mentioned components of uncertainty are determined by data-driven empirical methods. That is, the fitting bandwidth selection approach is not based on a specific circuit model of frequency dependence. A rigorous study of the possible dependence of the selected fitting bandwidth, selected model complexity, and reported components of uncertainty on choices of various experimental parameters is worthwhile, but beyond the scope of this paper.

In Ref. [37], the constrained grid search yields a minimum $\hat{\sigma}_{\text{tot}} = 4.55 \times 10^{-6}$ at $f_{\text{max}} = 350$ kHz and $d = 4$, with $\hat{\sigma}_{\text{stat}} = 4.25 \times 10^{-6}$ and $\hat{\sigma}_{\text{model}} = 1.63 \times 10^{-6}$; the additional component of uncertainty is $\hat{\sigma}_{f_{\text{max}}} = 1.31 \times 10^{-6}$. The estimate of the Boltzmann constant from the ratio of the power spectral densities S_R/S_Q is $k = 1.380\,642\,9 \times 10^{-23}$ J/K, with a relative offset of -4.05×10^{-6} from the CODATA 2014 recommended value and a combined relative standard uncertainty of 4.97×10^{-6} due to the combined effect of random measurement errors, model selection ambiguity, and fitting bandwidth ambiguity.

5.2 Estimation of Statistical Uncertainty

To better understand our measurement uncertainties, we can make a simple estimate of the expected statistical uncertainty similar to that in Ref. [34]. We start from the Dicke radiometer equation [14], expressed using mean-squared voltage $\langle V^2 \rangle \propto P$ instead of power,

$$\sqrt{\text{Var}[V^2]}/\langle V^2 \rangle = 1/\sqrt{B \tau}, \quad (18)$$

which relates the relative uncertainty in the measured mean-squared voltage noise to the bandwidth B of the measurement and the measurement time τ . This equation applies directly to the autocorrelation measurements that measure the total noise V^2 , which is the sum of the amplifier noise V_a^2 and the sense resistor noise $\langle V_R^2 \rangle \propto RT$, using the temperature T and resistance R of the sense resistor:

$$\sqrt{B \tau \text{Var}[V^2]}/\langle V^2 \rangle = 1 + \langle V_a^2 \rangle/\langle V_R^2 \rangle = 1 + \Gamma_a/(RT), \quad (19)$$

where we express the amplifier noise as Γ_a in units of ohms-kelvin. In our case, $\Gamma_a \approx 27\,000 \Omega \text{ K}$; that is, the amplifier noise is about half the $R = 200 \Omega$ sense resistor Johnson noise at the TPW.

This estimate for an autocorrelation measurement can be made more relevant to the JNT system by accounting for three effects: first, that we are making a ratio measurement; second, that we are measuring the two parts of the ratio sequentially instead of simultaneously; and third, that we are making a cross-correlation measurement.

The ratio measurement has two effects. First, the ratio P_R/P_Q has a relative variance that is the sum of the relative variances of $P_R \propto V_R^2$ and $P_Q \propto V_Q^2$, where the variable P denotes the power spectral density of the autocorrelation. Because the system is designed so that the magnitude of the QVNS synthesized signal is close to the magnitude of the thermal noise, the relative variances are approximately equal, and the uncertainty in the ratio is about $\sqrt{2}$ times the uncertainty in the measurement of a single source. Second, the sequential nature of the measurement means that the measurement time per source is only $\tau/2$, where τ is the total time spent collecting data. This results in another $\sqrt{2}$ increase in the uncertainty, based on the definition of τ :

$$\sqrt{B \tau \text{Var}[P_R/P_Q]}/(\langle P_R \rangle/\langle P_Q \rangle) = 2\sqrt{B \tau \text{Var}[V^2]}/\langle V_R^2 \rangle = 2 + 2 \Gamma_a/(RT). \quad (20)$$

Finally, the cross-correlation measurement reduces the effect of amplifier noise. To determine the magnitude of reduction, we calculate the variance in the product of the FFTs of the measured Johnson noise voltages in two cases: the autocorrelation measurement for a single channel, A or B; and the cross-correlation measurement between channels A and B. Again, the voltages are composed of two parts: thermal noise and amplifier noise. We assume that the thermal noise and amplifier noise signals are uncorrelated and Gaussian, with $E[V] = 0$, $E[V^2] = \sigma_V^2$, and $E[V^4] = 3\sigma_V^4$ for the thermal voltage noise V_R with uncertainty σ_{V_R} and amplifier noise voltages on channels A and B with equal uncertainties σ_{V_a} . We also assume that the two amplifier noises are uncorrelated when computing the variance of the cross-correlation. Using these simplifications, the autocorrelation measurements have a variance

$$\text{Var}[V_A^2] = \text{Var}[V_B^2] = 2(\sigma_{V_R}^2 + \sigma_{V_a}^2)^2 = 2\sigma_V^4, \quad (21)$$

and the cross-correlation measurement has a variance

$$\text{Var}[V_A V_B] = 2(\sigma_{V_R}^2 + \sigma_{V_a}^2)^2 - \sigma_{V_a}^2(2\sigma_{V_R}^2 + \sigma_{V_a}^2). \quad (22)$$

We therefore estimate the relative uncertainty in the cross-correlation measurement by scaling Eq. (20) by the ratio of the variances in Eq. (21) and Eq. (22):

$$\sqrt{B \tau \text{Var}[C_R/C_Q]}/(\langle C_R \rangle/\langle C_Q \rangle) = \sqrt{B \tau \text{Var}[P_R/P_Q]}/(\langle P_R \rangle/\langle P_Q \rangle) \times \sqrt{\text{Var}[V_A V_B]}/\sqrt{\text{Var}[V_A^2]}, \quad (23)$$

using the ratios

$$\sigma_{V_R}^2/\sigma_V^2 = RT/(\Gamma_a + RT) = 1/(1 + \gamma) \text{ and } \sigma_{V_a}^2/\sigma_V^2 = \Gamma_a/(\Gamma_a + RT) = \gamma/(1 + \gamma), \quad (24)$$

where $\gamma = \Gamma_a/(RT)$. Here, $\gamma \sim 0.49$ for our sense resistor and amplifiers. The relative uncertainty is approximately

$$\sqrt{\text{Var}[C_R/C_Q]} / (\langle C_R \rangle / \langle C_Q \rangle) \sim \sqrt{4 + 4\gamma + 2\gamma^2} / \sqrt{B\tau}. \tag{25}$$

However, the use of the Dicke radiometer equation in Eq. (18), which leads to Eq. (25), assumes that the ratio data $\Re[\langle C_R \rangle / \langle C_Q \rangle]$ are frequency-independent, so that it is appropriate to simply average over the bandwidth B ; that is, it assumes that the measurement channels are sufficiently matched $H_{A,R} H_{B,R}^* = H_{A,Q} H_{B,Q}^*$ with minimal undesired correlated noise $C_{n,R} = 0$. Since these assumptions represent the ideal JNT system, Eq. (25) is a lower bound on the statistical uncertainty. If the ratio data require a fit to a higher-order polynomial $d > 0$, then we would expect a larger uncertainty.

As an aside, Eq. (25) can be used to quantify how much of our statistical uncertainty is associated with amplifier noise, as opposed to the inherent randomness of the Johnson noise signal. In our experiment, $\gamma = \Gamma_a / (RT) \approx 0.49$, and the numerator of Eq. (25) is $\sqrt{4 + 4\gamma + 2\gamma^2} \approx 2.54$, while in an experiment with ideal noiseless amplifiers, $\gamma = 0$, and $\sqrt{4 + 4\gamma + 2\gamma^2} = 2$. Therefore, we predict that $0.54/2.54 \approx 21\%$ of our estimated statistical uncertainty is caused by amplifier noise. In other words, further amplifier optimization could reduce the statistical uncertainty of this experiment by a maximum of 21%; realistic improvements would probably offer uncertainty reduction much less than 21%. The remaining 79% of the expected statistical uncertainty could only be reduced by increasing bandwidth, increasing integration time, or improving other experimental factors not considered by this simple uncertainty estimation.

Applying the analysis of this section and Eq. (25) to our recent 50 d measurement [37], and assuming that the bandwidth of the system is equal to the alias filter bandwidth of 800 kHz, we estimate a lower bound to the relative statistical uncertainty of about 1.5×10^{-6} . However, the cross-validation method indicates an optimal bandwidth of only 350 kHz and a higher even-order polynomial model with $d = 4$. This reduced bandwidth would imply a lower bound of about 2.2×10^{-6} . Furthermore, since a $d = 4$ model was required, we realized only a relative statistical uncertainty of 4.25×10^{-6} in Ref. [37]. The statistical uncertainties of a few recent measurements are compared in Table 2.

