The programmable spectrum analyzer

Closing the gap between software-defined radio hardware and instrument-grade measurement.

ABSTRACT

Engineers building monitoring, signals intelligence and automated test systems work from two incomplete toolboxes: software-defined radio hardware that is programmable and uncalibrated, and instruments that are calibrated and closed. This paper sets out what separates a measurement from a reading, derives six requirements a programmable instrument must meet, and works through them against published figures.

WHO THIS PAPER IS FOR

RF and systems engineers building on SDR hardware who need results that survive review. **Test engineers** automating measurement. **Integrators** evaluating whether an instrument can sit inside their own product.

CONTENTS

1 Two toolboxes, neither complete

An engineer asked to build a spectrum monitoring system, a survey rig or an automated test sequence has two places to start, and both are wrong in a different direction.

Software-defined radio hardware is inexpensive, programmable to the sample, and surrounded by a large open ecosystem. What it lacks is a calibration certificate. Its absolute amplitude accuracy is unstated, its gain stages drift with temperature, its phase noise is whatever the synthesizer happened to deliver, and its image and IF rejection depend on a quadrature balance nobody characterized. It will give you a number. It will not tell you how wrong that number might be.

A conventional instrument is calibrated, specified and traceable, and its numbers hold up. It is also priced like it, and the price climbs steeply with frequency: at millimeter-wave the bench instrument can cost five to ten times what the receiver underneath it would. And what it will not do is let you inside. The processing is fixed, the measurement set is often licensed piece by piece, the command vocabulary changes between models, and the moment the application needs something the vendor did not anticipate, you are finished.

The obvious response is to calibrate the open hardware yourself. Done properly that means a traceable source, a calibrated attenuator, a stable thermal environment, characterization across every frequency and gain state in use, and a repeat per unit and at intervals. Even then you have a correction table rather than a specification, and no stated uncertainty.

2 What makes a reading a measurement

Three things separate an instrument result from a number on a screen, and none of them are about performance.

- **A traceable chain**, documented and unbroken back to a national standard. This is the substance of ISO/IEC 17025.
- **A stated uncertainty**: not an assurance of accuracy but a bound, under stated conditions.
- **Documented conditions.** A specification without conditions is an advertisement.

The instrument this paper works through, named in section 6, states its specifications after ten minutes of warm-up at 25 degrees C with spur rejection standard, and quotes amplitude accuracy as plus or minus 2.0 dB to 9.5 GHz and plus or minus 3.0 dB above. Those figures are typical for a field instrument, which is the point: they exist, they are bounded, and they are conditioned. An uncalibrated receiver does not have a worse number. It has no number.

The cheapest illustration is the frequency reference. A fractional error produces an absolute error that scales with the carrier:

$$\Delta f = f_c \times \delta \times 10^{-6} \qquad (1)$$

where

f_c is the carrier frequency, in hertz

δ is the fractional reference error, in parts per million

> **PRACTICAL RULE**
>
> The standard oscillator is specified below 1 ppm. That is a 1 kHz error at 1 GHz, invisible in most work, and a 40 kHz error at 40 GHz, which is not. The oven-controlled option, below 0.15 ppm, brings 40 GHz to roughly 6 kHz. Specify the reference for the top of your band, not the bottom.

3 Where the boundary sits

Every receiving instrument draws a line. On one side is hardware the user cannot change. On the other is processing that could in principle be rewritten. The history of the last thirty years is that line moving toward the antenna, and more of what sits behind it becoming the user's rather than the vendor's.

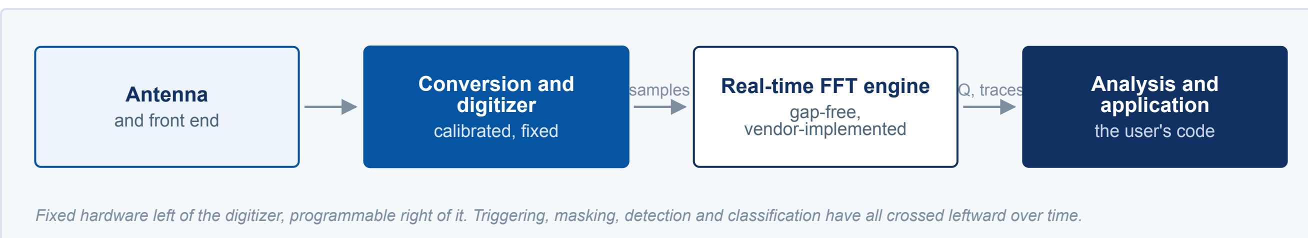

Figure 1. The boundary between fixed hardware and programmable processing is the architectural question that matters. An instrument is worth building on when the boundary sits early and everything to the right of it is documented. SCHEMATIC

Use Figure 1 as a scoring device: ask where an instrument's boundary sits and whether the crossing is documented. Section 7 shows what it looks like when the boundary sits early and the crossing is an open standard.