Table 2. Statistical uncertainty, integration time, and characteristics of the even-order polynomial fit used to determine the statistical uncertainty, for a number of different Boltzmann measurements.

| Measurement | Statistical uncertainty | Integration time | Polynomial fit order | Polynomial fit bandwidth |
|----------------|------------------------------------|------------------|----------------------|--------------------------|
| NIST 2011 [34] | 11.6×10^{-6} ^a | 10 d | 2 | 10 kHz to 650 kHz |
| NIM 2015 [35] | 3.2×10^{-6} | 33 d | 4 | 1.8 kHz to 575 kHz |
| NIM 2017 [36] | 2.37×10^{-6} | 100 d | 2 | 10 kHz to 369 kHz |
| NIST 2017 [37] | 4.25×10^{-6} | 50 d | 4 | 5 kHz to 350 kHz |
| NMIJ 2017 [39] | 9.85×10^{-6} | 5 d | 2 | 40 kHz to 380 kHz |

^aUnlike other measurements, the NIST 2011 [34] measurement included a dominant “spectral aberrations” term of 10.4×10^{-6} in the uncertainty budget along with a “statistics” term of 5.2×10^{-6} ; more recent measurements do not use the “spectral aberrations” term (see Sec. 5.5 below). In this table, we have combined the two terms and assigned a statistical uncertainty of 11.6×10^{-6} to the [34] measurement.

5.3 Electromagnetic Interference

We look for the effects of EMI in two different ways. First, we use a modified resistance probe where the configuration of resistors is the same as in the QVNS probe (see Fig. 1), that is, so that each measurement channel sees different resistors. In this configuration, only common-mode EMI will show up in the cross-correlation spectrum. This is the “null-measurement” mode, and it is necessary to uncover possible EMI that would ordinarily be hidden under the white noise spectrum of the sense resistor [45, 59]. However, this type of EMI measurement requires a long averaging time and cannot be performed at the same time as the typical JNT measurement.

Therefore, we also use a second technique where we average the complex QVNS cross-correlation spectrum and look at the frequencies between QVNS tones; this uses the same data analysis techniques as we used in Sec. 3.2 to measure the cryogenic sense resistor and find undesired correlated noise, but it is applied to the typical QVNS probe, which has different resistors on the two measurement channels. This is a sensitive way of looking for EMI that is picked up by the amplifier chain, relay board, or the QVNS itself.

However, this is not sensitive to any EMI that is coupled only to the R-probe (for example, the R-probe may be more sensitive to noise from the TPW maintenance system).

As in the previous section, we can estimate the expected uncertainty in both EMI measurement techniques. The calculation is the same until the estimation of the cross-validation variance. In the EMI measurement, there are two separate sense resistors with uncorrelated Johnson noise as well as uncorrelated amplifier noise. Therefore, V_A and V_B are also uncorrelated, and the variance of the cross-correlation is

$$\text{Var}[V_A V_B] = \text{Var}[V_A] \text{Var}[V_B] = (\sigma_{V_R}^2 + \sigma_{V_a}^2)^2. \quad (26)$$

Using this variance as in the previous sections, the relative uncertainty in the EMI ratio measurement is approximately

$$\sqrt{\text{Var}[C_R/C_Q]} / (C_R/C_Q) \sim \sqrt{2}(1 + \gamma) / \sqrt{B\tau}. \quad (27)$$

In the case of minimal EMI, the ratio $C_R/C_Q = 0$, independent of frequency. When using the first technique, that is, a modified resistance probe, Eq. (27) can be applied directly. In the case of the second technique, there is an additional scale factor because we are measuring 4 K resistors, but we are interested in the relative contribution to a measurement at the TPW. This means that γ in Eq. (27) should be evaluated at $T_{\text{cryo}} = 4$ K, but the relative uncertainty is scaled by a factor of $T_{\text{cryo}}/T_{\text{TPW}}$. Since the sense resistor noise at cryogenic temperatures is much less than the amplifier noise, Eq. (27) can be simplified to

$$\sqrt{\text{Var}[C_R/C_Q]} / (C_R/C_Q) \sim \sqrt{2} \times \Gamma_a / (R T_{\text{TPW}} \sqrt{B\tau}) \quad (28)$$

because, as defined in Sec. 5.2, the amplifier noise temperature is approximately equal to the TPW.

During our recent measurement of the Boltzmann constant [37], we interrupted the main measurement for 10 h to measure EMI using a modified resistance probe. From Eq. (28) and a measurement bandwidth that matches the f_{max} chosen by the cross-validation method $B = 350$ kHz, we would expect an uncertainty of 19×10^{-6} . The real and imaginary parts of the complex ratio C_R/C_Q for this EMI measurement are shown in Fig. 14, and we obtain a relative offset of $-2/+8 \times 10^{-6}$ with a relative uncertainty of $14/13 \times 10^{-6}$ in the real and imaginary parts of the ratio, respectively. These offsets are consistent with no EMI, but to reduce the uncertainty in the measurement to 1×10^{-6} , we would have to average for approximately two orders of magnitude longer, that is, about 50 d.

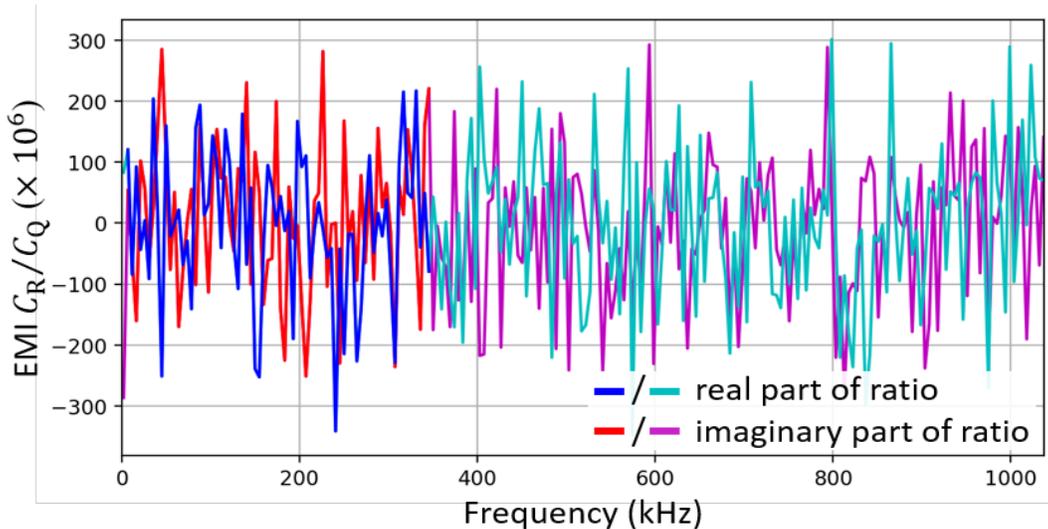


Fig. 14. Measured EMI using modified resistance probe averaged over 10 h.

The long averaging time required when using the modified resistance probe motivated the use of the second technique and the QVNS data. In Fig. 15, we plot the results of the EMI test using the QVNS measurement over the entire 50 d of collected data. From Eq. (25) and a measurement bandwidth that matches the f_{\max} chosen by our uncertainty minimization method (Sec. 5.1), $B = 350$ kHz, we would expect an uncertainty of 0.6×10^{-6} ; we observe a relative offset of $-1.3/+1.3 \times 10^{-6}$ with a relative uncertainty of $0.5/0.4 \times 10^{-6}$ in the real and imaginary parts of the ratio, respectively.

From these measurements, we assign an uncertainty of 1×10^{-6} to the EMI contribution. However, in the lower noise QVNS data (Fig. 15), we observe a structure with a magnitude of $\sim 15 \times 10^{-6}$ near 700 kHz. This structure is outside of our current measurement bandwidth, but future experiments will need to deal with this possibility more carefully.

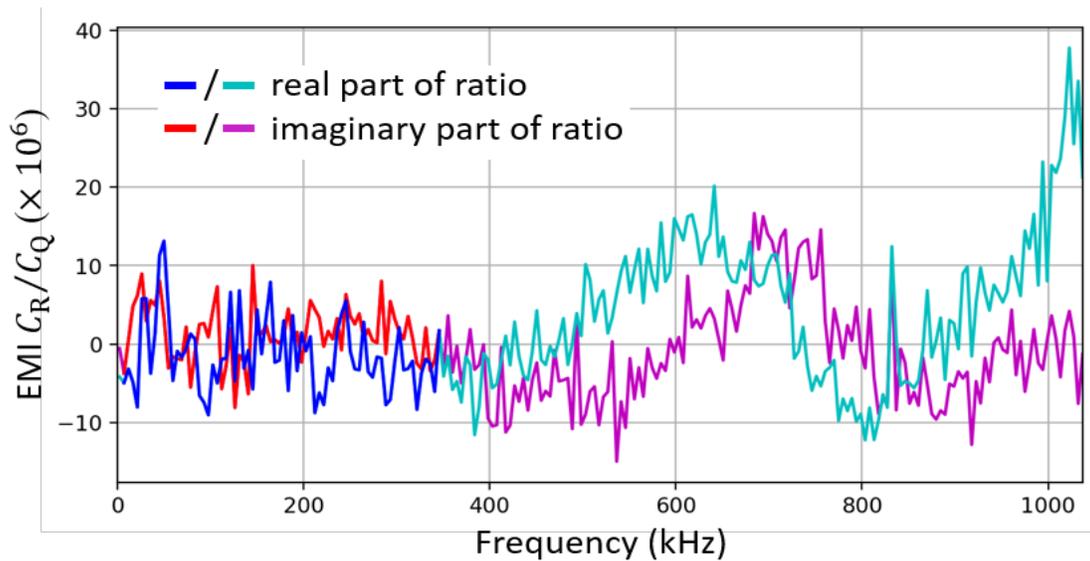


Fig. 15. Measured EMI using QVNS probe noise averaged over 50 d.

5.4 Nonlinearity

Two differences between the QVNS voltage and the Johnson noise voltage, in combination with measurement nonlinearities, can create errors in the results of the k measurement: differences in the dc voltage and differences in the ac magnitude of the signals [60]. The QVNS pattern includes an inherent dc voltage offset as a side effect of the properties of the pulse generator and the zero-compensation technique, which force the QVNS output to be either entirely positive or entirely negative. For comparison, the average dc voltage of the Johnson noise is zero. This difference between the dc voltages of the two sources has negligible effect on the JNT measurement because the amplifiers are ac-coupled after the first input stage, and thus the digitizer is ac-coupled. In general, the amplifiers are more linear than the digitizers.

Any differences in the magnitudes of the QVNS waveform and the Johnson noise can also create measurement errors through the nonlinearities of both the amplifiers and digitizers. We therefore attempt to match the magnitude of the QVNS to the magnitude of the Johnson noise within the bandwidth of the ADC and anti-alias filter. This matching can take a number of forms, including matching the magnitude of the average autocorrelation of each channel and matching the magnitude of the cross-correlation. In past experiments, additional small resistors were added to the various channels of the various sources to improve this matching.

In our recent measurement [37], we matched the dc resistance of all the channels. We then chose a QVNS pattern that matched the magnitude of the QVNS cross-correlation to that of the sense resistor. However, this match was not exact and did not explicitly match the magnitude of the autocorrelation measurements, so we tested the sensitivity of the system by scaling the magnitude of the QVNS pattern by

a factor α and extracting the relative change in the magnitude of the measured QVNS peaks β ; this quantity was also divided by the measured sense resistor noise to remove the effect of amplifier gain drift:

$$\beta(\alpha) = C_Q / (|v_{JJ}(\alpha)|^2 C_R). \tag{29}$$

Because we did not change the system parameters during this measurement, the relative change in the peak magnitudes $\beta(\alpha)/\beta(1)$ is sensitive only to nonlinear effects and is not sensitive to the linear transfer functions of the system.

In Fig. 16, we plot $\beta(\alpha)/\beta(1)$ for $\alpha = 0.92$ and $\alpha = 1.08$ averaged over frequencies between 5 kHz to 850 kHz. Assuming a linear dependence for small changes in QVNS magnitude, we observe a slope of 0.33×10^{-6} change in the measured QVNS tones per percent change in the QVNS magnitude with an uncertainty of 0.22×10^{-6} . We also plot the ratio of the total power in the QVNS autocorrelation measurements relative to the source resistor measurements as a function of α . We see that a perfect match in total power would result in a shift in α of 1 %, which would, based on the measured nonlinearity, change M_{JJ} by $< 0.5 \times 10^{-6}$. Thus, we conservatively assign an uncertainty due to nonlinearity of 0.5×10^{-6} .

In the future, new pulse generators will let us drastically decrease the tone spacing to < 10 Hz by increasing the length of the pattern. In that case, a better nonlinearity test would include changing the magnitude and frequency spacing of the tones while keeping the power in the comb constant, as well as changing the relative phase of the tones.

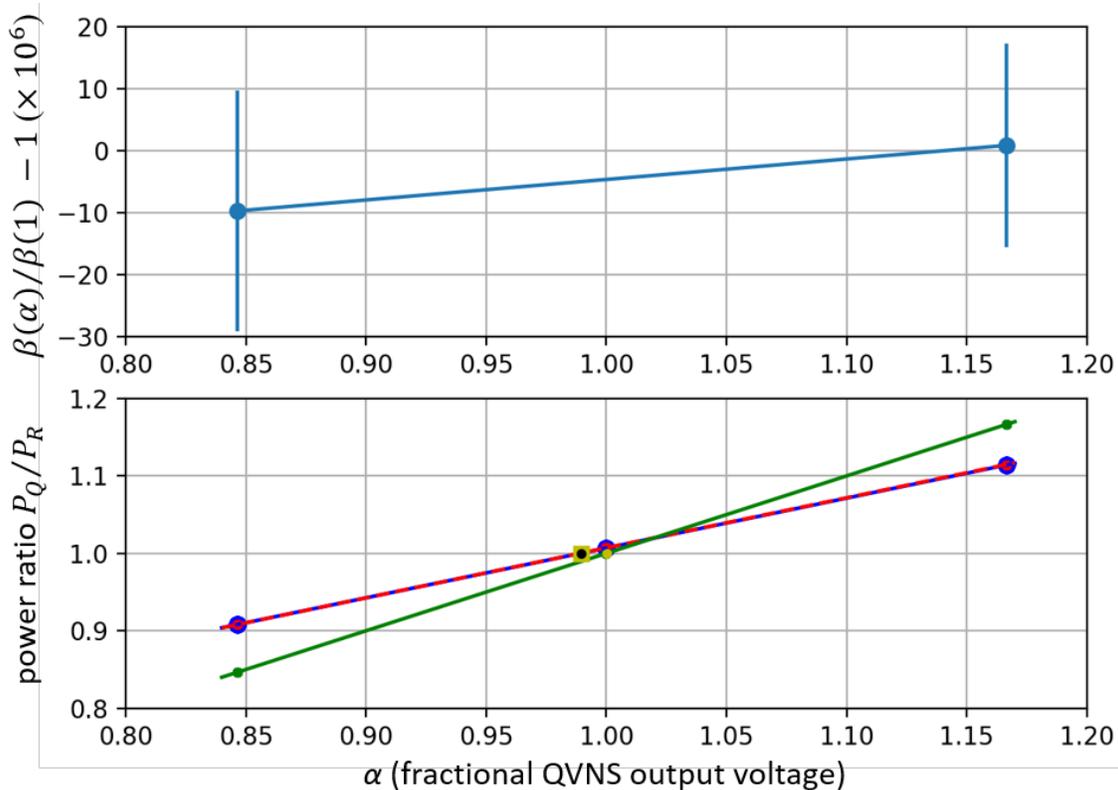


Fig. 16. (Top) For QVNS patterns where the calculated magnitude differs from the typical magnitude (x -axis), we plot the measured QVNS magnitude normalized by both the calculated magnitude for that pattern and the measured magnitude of the typical pattern. System nonlinearities would offset the normalized magnitude from 1. (Bottom) We plot the ratio of the measured power from the QVNS source and the Johnson noise source (y -axis) for the channel A autocorrelation (blue dots and solid blue line), channel B autocorrelation (red dots and dashed red line), and cross-correlation (green dots and solid green line). We highlight the QVNS source magnitude at which the QVNS and Johnson noise source powers would match with a yellow square (black dot) for the channel A (B) autocorrelation and a yellow dot for the cross-correlation.

5.5 Spectral Aberrations

In our previous publications [34], an uncertainty component, referred to as “spectral aberrations,” was included to account for sources of frequency dependence that could not be modelled by the simple second-order filter-response functions $a_0 + a_2 f^2$. The scope of this uncertainty category has narrowed in the last few years as more sophisticated statistical analysis has become available [40]. Prior to the development of the spectral model selection methods described in Sec. 4.2, the simple statistical uncertainty estimates used for JNT depended on the combined effects of random measurement errors and spectral anomalies, and did not distinguish the two. This had the effect of introducing strong model and bandwidth dependence in the results, which had to be accounted for in a less rigorous manner. In our new statistical analysis, we quantify two components of uncertainty produced by spectral aberrations; one component accounts for model ambiguity and the other accounts for bandwidth ambiguity. In general, we expect our new statistical analysis methods to yield a better understanding of spectral aberration effects in similar JNT experiments. Consequently, the spectral aberrations category has been removed from this and other [36] uncertainty budgets for contemporary JNT results. We have retained the entry in Table 3 only in reference to the previous work [34].

5.6 Dielectric Loss

The potentially most significant source of low-frequency aberrations has been and remains stray shunt capacitance that has appreciable loss, expressed as the frequency-dependent loss tangent $\tan(\varphi(f))$. For any such complex shunt capacitance $C_S \approx C_S'(1 - j \tan(\varphi(f)))$ coupled to a resistance R , there will be two quasi-linear frequency-dependent error terms. We refer to this type of dependence as “quasi-linear” since the extra terms will enter into the expansions for H_R and H_{JJ} as terms proportional to $RCf \tan(\varphi(f))$, where the frequency dependence of $\tan(\varphi(f))$ is usually weak. By superposition, one term will be the lossy capacitance filtering of the Johnson noise from R , and the other term will be the thermal noise from the dissipative part of the capacitance as attenuated by a noiseless R . When the system is connected to the QVNS, the array will decouple the two channels such that the capacitance noise will be uncorrelated and average to zero. When coupled to the R-probe, however, the capacitance noise will be correlated and remain in the correlated noise spectra.