4 Why the conventional instrument falls short

The conventional instrument does exactly what it was designed to do, which is an operator making a bounded set of measurements on a bench. It fails in three ways when it is asked to become part of a system.

- **The measurement set is sold by the piece.** A function already in the firmware is withheld until purchased, so a fleet of nominally identical units is not identical.
- **The command vocabulary changes between models**, so code written against one instrument often will not run on the next tier up.
- **The processing is a black box.** Settings can be changed; the algorithm cannot, and the samples that fed it are usually unreachable.

Those three are inconveniences. The fourth is worse, because it does not announce itself. A closed instrument answers every question you ask it and none that you did not, and the answers look identical either way. Ask it for the level in a channel and it returns a number; if a 2 microsecond transient sat inside the measurement window and the detector averaged across it, the number is low and nothing on the display says so. The reading is not flagged, not caveated and not obviously wrong. It is simply the wrong quantity, reported with the confidence of the right one.

THE FALSE NEGATIVE THIS PRODUCES

A device is tested for spurious emissions and passes. Six months later a customer reports interference that the test should have caught. The test was run correctly, on a calibrated instrument, by someone competent, and it still missed the emission, because the detector integrated across a burst that was present for less than one percent of the dwell. **On a closed instrument that failure is undetectable after the fact**, because the samples that produced the number no longer exist. On an open one it is a question you can answer in an afternoon by reprocessing the record. The difference between the two is not accuracy. It is whether the measurement can be audited at all.

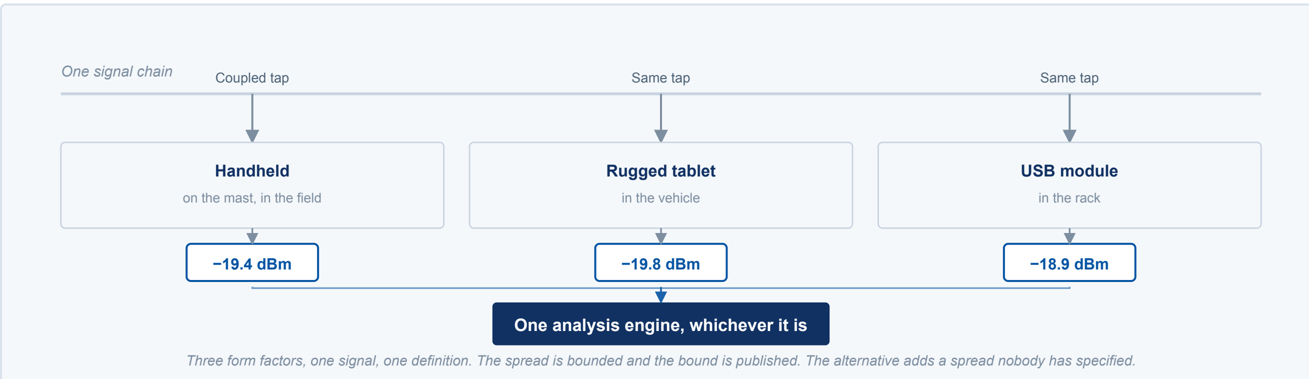

Figure 2. The same tap measured by three form factors under one analysis engine. The three readings differ, and they are allowed to: all three are within one stated amplitude bound and they are the same quantity, so the spread is known. Assemble the fleet from three vendors and the readings differ by that bound *plus* each vendor's reference conventions, which is a difference no calibration certificate removes because every one of the three is individually correct. SCHEMATIC

Figure 2 is the consequence drawn out, and it is the whole of the procurement argument: standardizing on one analysis engine across every form factor is not a preference, it is what makes two numbers mean the same thing.

5 Derived requirements

Six requirements follow from the argument so far, and each is checkable against any vendor's published data. Table 1 sets them out.

Table 1. Requirements derived from sections 2 to 4. Everything after this is an assessment against these six lines.

#	Requirement	Why not something looser
R1	Bounded amplitude accuracy under documented conditions	Without it the result cannot enter a test record.
R2	Gap-free capture with a published probability of intercept	A detector cannot be better than the data feeding it.
R3	Raw IQ, with published rate, depth and trigger behavior	The user's algorithm needs samples, not the vendor's trace.
R4	One documented API across every form factor	Code validated on the bench must run in the field unchanged.
R5	The measurement set included, not licensed piece by piece	Anything withheld is not scriptable, so it is not usable.
R6	A frequency reference specified for the top of the band	Reference error scales with carrier frequency.

One application, carried through

Requirements derived from a grievance are easy to agree with and hard to use, so run them against a job. A research group characterizing a wideband emitter across twenty units wants results that can go into a paper. R1 sets the floor: every amplitude needs a bound and a condition, so the reference plane, the warm-up and the temperature go into the method section or the numbers are decorative. R4 decides whether the work is feasible at all, since a routine revalidated per unit turns twenty instruments into twenty experiments. R5 decides the budget, because a measurement withheld behind a license on nineteen of the twenty is one the study cannot use.

R2 and R3 decide what the group can claim. If the emitter is continuous, the intercept guarantee never binds and a swept instrument would have done; that is the case where this paper argues against itself, and it is common in teaching. If the emitter is bursty, R2 sets the shortest event whose amplitude is defensible and R3 is what lets a reviewer ask for the raw samples, which is the request that separates a result from a screenshot. **The requirements do not change with the job. Which of them binds does**, and knowing which one binds is most of what specifying an instrument is.

6 Meeting the requirements

Sections 2 to 4 were about what to require, and nothing in them depends on who makes the instrument. This section is about whether anything meets the requirement, which is a different kind of question: it is about a particular instrument rather than about a class. From here the paper works one example and shows its published numbers, and a reader who prefers a different one can run the same six requirements against it.

A real-time analyzer built as a programmable platform, such as the Berkeley Nucleonics ICX-FieldHawk family, is designed against roughly this list.

R1 and R6 are already settled by section 2: amplitude accuracy is bounded and conditioned, and the frequency reference is specified in two grades so it can be chosen for the top of the band rather than the bottom. The remaining four are the interesting ones.

R2: the real-time engine

The engine transforms a continuous stream in an FPGA with no missing samples. Two published relations govern its timing. The frame rate is the reciprocal of the transform span:

$$R_f = 1 / (N \times D \times 8\ \text{ns}) \tag{2}$$

where

R_f is the frame rate, in frames per second

N is the transform size, in points

D is the decimation factor

$8\ ns$ is the engine's sample period, so $N\,D$ × 8 ns is the span of one transform

At N = 32 with no decimation the transform spans 256 ns, so the engine produces 3,906,250 frames per second. At N = 2048 it spans 16.384 microseconds and produces 61,035.

And the duration a signal must be present for full amplitude accuracy is:

$$\text{POI} = 2 \times N \times D \times 8\ \text{ns} \tag{3}$$

where

POI is the 100 percent probability of intercept, in seconds

Prove the factor of two from the least flattering arrival. A burst beginning one sample before a transform boundary leaves that frame holding a single sample of it, reported low by 20 $\log_{10}(N)$, or 66 dB at N = 2048, and detected all the same. Only a burst long enough to fill some later frame end to end is measured correctly, and the shortest burst for which that holds regardless of arrival phase is two frames. The factor is not a margin. It is the worst arrival phase, and any figure quoted without it is quoting the best case.

Detecting a pulse and measuring it are different claims, and only the second is worth quoting. It is where published intercept figures quietly diverge from each other.