The possible existence of small quasi-linear frequency dependence in the noise spectra presents a problem from a statistical standpoint. It is not useful to try to extract a small linear term a_1 as a fitting parameter in the presence of competing random noise. First, such attempts can easily result in fitted values for a linear coefficient that are simply unphysical in magnitude and/or sign when compared to the realistic upper limits set by equivalent circuit models and knowledge of the characteristics of the input circuit impedances. Second, there will be a significant covariance between a_0 and a_1 , which will have the effect of expanding the uncertainty in a_0 much more than it would by simply adding another fitting parameter of even order. Hence, *ab initio* estimates using lumped-parameter equivalent circuit models are necessary to estimate the uncertainty due to these particular aberrations.

We consider two specific potential sources: (1) the lossy capacitance C_{FR4} associated with the FR4 fiberglass composite [61] that makes up our printed circuit switchboard (coupling to both R- and Q-probes); and (2) the small stray capacitance C_{LCC} associated with certain lossy ceramic materials [62] used in the construction of the LCC package inside of the R-probe (which only couples to the R-probe resistors). We use a simple equivalent circuit representation to predict an *ab initio* value for an effective quasi-linear error term a_{1-Est} approximated by

$$a_{1-Est} \approx 2\pi \left\{ [2R_Q - (2 - T_{FR4}/T)2R_T] C_{FR4} \tan(\varphi_{FR4}(f)) - (2 - T_{LCC}/T)2R_T C_{LCC} \tan(\varphi_{LCC}(f)) \right\}. \quad (30)$$

There are five distinct terms in Eq. (30). The first two terms are essentially filtering effects due to the lossy FR4 capacitance coupling to either the QVNS terminating resistors R_Q or the R-probe resistors R_T . The next term proportional to the temperature $T_{FR4} \approx 290$ K is the correlated capacitance noise from the FR4 dielectric coupled to the R-probe resistance R_T . The fourth term is another filtering term specific to the LCC insulator capacitances, and the final term is the correlated noise from those capacitances coupled to

the R-probe resistors for which $T_{LCC} = T = 273.16$ K. In this work, we constrain the system with an impedance matching condition such that $R_Q = R_T \equiv R = 100 \Omega$, which is the single-ended resistance as shown in Fig. 1. In this special case, there is some significant cancellation in Eq. (30), which then simplifies to

$$a_{1-\text{Est}} \approx 4\pi R \{0.062 C_{FR4} \tan(\varphi_{FR4}(f)) - C_{LCC} \tan(\varphi_{LCC}(f))\}. \quad (31)$$

We estimate $C_{FR4} \tan(\varphi_{FR4}) \approx 0.12(4)$ pF and $C_{LCC} \tan(\varphi_{LCC}) \approx 0.004(1)$ pF, which yields $a_{1-\text{Est}} \approx 4.3 \times 10^{-9} \text{ kHz}^{-1}$. This is approximately a factor of 0.53 times our estimate for $a_{1-\text{Est}}$ made in our previous work [34]. In that work, we separated the empirical spectral aberration uncertainty from the strictly *ab initio* dielectric loss uncertainty (then listed as a ‘‘Correction Model’’ uncertainty), and the latter estimate was 2×10^{-6} .

Based on these simple assumptions and rough calculations, we estimate an uncertainty due to spectral aberrations to be 1×10^{-6} . It should also be noted that this smaller estimated uncertainty has been achieved in part by constraining the system to be impedance matched first, and then tuning the time constants separately with lossless trim capacitors. As discussed in Sec. 3, this is different from the $R_Q = 2R_T$ matching strategy employed in our previous work.

5.7 QVNS Waveform Synthesis

The QVNS produces a waveform using quantized pulses produced by Josephson junctions [29, 41, 42]. Therefore, at frequencies much less than the inverse width of the quantized pulse, the waveform is exactly proportional to the Josephson constant K_J and the pulse generator clock. The width of the JJ pulses is less than 50 ps (corresponding to the 20 GHz characteristic frequency of the JJs), and since the 2 MHz frequency of interest is four orders of magnitude smaller than the characteristic frequency, the effect of the JJ pulse width is negligible. The relative uncertainty in K_J , which is 6.1×10^{-9} [16], is also negligible compared to the statistical uncertainty in this measurement. Finally, we use a Stanford Research Systems FS725 rubidium frequency standard to provide a 10 MHz clock for the JNT system. This standard has been compared to a NIST maser frequency standard and agrees to better than 0.01×10^{-6} , which makes the clock a negligible source of error.

5.8 Temperature

All TPW cells are subject to uncertainties of both thermophysical and chemical origin. The TPW working cell used here is subject to additional uncertainties due to the use of the dry-well maintenance system.

5.8.1 Chemical Impurities

Chemical impurities, both dissolved gases and metals, may exist in the liquid phase and would depress the melting point temperature. Most modern TPW cells can be produced where the initial impurity levels are low enough to be almost undetectable by temperature alone. Over time, these impurities may slowly increase as the cell ages, so occasional comparisons with newer cells and or cells made from inert quartz glass can provide upper limits on the extent of the chemical aging [63]. The cell comparisons described in Sec. 4.4 suggest that an uncertainty of 0.05 mK, or 0.18 $\mu\text{K/K}$, is sufficient to account for the combined effects of air and dissolved metals in the water of the working cell.

5.8.2 Pressure Head Correction

The pressure from the column of water, or hydrostatic head, over the effective immersion depth will lower the temperature of the sense resistor as it equilibrates with the nearby region of solid-liquid interface that exists at slightly elevated pressure. Based on the dimensions of the thermowell and the heat sink within the JT4B probe, we estimate an effective immersion depth of 24 cm. We then apply a correction of

-0.0073 mK/cm [64], which yields a correction of -0.18 mK for a realization temperature of $273.159\,82$ K. We assign an uncertainty of 10 % for the head correction coefficient plus 2 cm uncertainty in effective depth, which yields an uncertainty of 0.023 mK, or 0.084 μ K/K.

5.8.3 Immersion Errors

Immersion errors can occur for a temperature probe in any immersion-type fixed point cell when either the axial heat flux within the probe is too large, or the radial heat transfer away from the probe is too low. The ratio of these two heat-exchange rates determines a characteristic length d_{im} , which in turn governs the degree of equilibrium achieved with increasing immersion depth z via an exponential characteristic $\sim \exp(z/d_{\text{im}})$ [65]. Once sufficient immersion has been achieved, it is possible to obtain the linear pressure-induced effect with increasing immersion depth as described in Sec. 4.3 and Sec. 5.7.2. A well-designed probe and cell combination may achieve this condition if the wall temperature of the thermowell is also sufficiently uniform, as is usually achieved when there is a full length of solid ice mantle to provide a liquid-solid interface near the exterior thermowell wall. In this condition, the immersion error is considered negligible, but departures from this condition imply that a finite immersion error will exist, and this error will depend on ambient or otherwise external temperatures.

Figure 12 illustrates the degree of immersion achievable using our JT4B probe in the working cell with a full-length ice mantle. Under these “best-case scenario” conditions, we would assign an uncertainty of 0.05 mK due to any remaining immersion error, or 0.18 μ K/K. Some loss of effective cooling in the dry-well cooler maintenance system created accelerated melting rates for some of the absolute measurement runs in May. In those less than ideal cases, partial melting of the ice mantle will produce larger immersion errors, and the uncertainty must be expanded up to as high as 0.3 mK (1.1 μ K/K) for run index numbers between 61 and 80. The uncertainty of 0.65 μ K/K as shown in Table 3 is derived from a pooled variance of variable immersion error uncertainties between these two limits.

5.8.4 Isotopic Variations

The *mise en pratique* for the kelvin [66] specifies that the water inside of a TPW cell must be of a certain isotopic composition, equivalent to Vienna Standard Mean Ocean Water (VSMOW) [67]. The water inside the working cell used in this work does not have any certification or other equivalent traceability to establish this composition. However, it has been compared with at least one other cell that does have that certified isotopic traceability. As shown in Table 1, our November 2015 comparison of the working cell temperature differs with that of the quartz glass transfer cell Q1454 by $-0.061(27)$ mK. The cell Q1454 was a new cell in 2014 that had recently been compared to other NIST quartz glass cells with traceable isotopic certifications. The certificate for Q1454 states that the isotopic variation for that cell will correspond to temperature variations within ± 0.02 mK. However, the apparent relative depression in the working cell is most likely due to a combination of both chemical and isotopic effects, both of which will normally lower the temperature. Based on these results, and given that we have already assigned a standard uncertainty of 0.05 mK for the effects of chemical impurities in the working cell, we assign a standard uncertainty component of 0.03 mK (0.1 μ K/K) due to an unknown isotopic variation in the working cell.