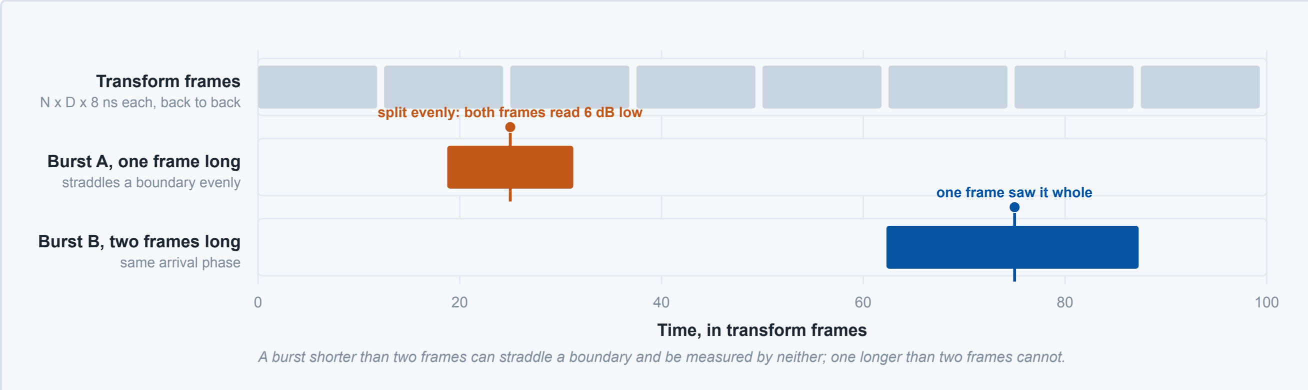

Figure 3. Why the intercept figure carries a factor of two. Burst A is exactly one frame long and straddles a boundary evenly, so each frame holds half of it and both report 6.02 dB low. Burst B is two frames long, so whatever its arrival phase, some frame contains it end to end. The factor is the worst arrival phase rather than a safety margin, which is why an intercept figure quoted without it is quoting the best case. SCHEMATIC

Figure 3 draws that argument. Read the two event markers together: detection and measurement are different claims, and the gap between them is exactly one frame wide.

R3: getting at the samples

Raw IQ is a first-class output. The sample rate reaches 125 MSPS with decimation from 1 to 4096 in powers of two, burst recording covers the full 100 MHz, continuous recording covers 25 MHz, the capture buffer is 128 Mbyte, and external triggers are accepted up to 500 times per second. The data rate follows from the sample rate and format:

$$r = f_s \times 2 \times b/8 \qquad (4)$$

where

r is the data rate, in bytes per second

f_s is the complex sample rate

b is the bits per component, 16 in a conventional interleaved format

> **THE ONE NUMBER WORTH REMEMBERING**
>
> At 125 MSPS with 16-bit components that is 500 Mbyte per second, so the 128 Mbyte buffer holds about 0.26 seconds of gap-free 100 MHz capture. With a 0.512 microsecond intercept, everything inside that quarter second lasting longer than half a microsecond was measured rather than glimpsed: roughly half a million resolvable intervals in one trigger, all of them reachable from your own code.

That is also why continuous recording is specified at 25 MHz rather than 100 MHz: sustained transfer to a host is a different constraint from filling a buffer, and the datasheet separates them instead of quoting the flattering number twice.

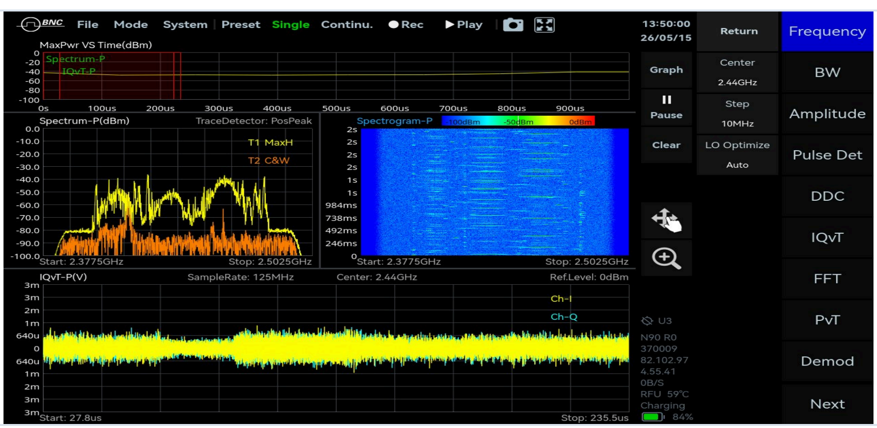

Figure 4. Raw IQ as an instrument mode rather than a host feature: I and Q against time across a 208 microsecond window, with a sample-rate field reading 125 MHz, above a spectrum and a two-second spectrogram computed from the same acquisition. The span reads 2.3775 to 2.5025 GHz, which is 125 MHz, the sample rate rather than the usable 100 MHz. This is one acquisition into the buffer, so it is the burst case above and not the sustained one. MEASURED

Figure 4 is R3 on the glass before any host is involved, which is the half of the claim the flowgraph screens later in this paper cannot make.

> **HONEST LIMITATION**
>
> This instrument does not stream 100 MHz to disk indefinitely, and nothing in its class does. If your application genuinely requires continuous wideband recording for minutes at a time, you need a recorder with a storage array behind it.

R4 and R5: one API, nothing held back

The ICX-FieldHawk host software is SpectraCore, and the published programming interfaces beneath it are C, C++, C#, Python, MATLAB, Qt and LabVIEW, with SCPI as standard, on Windows and Linux, for x64 and AArch64 hosts. The same interface serves the USB module, the LAN module, the handheld and the rugged tablet, so a routine developed on the bench and the same routine in the field produce comparable numbers. Comparability is a property of the software, not of the boxes.

That also answers the licensing question. Channel power, occupied bandwidth, x dB bandwidth, ACPR, IM3, SEM, AM and FM demodulation, antenna factor, amplitude offset, signal track, peak table, record and playback, multiple unit display and amplitude correction all ship in SpectraCore, alongside the spectrum, real-time, power detection, phase noise and harmonics modes. Pulse detection and digital demodulation are options. What matters for R5 is not the length of that list but that nothing on it is withheld behind a license, so a routine written on one unit runs on every unit in the fleet.

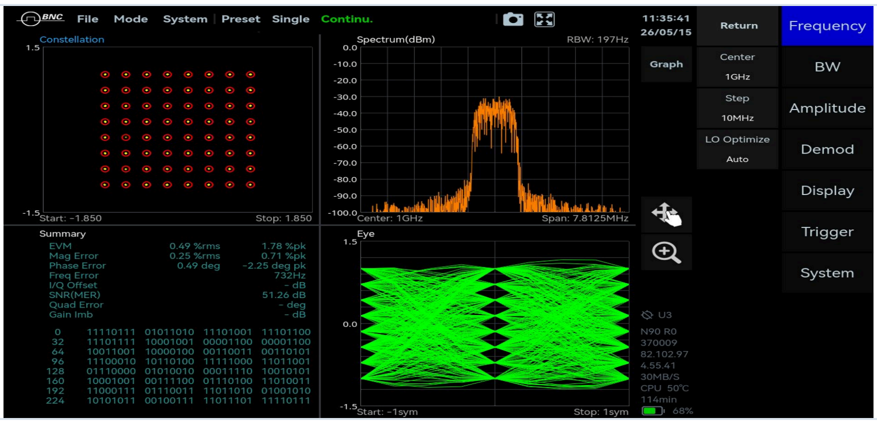

Figure 5. The same class of measurement on the instrument's own side of the boundary: a 64-QAM carrier at 1 GHz in a 7.8125 MHz span at 197 Hz resolution, with constellation, spectrum, eye and recovered symbols. Two summary fields repay reading together. The 732 Hz frequency error at 1 GHz is 0.73 ppm, inside the instrument's own sub-1 ppm reference bound, which is the self-consistency to look for. But the 0.49 percent rms error vector magnitude and the 51.26 dB modulation error ratio are *not* related by MER = −20 log(EVM), which pairs 0.49 percent with 46.2 dB: the panels are referenced differently and neither should be quoted without establishing which. **MEASURED**

Figure 5 is worth setting against the open-flowgraph constellation in section 7. The instrument computes that summary itself, and it will also hand over the samples it computed it from. A closed instrument offers the first and not the second, which is a difference that only matters on the day the summary is wrong.

7 The open ecosystem, with the calibration attached

Everything so far has been an argument that the instrument is programmable. This section is the evidence, and it comes from outside the instrument: the receiver now presents itself to the open software-defined radio ecosystem as an ordinary device.

SoapySDR is a vendor-neutral hardware abstraction layer. An application written against it runs against any receiver that supplies a driver module, with no knowledge of the hardware underneath. That is how GNU Radio, Gqrx and a large body of open signal processing code reach hardware at all. A driver module for the ICX-FieldHawk family translates standard SoapySDR calls into the instrument's own interface, so the analyzer is discovered, tuned, configured and streamed exactly like any other SoapySDR device.