5.8.5 Thermal Equilibrium

Ideally, the sense resistors within the LCC package should equilibrate with the solid-liquid interface of the water within the TPW cell. A series of thermal impedances exist over various other interfacial surfaces between those regions and create thermal time constants that govern the approach to equilibrium. We have used the diagnostic thermistor inside of the R-probe to study transient behavior and departures from equilibrium. In the absence of any external heat flux, equilibration takes place within 600 s following minor perturbations associated with adjustments in the R-probe or the cell from their normal positions. However, some small static stray heat flux through portions of the normal thermal path can remain. In practice, any such departure from ideal equilibrium would be indistinguishable from an immersion error. Our lower limit

of 0.05 mK for the minimum possible remaining immersion error as already given above is sufficient to account for these ordinary heat leaks.

The R-probe is wetted by alcohol contained within the thermowell of the TPW cell, which allows for effective heat transfer between those surfaces. In contrast, the TPW cell itself is by design installed inside of a so-called “dry-well” of the thermoelectrically cooled maintenance system. In this case, only air gaps of variable dimensions provide the heat transfer to the external glass surfaces, and the larger thermal impedance can result in larger thermal gradients and stray heat flux across portions of the cell’s interior. The influence of these gradients on the as-realized equilibrium temperature will depend on the condition of the ice mantle as well as on the system controller set point. We assign an additional uncertainty of 0.05 mK or 0.18×10^{-6} to account for this cell-dry well-system interaction. This contribution remains unchanged from our 2011 uncertainty budget [34].

5.9 Resistance

The goal of the resistance measurements is to establish values for the R-probe resistance in units of Ω_{90} , which are traceable to realization of the QHE. The calibration of the 1202T standard achieves this traceability with calibration uncertainties that are below $0.01 \mu\Omega/\Omega$ [46], but our comparisons of the R-probe resistance to the standard are not capable of approaching that level of uncertainty. Here, we describe the limiting sources of uncertainty in both resistance measurements and in the underlying assumptions behind the methods.

5.9.1 Bridge Statistics

Both the DCC and the DSB are capable of measuring resistance ratios using signals of 50 mV with statistical uncertainties of $0.1 \mu\Omega/\Omega$, with only a few exceptions. Those exceptions occurred when the DCC was used in a quick-measurement mode, but this did not impact most of the R-probe resistance data. A 50 mV signal was produced by biasing the 200Ω probe resistance at $250 \mu\text{A}$, which was shown to cause negligible probe heating in a test where resistance was measured at multiple currents. In principle, both bridges are also subject to ratio-dependent systematic errors. Our experience is that these errors are much smaller than the statistics when using these bridges with the modest signal levels of 50 mV, and so we assign $0.1 \mu\Omega/\Omega$ uncertainty for all of the bridge-specific ratio measurements.

5.9.2 Standard Resistor Calibration Stability

The 1202T standard resistor has a finite stability with time and is known to exhibit a small positive drift rate. It has been hand-carried between NIST’s Gaithersburg and Boulder facilities on several occasions with no indication of transit-induced shifts. Resistance standards of the type (SR102) are also known to improve with age in the sense that the initial drift rate will slowly diminish. The most recent calibration was in August of 2015 via a 100Ω resistor bank of five 100Ω standards that are maintained at the NIST Quantum Measurement Division’s facilities in Gaithersburg. That bank of standards is regularly compared against the QHE Ohm and is also capable of achieving uncertainties below $0.01 \mu\Omega/\Omega$ [68]. At that time, the apparent drift rate of 1202T was estimated to be $0.02 \mu\Omega/\Omega$ per year. At approximately 1.67 years later, on April 20, 2017, this would amount to an incremental increase of only $0.033 \mu\Omega/\Omega$. Even if this drift rate were uncertain at 30 %, the impact on the extrapolated calibration uncertainty would still be only $0.01 \mu\Omega/\Omega$, which is negligible.

Some additional checks were performed in Boulder in November of 2015 by comparing the 1202T standard with a recently calibrated 200Ω reference resistor, and those checks were consistent at the $0.1 \mu\Omega/\Omega$ level using the DCC bridge. In the absence of any other independent information, we assign an uncertainty of $0.1 \mu\Omega/\Omega$ to account for a possible unknown instability in the transfer standard since its last calibration.

5.9.3 Standard Resistor Temperature

As already discussed in Sec. 4.5, the 1202T standard was subject to corrections as large as $1.1 \mu\Omega/\Omega$ due to the fact it was used at temperatures as much as 6.6°C colder than its normal 23.0°C calibration temperature. For data obtained prior to 19 April 2017, this produced relatively large uncertainties due to the large (i.e., 1 K) uncertainties in the actual temperature under use. Once the calibrated SPRT was installed inside the 1202T thermowell on 19 April, this uncertainty became negligible, since accurate corrections to the standard resistance could be determined from the simultaneously measured temperature. In Fig. 13, these additional uncertainties are reflected in the larger error bars for most of the data as shown prior to 19 April. These additional uncertainty components were variable, depending on the ambient temperature in the room during each measurement, but they were typically about equal to the statistical uncertainties, or $\sim 0.1 \mu\Omega/\Omega$.

5.9.4 Sense Resistor Relaxation and Stability

As already discussed in Sec. 4.4, the sense resistors are subject to thermally induced stress relaxation effects. These relaxation effects are to a certain degree predictable, since the relaxation exhibits an exponential relaxation in time. The instabilities produce some additional uncertainties, which vary from a lower limit of $0.031 \mu\Omega/\Omega$ to as large as $0.38 \mu\Omega/\Omega$ in the special case of data taken prior to 29 March (see Fig. 13). The uncertainty of $0.11 \mu\text{K}/\text{K}$ as shown in Table 3 is derived from a pooled variance of weighted uncertainties.

5.9.5 Frequency Dependence

A basic underlying assumption in spectral ratio noise thermometry is that a low-frequency measurement of the sense resistance is sufficient to predict the effective dissipation (and hence the fluctuations) at much higher frequencies. This is a valid assumption if whatever frequency dependence happens to be exhibited by the sense resistor is indistinguishable from frequency dependence that is already incorporated into the model. In the lumped-parameter approximation, any simple impedance-based time constants such as RC or $(LC)^{-1/2}$ that are specific to the sense resistor will effectively attenuate or filter the Johnson noise in the same way that the transmission line does, and so there is no practical way to distinguish the two effects.

Given a sense resistor with a frequency-dependent impedance $Z(f) = R_0 + jX$, the dissipative component that produces the fluctuations is $\Re[Z(f)]$, but the low-frequency bridge measurements simply measure R_0 . This means that errors could be produced if there is any frequency dependence in $\Delta Z = \Re[Z(f) - R_0]$ that does not take the same general form as the transfer function ratio as shown in Eq. (14). There are three known mechanisms that, in principle, could produce this type of anomalous frequency dependence. The first two mechanisms are discussed in this section, while the third, more subtle effect is discussed separately in the next section below.

The first anomalous frequency-dependent mechanism is due to lossy dielectrics that couple to sense resistor resistance, and would produce error terms proportional to the product of the lossy capacitance C_1 and the loss tangent $\tan(\delta)$ associated with just the sense resistor. These errors would be equivalent to those already discussed in Sec. 5.3 above in the case of lossy shunt capacitance C_s , except that the resistance measurements would be limited to very low frequencies.

The second mechanism is due to the well-known “skin effect,” where dependence on the frequency f arises via the characteristic skin depth $\delta_s = (\rho/\pi f \mu)^{1/2}$, for a conductor of resistivity ρ and permeability μ . The nickel-chrome-based alloy used in the construction of our foil resistors is known to have high resistivity, typically $\rho \approx 100 \mu\Omega\cdot\text{cm}$ or higher. This has the effect of creating relatively large skin depths; even at a frequency of $f = 1 \text{ MHz}$, we expect to have $\delta_s = 0.5 \text{ mm}$. The metal foil resistors have thickness $t \approx 0.025 \text{ mm}$ or less, so that $t/\delta_s \ll 1$ within our bandwidth. In this limit, the expected skin-effect frequency dependence of a foil resistor is very weak, and the first real nonvanishing term in an expansion in powers of t/δ_s would be of order $(t/2\delta_s)^4$, which is proportional to f^2 . Again, we would expect that the skin-effect

contributions to the observed frequency dependence in the ratio spectrum are small and indistinguishable from larger terms already accounted for in the model.

5.9.6 Second-Order Thermoelectric Effects

The third mechanism that could result in anomalous frequency dependence is sometimes referred to as a Peltier effect in dc measurements of resistors [69]. While ordinary thermal voltages will cancel upon current reversal in a resistance measurement, this second-order effect is polarity independent and is additive under current reversal. In this case, there is actually an error in the dc measurement of R_0 itself, and this is associated with a complex frequency dependence that is manifested only at low frequencies. This is purely a thermal effect and is of second order in the Seebeck coefficient S associated with junctions of dissimilar metals used in the resistor construction. In this case, the frequency dependence enters via a thermal skin depth $\delta_T = (\alpha/\pi f)^{1/2}$ and produces maximal errors in the dc limit of relative magnitude $S^2 T/\lambda \rho$ for thermal diffusivity α and thermal conductivity λ . For four-terminal junctions of gold bonding wire and Ni-Cr base metal, we estimate $S \approx 18 \mu\text{V/K}$, so that an idealized upper limit on the size of the effect is approximately 0.5 %. In practice, the effect is greatly attenuated by lateral (transverse to the current density) heat transport through the alumina substrates, and we are unaware of any examples of Peltier errors greater than $1 \mu\Omega/\Omega$ in similar foil resistor constructions. The actual error for our particular LCC package can be bounded by comparing the results of bridge measurements with different operating frequencies. Our experience with these foil resistors using bridges operating at both 30 Hz and 90 Hz sine wave and 6 Hz square wave excitation agrees with the results of bipolar dc measurement systems at the $0.1 \mu\Omega/\Omega$ level. We assign an uncertainty of $0.1 \mu\Omega/\Omega$ to account for any otherwise undetected resistance errors due to this effect.

Table 3. Uncertainty budget summary expressed in relative dimensionless units $\times 10^{-6}$ (parts per million).

| Category | Component | NIST 2011 | NIST 2017 |
|---------------------------------------|---|--------------|----------------|
| Ratio of the power spectral densities | Statistical | 5.2 | 4.25 |
| | Model ambiguity | See Sec. 5.5 | 1.63 |
| | Bandwidth ambiguity | — | 1.31 |
| | Spectral aberrations | 10.4 | — |
| | Dielectric loss | 2 | 1 |
| | EMI | 2.0 | 1 |
| | Nonlinearity and offsets | 1.0 | 0.5 |
| | Total Ratio | 12.0 | 4.97 |
| QVNS waveform | Frequency reference | < 0.001 | < 0.01 |
| | Quantization effects | < 0.5 | < 0.1 |
| | Total QVNS | | <0.1 |
| TPW temperature realization | Chemical impurities | 0.18 | 0.18 |
| | Pressure head correction | 0.18 | 0.08 |
| | Immersion error | 0.73 | 0.65 |
| | Isotopic variation | 0.26 | 0.1 |
| | Thermal equilibrium | 0.18 | 0.18 |
| | Total Temperature | 0.84 | 0.71 |
| Resistance | Bridge statistics | 0.4 | 0.1 |
| | Standard resistor calibration stability | 0.15 | 0.1 |
| | Standard resistor temperature | — | 0.05 |
| | Sense resistor relaxation/stability | 0.50 | 0.11 |
| | Frequency dependence | 0.10 | 0.05 |
| | Second-order thermoelectric effects | 0.10 | 0.1 |
| | Total Resistance | 0.67 | 0.22 |
| Total Combined Uncertainty | | 12.1 | 5.0 |

6. JNT System Hardware

The QVNS approach to JNT has been utilized for previous determinations of the Boltzmann constant starting in 2011 (see Table 4). The present work and these previous works have some core aspects in

common, but they are also different in many details, as the technology has continuously improved and methods have evolved.

Table 4. Comparison of methods, parameters, and technologies used in recent QVNS Boltzmann constant determinations.

| Feature (section) | NIST 2011 [34] | NIM 2015 [35] | NIM 2017 [36] | NIST 2017 [37] | NMIJ 2017 [39] |
|----------------------------------|--|--|--|--|---|
| Location | NIST Boulder | NIM Beijing | NIM Beijing | NIST Boulder | NMIJ Tsukuba |
| TPW (2.1) | Dry-well maintenance system | Ice bath | Ice bath | Dry-well maintenance system | Ice bath |
| Sense resistance R total (2.1) | 100 Ω | 200 Ω | 100 Ω | 200 Ω | 100 Ω |
| QVNS chip number (2.2) | 71218-32 (7a2) | 71218-54 (7A1) | 71218-54 (7A1) | 150519-44 (JNT 9c) | Not applicable |
| On-chip resistance $2R_0$ (2.2) | 200 Ω | 200 Ω | 200 Ω | 200 Ω | 200 Ω |
| Junction count (2.2) | 8 | 20 | 20 | 20 | 8 SQUIDS (output stage) ^a |
| QVNS CPW coupling (2.2) | Direct coupling | Direct coupling | Direct coupling | On-chip ac coupling | Inductive coupling to SQUIDS ^b |
| QVNS pulse bias type (2.2) | Two-level | Three-level Zero-comp | Three-level Zero-comp | Three-level Zero-comp | Two-level |
| Grounding approach (2.5) | Through QVNS semirigid coax and pulse bias generator | Through QVNS semirigid coax and pulse bias generator | Through QVNS semirigid coax and pulse bias generator | Through Dewar in star config. with TPW maintenance and batteries | Through amplifier enclosure |
| Matching method (2.3) | Trim R and coax lengths | Trim R and C | Coax lengths | Trim R and C | Trim R and coax lengths |
| Decoupling ferrites (2.3) | Ferrites between amps and relay board | Unknown | Ferrites on all cable pairs | Ferrites on R-probe or removed | Ferrites on all cable pairs |
| Connectors (2.3) | One 12 pin connector | Two pairs of 3 pin LEMO connectors | Two pairs of 3 pin LEMO connectors | Two pairs of 3 pin LEMO connectors | Two pairs of 3 pin LEMO connectors |
| Cabling (2.3) | Twisted pair | Twisted pair | (Micro) Coax | Twisted pair with shield | Twisted pair |
| Facility (2.5) | Unshielded room | Shielded room | Shielded room | Shielded room | Unshielded room |

^aTotal of 1131 junctions, including the code generator.

^bOn-chip high-frequency signals transmitted with rapid single flux quantum microstrip, not CPWs.

Table 4 lists, for comparison, the various features, components, and conditions for the present work (NIST 2017) and three other preceding QVNS-JNT measurements that have produced determinations of k . Certain major differences in hardware configurations (e.g., matching approach, ferrites, and cabling) are noted, which combine to create a unique configuration to each determination. The experimental feature under comparison is listed under the first (leftmost) column of the table along with the section of the text with the relevant descriptions.

7. Conclusion

In this article, we have described in detail the hardware, analysis, and individual components of the uncertainty budget for our recent Boltzmann constant determination. The most important limiting factor in the NIST measurements was the need to perform the measurements in a high-quality, EMI-shielded environment. NIST attempted measurements in five different locations: three in Boulder, Colorado, and two in Gaithersburg, Maryland, before the final measurements were completed in a shielded room in Boulder in the spring of 2017. After this measurement, which was stringently time-limited by the CIPM deadline of July 1, 2017, we used a modified QVNS probe to measure the undesired correlated noise in the system, that is, correlated noise that is not related to the Johnson noise of the sense resistor. This new measurement suggests that a number of the matching conditions used in our Boltzmann constant measurement [37] were not optimal, which limited our relative uncertainty, but this knowledge will also enable us to improve the matching conditions in future measurements. More generally, we hope that the details, observations, and insights that we provided here will assist metrologists and technologists in further

advancing Johnson noise thermometry to new applications, and perhaps even to develop practical, quantum-based, programmable primary thermometers for disseminating temperature and the kelvin.

Acknowledgments

We are grateful for the advice and support we received from our current collaborators: Paul Dresselhaus, Jifeng Qu, Rod White, John Martinis, Sae Woo Nam, John Labenski, Chiharu Urano, Kazuaki Yamazawa, and Jason Underwood. We acknowledge the staff and researchers who support and develop processes in the Boulder Microfabrication Facility. We thank Dave Howe, Corey Barnes, and Josh Savory for clock calibration, John Kitching for use of the shielded room, and Rand Elmquist and Marlin Kraft for resistance standard calibration.