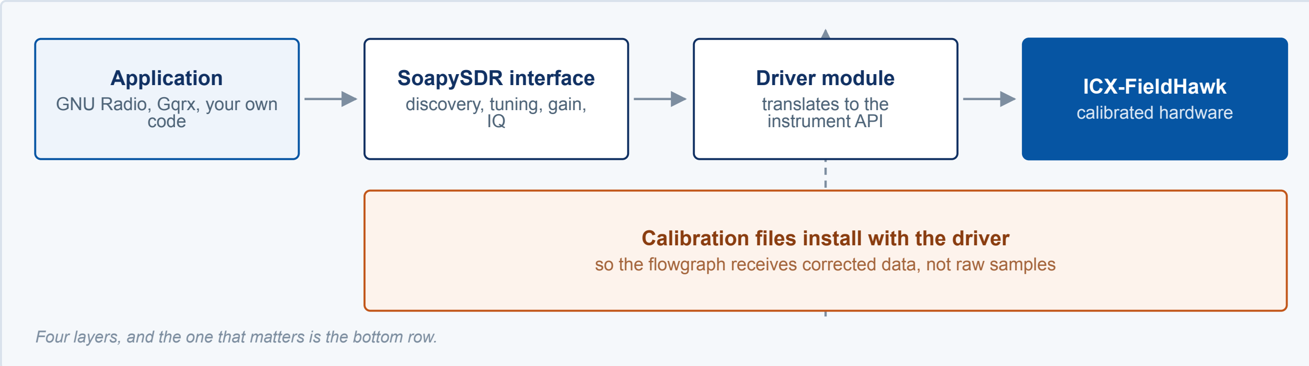

Figure 6. The abstraction layer itself is standard and widely used. What is unusual is the bottom row: the device calibration installs with the driver, so an open-source flowgraph is handed the instrument's corrected data rather than raw converter counts. SCHEMATIC

> **WHAT THE CALIBRATION IS WORTH, IN DB**
>
> Figure 6 looks like every other SoapySDR stack until you ask what reaches the first block of the flowgraph. An uncorrected receiver reports full scale. Turning that into dBm needs a gain-state table the user builds and maintains for every gain setting and every frequency in use, and re-checks when it drifts. Here that table installs with the driver, so a level read inside a flowgraph carries the same plus or minus 2.0 dB bound as a level read on the instrument's own display.

The integration has been demonstrated end to end rather than merely asserted. Table 2 lists what has been shown, with the signal used in each case.

Table 2. Demonstrated applications, each with a real signal and a captured result rather than a claim. AM, narrowband and wideband FM, CW, FSK and PSK run on the same path.

Application	Signal used	Result
16-QAM	1 GHz at **−80 dBm**, 500 kSym/s, root-raised-cosine 0.35	clean constellation, IQ, spectrum
WLAN 802.11a/g	2.412 GHz at −40 dBm, 12 Mb/s, QPSK, rate 1/2	constellation, and frames **decoded to a PCAP file**
ADS-B 1090ES	live air traffic	aircraft address, callsign, altitude, speed, heading and position

Source: the platform integration guide, August 2026. Demonstrated rather than specified; see the verification note.