8. References

- [1] Consultative Committee for Thermometry (2014) Report of the 27th Meeting (21–23 May 2014) to the International Committee for Weights and Measures, Recommendation T1. International Bureau of Weights and Measures (BIPM), Sèvres, France. <https://www.bipm.org/utis/common/pdf/CC/CCT/CCT27.pdf>.
- [2] Pitre L, Sparasci F, Truong D, Guillou A, Risehari L, Himbert M (2011) Determination of the Boltzmann constant using a quasi-spherical acoustic resonator. *International Journal of Thermophysics* 32:1825–1886. <https://doi.org/10.1098/rsta.2011.0197>.
- [3] de Podesta M, Underwood R, Sutton G, Morantz P, Harris P, Mark D, Stuart F, Vargha G, Machin G (2013) A low-uncertainty measurement of the Boltzmann constant. *Metrologia* 50(4):354. <https://doi.org/10.1088/0026-1394/50/4/354>.
- [4] Gaiser C, Zandt T, Fellmuth B (2015) Dielectric-constant gas thermometry. *Metrologia* 52:S217–S226. <https://doi.org/10.1088/0026-1394/52/5/S217>.
- [5] Zandt T, Sabuga W, Gaiser C, Fellmuth B (2015) Measurement of pressures up to 7 MPa applying pressure balances for dielectric-constant gas thermometry. *Metrologia* 52:S305–S313. <https://doi.org/10.1088/0026-1394/52/5/S305>.
- [6] Gaiser C, Fellmuth B, Haft N, Kuhn A, Thiele-Krivoi B, Zandt T, Fischer J, Jusko O, Sabuga W (2017) Final determination of the Boltzmann constant by dielectric constant gas thermometry. *Metrologia* 54:280–289. <https://doi.org/10.1088/1681-7575/aa62e3>.
- [7] Johnson J (1928) Thermal agitation of electricity in conductors. *Physical Review* 32:97–109. <https://doi.org/10.1103/PhysRev.32.97>.
- [8] Nyquist H (1928) Thermal agitation of electric charge in conductors. *Physical Review* 32:110–113. <https://doi.org/10.1103/PhysRev.32.110>.
- [9] Johnson J (1927) Thermal agitation of electricity in conductors. *Nature* 119:50–51. <https://doi.org/10.1038/119050c0>.
- [10] Callen H, Welton T (1951) Irreversibility and generalized noise. *Physical Review* 83:34. <https://doi.org/10.1103/PhysRev.83.34>.
- [11] Kubo R (1966) The fluctuation-dissipation theorem. *Reports on Progress in Physics* 29:255–284. <https://doi.org/10.1088/0034-4885/29/1/306>.
- [12] Fink H (1959) A new absolute noise thermometer at low temperatures. *Canadian Journal of Physics* 37:1397–1406. <https://doi.org/10.1139/p59-161>.
- [13] Brixy H, Hecker R, Oehmen J, Rittinghaus K, Setiawan W, Zimmermann E (1992) Noise thermometry for industrial and metrological applications. *Temperature, Its Measurement and Control in Science and Industry* (American Institute of Physics, New York), pp 993–996.
- [14] Dicke R (1946) The measurement of thermal radiation at microwave frequencies. *Review of Scientific Instruments* 17:268–275. <http://doi.org/10.1063/1.1770483>.
- [15] White D, Galleano R, Actis A, Brixy H, Groot M, Dubbeldam J, Reesink A, Edler F, Sakurai H, Shepard R, Gallop J (1996) The status of Johnson noise thermometry. *Metrologia* 33:325–335. <https://doi.org/10.1088/0026-1394/33/4/6>.
- [16] Mohr P, Newell D, Taylor B (2016) 2016 CODATA recommended values of the fundamental constants: 2014. *Reviews of Modern Physics* 88:035009. <https://doi.org/10.1103/RevModPhys.88.035009>.
- [17] Benz S, Martinis J, Nam SW, Tew W, White D (2001) A new approach to Johnson noise thermometry using a Josephson quantized voltage source for calibration. *Proceedings of TEMPMEKO 2001, International Journal of Thermophysics*, pp 37–44.
- [18] Benz S, Dresselhaus P, Martinis J (2003) An ac Josephson source for Johnson noise thermometry. *IEEE Transactions on Instrumentation and Measurement* 52:545–549. <https://doi.org/10.1109/TIM.2003.811687>.
- [19] Nam SW, Benz S, Dresselhaus P, Burroughs C, Tew W, White D, Martinis J (2005) Progress on Johnson noise thermometry using a quantum voltage noise source for calibration. *IEEE Transactions on Instrumentation and Measurement* 54(2):653–657. <https://doi.org/10.1109/TIM.2005.843574>.
- [20] Labenski J, Tew W, Nam SW, Benz S, Dresselhaus P, Burroughs C (2007) Resistance-based scaling of LF- and MF-band thermal noise powers. *IEEE Transactions on Instrumentation and Measurement* 56(2):481–485. <https://doi.org/10.1109/TIM.2007.891070>.
- [21] Tew W, Labenski J, Nam SW, Benz S, Dresselhaus P, Burroughs C (2007) Johnson noise thermometry near the zinc freezing point using resistance-based scaling. *Sixteenth Symposium on Thermophysical Properties, International Journal of Thermophysics* 28(2):629–645. <https://doi.org/10.1007/s10765-007-0196-9>.

- [22] Labenski J, Tew W, Benz S, Nam SW, Dresselhaus P (2008) A determination of the ratio of the zinc freezing point to the tin freezing point by noise thermometry. *Proceedings of TEMPMEKO 2007. International Journal of Thermophysics* 29:1–17. <https://doi.org/10.1007/s10765-007-0249-0>.
- [23] White D, Benz S (2008) Constraints on a synthetic-noise source for Johnson noise thermometry. *Metrologia* 45:93–101. <https://doi.org/10.1088/0026-1394/45/1/013>.
- [24] White D, Benz S, Labenski J, Nam SW, Qu J, Rogalla H, Tew W (2008) Measurement time and statistics for a noise thermometer with a synthetic-noise reference. *Metrologia* 45:395–405. <https://doi.org/10.1088/0026-1394/45/4/004>.
- [25] Toonen R, Benz S (2009) Nonlinear behavior of electronic components characterized with precision multitones from a Josephson arbitrary waveform synthesizer. *IEEE Transactions on Applied Superconductivity* 19:715–718. <https://doi.org/10.1109/TASC.2009.2019051>.
- [26] Benz S, Qu J, Rogalla H, White D, Dresselhaus P, Tew W, Nam SW (2009) Improvements in the NIST Johnson noise thermometry system. *IEEE Transactions on Instrumentation and Measurement* 58(4):884–890. <https://doi.org/10.1109/TIM.2008.2007027>.
- [27] Qu J, Benz S, Rogalla H, White D (2009) Reduced nonlinearities and improved temperature measurements for the NIST Johnson noise thermometer. *Metrologia* 46:512–524. <http://doi.org/10.1088/0026-1394/46/5/016>.
- [28] Benz S, White D, Qu J, Rogalla H, Tew W (2009) Electronic measurement of the Boltzmann constant with a quantum-voltage-calibrated Johnson-noise thermometer. *Comptes Rendus de l'Academie des Sciences* 10(2):849–858. <https://doi.org/10.1016/j.crhy.2009.10.008>.
- [29] Benz S, Dresselhaus P, Burroughs C (2011) Multitone waveform synthesis with a quantum voltage noise source. *IEEE Transactions on Applied Superconductivity* 21(3):681–686. <https://doi.org/10.1109/TASC.2010.2083616>.
- [30] Qu J, Benz S, Pollarolo A, Rogalla H (2011) Reduced nonlinearity effect on the electronic measurement of the Boltzmann constant. *IEEE Transactions on Instrumentation and Measurement* 60(7):2427–2433. <https://doi.org/10.1109/TIM.2010.2099310>.
- [31] Qu J, Zhang J, Fu Y, Rogalla H, Pollarolo A, Benz S (2013) Development of a quantum-voltage-calibrated noise thermometer at NIM. *Temperature: Its Measurement and Control in Science and Industry*, Vol. 8, ed Ripple DC (AIP Conference Proceedings, Melville, New York), Part 1, pp 29–33. <http://doi.org/10.1063/1.4821369>.
- [32] Qu J, Benz S, Fu Y, Zhang J, Rogalla H, Pollarolo A (2013) Flat frequency response in the electronic measurement of the Boltzmann constant. *IEEE Transactions on Instrumentation and Measurement* 62(6):1518–1523. <https://doi.org/10.1109/TIM.2013.2238431>.
- [33] Zhou Z, Qu J, Benz S (2015) Zero-compensation method and reduced inductive voltage error for the ac Josephson voltage standard. *IEEE Transactions on Applied Superconductivity* 25(5):1400806–6. <https://doi.org/10.1109/TASC.2015.2470684>.
- [34] Benz S, Pollarolo A, Qu J, Rogalla H, Urano C, Tew W, Dresselhaus P, White D (2011) An electronic measurement of the Boltzmann constant. *Metrologia* 48:142–153. <https://doi.org/10.1088/0026-1394/48/3/008>.
- [35] Qu J, Benz S, Pollarolo A, Rogalla H, Tew W, White R, Zhou K (2015) Improved electronic measurement of the Boltzmann constant by Johnson noise thermometry. *Metrologia* 52:S242–S256. <https://doi.org/10.1088/0026-1394/52/5/S242>.
- [36] Qu J, Benz S, Coakley K, Rogalla H, Tew W, White R, Zhou K, Zhou Z (2017) An improved electronic determination of the Boltzmann constant. *Metrologia* 54:549–558. <https://doi.org/10.1088/1681-7575/aa781e>.
- [37] Flowers-Jacobs N, Pollarolo A, Coakley K, Fox A, Rogalla H, Tew W, Benz S (2017) A Boltzmann constant determination based on Johnson noise thermometry. *Metrologia* 54:730–737. <https://doi.org/10.1088/1681-7575/aa7b3f>.
- [38] Maezawa M, Yamada T, Urano C (2016) Improved design of integrated quantum voltage noise source. *IEEE Transactions on Applied Superconductivity* 26(3):1800504. <https://doi.org/10.1109/TASC.2016.2525990>.
- [39] Urano C, Yamazawa K, Kaneko NH (2017) Measurement of the Boltzmann constant by Johnson noise thermometry using superconducting integrated circuit. *Metrologia* 54(6). <https://doi.org/10.1088/1681-7575/aa7cdd>. Retrieved 8 September 2017.
- [40] Coakley K, Qu J (2017) Spectral model selection in the electronic measurement of the Boltzmann constant by Johnson noise thermometry. *Metrologia* 54:204–217. <https://doi.org/10.1088/1681-7575/aa5d21>.
- [41] Benz S, Hamilton C (1996) A pulse-driven programmable Josephson voltage standard. *Applied Physics Letters* 68:3171–3173. <http://doi.org/10.1063/1.115814>.
- [42] Benz S, Hamilton C, Burroughs C, Harvey T, Christian L, Przybysz J (1998) Pulse-driven Josephson digital/analog converter. *IEEE Transactions on Applied Superconductivity* 8:42–47. <https://doi.org/10.1109/77.678440>.
- [43] Van der Ziel A (1970) *Noise: Sources, Characterization, Measurement* (Prentice Hall, Englewood Cliffs, NJ), 1st Ed.
- [44] Sampietro M, Fasoli L, Ferrari G (1999) Spectrum analyzer with noise reduction by cross-correlation technique on two channels. *Review of Scientific Instruments* 70:2520–2525. <https://doi.org/10.1063/1.1149785>.
- [45] Pollarolo A, Tew W, Rogalla H, Underwood J, Benz S (2014) Systematic error resolved in NIST Johnson noise thermometer. *29th Conference on Precision Electromagnetic Measurements (CPEM 2014)*, Rio de Janeiro, Brazil. <https://doi.org/10.1109/CPEM.2014.6898241>.
- [46] Elmquist R, Jarrett D, Jones G, Kraft M, Shields S, Dziuba R (2004) NIST Measurement Service for dc Standard Resistors. NIST Technical Note 1458. <https://www.nist.gov/sites/default/files/documents/calibrations/tn1458.pdf>.
- [47] Guildline (2002) Automatic DCC resistance/temperature bridge. *6675AV1.0 Datasheet* (Guildline Instruments Limited, Smith Falls, Ontario, Canada).
- [48] Bramley P, Pickering J (2006) Better accuracy in temperature calibration and measurement through a new type of analog-to-digital converter. *Cal Lab Magazine* 10/11:21–26.
- [49] Benz S (1995) Superconductor-normal-superconductor junctions for programmable voltage standards. *Applied Physics Letters* 67:2714–2716. <https://doi.org/10.1063/1.114302>.
- [50] Van Duzer T, Turner C (1999) *Principles of Superconductive Devices and Circuits* (Prentice Hall, Upper Saddle River, NJ), 2nd Ed.