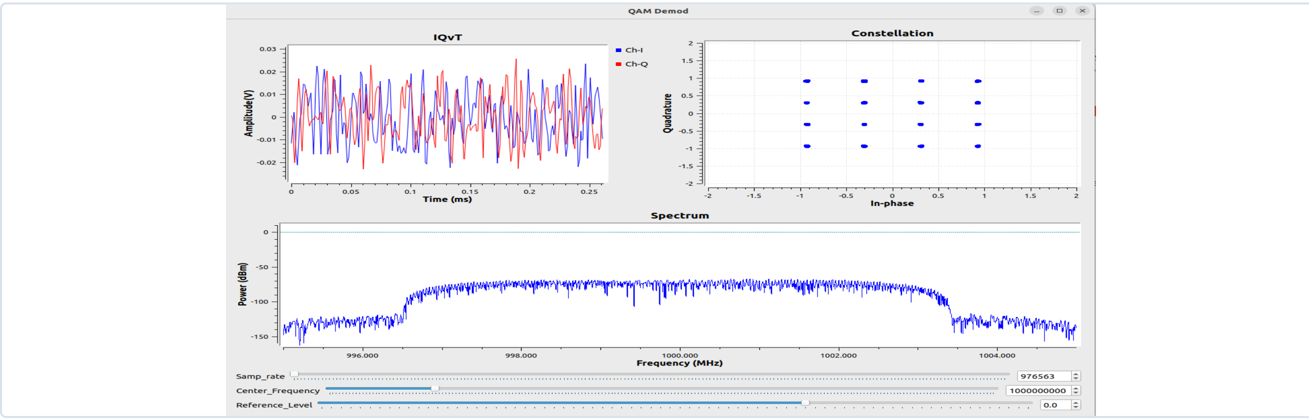

Figure 7. A 16-QAM constellation recovered in GNU Radio, with the IQ record and the spectrum. The flowgraph controls read a 1 GHz center, 976.563 kSPS and a 0 dBm reference level; the −80 dBm source level is the integration guide's figure, not a reading off this display. What the calibration buys is not the constellation's shape, which comes from noise figure and phase noise, but the axis it is plotted against: the spectrum reads in dBm, an absolute level with a stated bound, where an uncorrected board reports a fraction of full scale. **MEASURED**

The Wi-Fi case goes further, because it ends in a file another tool can open. The receive chain performs packet detection, synchronization, transform, channel equalization and MAC decoding, and writes frames out for Wireshark. Every block in that chain downstream of the source is open-source and replaceable, which is the point: the instrument supplies calibrated samples and stays out of the way.

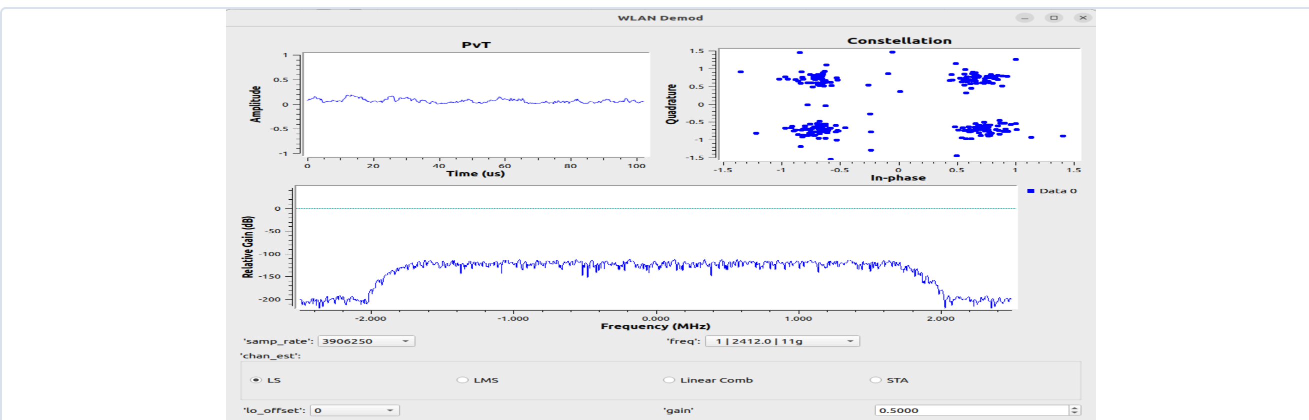

Figure 8. The 802.11a/g OFDM physical layer received at 2.412 GHz, the flowgraph's channel selector reading 11g and its least-squares channel estimator selected: four clusters in the constellation pane and the channel response below them. The sample-rate field reads 3,906,250, which is the instrument's 125 MSPS divided by 32, so the host is running an exact binary decimation of the converter rate rather than a resampled approximation of it. That the decoded frames leave as a PCAP file is the integration guide's claim rather than a field on this window. **MEASURED**

Figure 7 and Figure 8 share a detail worth pausing on. One host reports 976.563 kSPS and the other 3,906,250, which are 125 MSPS divided by 128 and by 32: the powers-of-two decimation of R3, selected from a flowgraph and reported back by it. Neither number was chosen to make that point. A protocol decoder, a classifier or a logging path is another block in the same graph, and the amplitudes underneath stay traceable.