- [51] Howe L, Fox A, Rufenacht A, Burroughs C, Dresselhaus P, Benz S (2015) NIST 10 V programmable Josephson voltage standard system using a low-capacity cryocooler. *IEEE Transactions on Applied Superconductivity* 25(3):1400404. <https://doi.org/10.1109/TASC.2014.2367531>.
- [52] Pollarolo A, Jeong T, Benz S, Rogalla H (2013) Johnson noise thermometry measurement of the Boltzmann constant with a 200 Ω sense resistor. *IEEE Transactions on Instrumentation and Measurement* 62(6):1512–1517. <https://doi.org/10.1109/CPEM.2012.6250878>.
- [53] Pollarolo A, Jeong T, Tew W, Benz S, Rogalla H (2013) Johnson noise thermometry based on a quantized-voltage noise source at NIST. *AIP Conference Proceedings: Temperature: Its Measurement and Control in Science and Industry*, Vol. 8, pp 23–28. <https://doi.org/10.1063/1.4821368>.
- [54] Pickup CP (1975) A high-resolution noise thermometer for the temperature range 90–100 K. *Metrologia* 11:151–159. <https://doi.org/10.1088/0026-1394/11/4/002>.
- [55] White D (1984) Systematic errors in a high-accuracy Johnson noise thermometer. *Metrologia* 20:1–9. <https://doi.org/10.1088/0026-1394/20/1/002>.
- [56] White D, Zimmermann E (2000) Preamplifier limitations on the accuracy of Johnson noise thermometers. *Metrologia* 37(1):11–23. <https://doi.org/10.1088/0026-1394/37/1/3>.
- [57] Callegaro L, Pisani M, Ortolano M (2010) Systematic errors in the correlation method for Johnson Noise thermometry: Residual correlations due to amplifiers. *Metrologia* 47(3):272–278. <https://doi.org/10.1088/0026-1394/47/3/018>.
- [58] Draper N, Smith H (1998) *Applied Regression Analysis* (Wiley, New York, NY), 3rd Ed.
- [59] White D, Mason R (2005) An EMI test for Johnson noise thermometry. *Proceedings of TEMPMEKO 2004*, ed D Zvizdic *et al* (Faculty of Mechanical and Naval Architecture, Zagreb, Croatia), pp 485–490.
- [60] White D (2012) Non-linearity in Johnson noise thermometry. *Metrologia* 49:651–665. <https://doi.org/10.1088/0026-1394/49/6/651>.
- [61] ASTM International (2017) Standard specifications for laminated thermosetting materials. *Active Standard ASTM D709* (ASTM International, West Conshohocken, PA).
- [62] Kyocera (2015) Characteristics of Kyocera Technical Ceramics, Materials Datasheet (Kyocera, San Diego, CA).
- [63] Hill K (2001) Is there a long-term drift in triple point of water cells? *Metrologia* 38(1):79–82. <https://doi.org/10.1088/0026-1394/38/1/7>.
- [64] Preston-Thomas H (1990) The international temperature scale of 1990 (ITS-90). *Metrologia* 27(1):3–10. <https://doi.org/10.1088/0026-1394/27/1/002>.
- [65] White D, Jongenelen C (2010) The immersion characteristics of industrial PRTs. *International Journal of Thermophysics* 31:1685–1695. <https://doi.org/10.1007/s10765-010-0816-7>.
- [66] Consultative Committee for Thermometry (2011) *Mise en pratique* for the definition of the kelvin. International Bureau of Weights and Measures (BIPM), Sèvres, France.
- [67] Gonfiantini R (1978) Standards for stable isotope measurements in natural compounds. *Nature* 271:534–536. <https://doi.org/10.1038/271534a0>.
- [68] Jarret D, Elmquist R, Kraft M, Jones G, Payagala S, Seifert F, Haddad D, Schlamming S (2016) Quantum Hall resistance traceability for the NIST-4 watt balance. *30th Conference on Precision Electromagnetic Measurements (CPEM 2016)*, Ottawa, ON, 2016. <https://doi.org/10.1109/CPEM.2016.7540787>.
- [69] Kirby C, Laubitz M (1973) The error due to the Peltier effect in direct-current measurements of resistance. *Metrologia* 9(3):103–106. <https://doi.org/10.1088/0026-1394/9/3/001>.

About the authors: Nathan E. Flowers-Jacobs is a physicist in the Superconductive Electronics Group in the Quantum Electromagnetics Division at NIST. He has been working on development, characterization, and applications of the Josephson arbitrary waveform synthesizer (JAWS), an ac Josephson voltage standard based on pulse-biased arrays of Josephson junctions.

Alessio Pollarolo is an electrical engineer, and from 2009 to 2017, he was a guest researcher at NIST working on the Johnson Noise Thermometry program in the Quantum Electromagnetics Division. He is currently the vice president of measurement science at Measurements International, Ltd.

Kevin J. Coakley is a mathematical statistician in the Statistical Engineering Division at NIST. His research interests include statistical signal processing, computer intensive statistical methods, and planning and analysis of experiments in physical science.

Adam C. Weis is a postdoctoral research associate working on the Johnson Noise Thermometry program in NIST's Quantum Electromagnetics Division. He is working on developing a noise thermometer for accessibility and dissemination.

Anna E. Fox is a researcher working in the Quantum Voltage program in the Quantum Electromagnetics Division at NIST. She is developing designs and fabrication processes for voltage standard devices such as the programmable Josephson voltage standard and the ac Josephson voltage standard.

Horst Rogalla has been a senior scientist at NIST since 2007 and is the project leader of the Johnson Noise Thermometry program in the Quantum Electromagnetics Division. His research interests are in superconducting electronics and materials science, especially related to thin film growth and properties.

Weston L. Tew is a physicist in the NIST Process Measurements Division and Sensor Science Division. His research focuses on low-temperature thermometry standards, gas-based fixed points, resistance thermometry, and noise thermometry.

Samuel P. Benz is the group leader of the Superconductive Electronics Group in the Quantum Electromagnetics Division at NIST. He has worked on a broad range of topics within the field of superconducting electronics, including Josephson junction array oscillators, single flux quantum logic, ac and dc Josephson voltage standards, Josephson waveform synthesis, and noise thermometry.

The National Institute of Standards and Technology is an agency of the U.S. Department of Commerce.