> **HONEST LIMITATION**
>
> This path has been demonstrated on x86_64 hosts running Ubuntu 22.04 or later with GNU Radio 3.9 or later over USB 3.0. It is not a claim about every operating system, host architecture or transport. If your environment differs, ask before you plan around it.

One more concession, and it is the one that decides some purchases. If the work is algorithm development where absolute level never has to be defended, if the band is narrow and fixed, and if the unit count is high enough that price dominates, bare software-defined radio hardware is the right buy and this paper is an argument against your own interest. The calibration, the specified conditions and the traceable amplitude cost money, and they earn it only when a number has to leave the building with your name on it.

8 Measurement practice

None of the following depends on which analyzer you buy. All of it determines whether the result is worth anything.

1. **Let it warm up.** The published accuracy figures assume ten minutes. Measurements from a cold start are covered by no specification on the datasheet.
2. **Set the reference level deliberately.** Amplitude accuracy, intermodulation performance and displayed noise floor all move with it. Automatic reference level during a comparison invalidates the comparison.
3. **Record the state, not just the trace.** Center frequency, span, resolution bandwidth, detector, reference level, attenuation, preamplifier state, transform size and temperature. A trace is a conclusion someone else already drew; IQ plus its state is the evidence.
4. **Check the sample format before parsing an IQ file.** Interleaving order, component width and scaling silently corrupt an otherwise perfect capture.

Field evidence

The buyers for this arrangement are universities, research groups and experimenters who need a result that can go into a paper rather than into a demonstration, and what they report back is not about throughput or block count. It is that the amplitude axis survives the boundary: a level read inside a flowgraph is the level the instrument would have displayed, so a plot produced by student code carries the same bound as one produced by the vendor's. That is a narrower claim than the section 7 demonstrations make and it is the one that decides whether a result is publishable.

9 Summary and selection

Table 3 collects the governing relations and published values a reader is most likely to want again.

Table 3. Governing relations and published values, each with its condition. Every hero number in this paper appears here with its condition attached.

Quantity	Relation or value	Condition
Frame rate	1 / (N × D × 8 ns)	3,906,250 /s at N = 32; 61,035 /s at N = 2048, D = 1
100 percent POI	2 × N × D × 8 ns	full amplitude accuracy
Published operating points	0.512 us at N = 32; 32.768 us at N = 2048	D = 1
Gap-free analysis bandwidth	100 MHz where fitted	the 4.5 to 9 GHz handheld models ship with 50 MHz as standard, 100 MHz optional
IQ sample rate	up to 125 MSPS, decimation 1 to 4096	powers of two
Capture buffer	128 Mbyte	about 0.26 s at 100 MHz, 16-bit
Continuous recording	25 MHz	sustained to host
Amplitude accuracy	±2.0 dB to 9.5 GHz; ±3.0 dB above	25 °C, 10 min warm-up, spur reject on
Reference	TCXO < 1 ppm; OCXO option < 0.15 ppm	temperature stability

Source: the ICX-FieldHawk handheld, rugged and USB datasheets. Preliminary values on the portfolio and overview pages should be verified prior to application deployment.

Choosing a form factor. Build systems around the USB or LAN module. Develop on the module and deploy to the handheld, since one API works for all platforms and your validated routines remain unchanged. If you are working above 9.5 GHz you need to use the 40 GHz variant. Those are the two frequency tiers which include published specifications, so any requirement written against a number in this paper has to fall back to one of those two models.

10 Definitions, symbols and further reading

Calibration
Comparison against a reference of known accuracy, documented, with the reference itself traceable to a national standard.

Gap-free capture
Acquisition in which consecutive transforms overlap sufficiently that no interval of time is unobserved.

Probability of intercept
The minimum duration a signal must be present for its amplitude to be measured correctly. A stronger claim than the duration required merely to detect it.

Symbols used in this paper.

Symbol	Meaning	Units
f_c	Carrier frequency	Hz
f_s	Complex sample rate	samples per second
δ	Fractional reference error	parts per million
N	Transform size	points
D	Decimation factor	dimensionless
8 ns	Engine sample period	s

FFT fast Fourier transform
FPGA field-programmable gate array
IQ in-phase and quadrature
OCXO oven-controlled crystal oscillator
POI probability of intercept
SCPI Standard Commands for Programmable Instruments
SDR software-defined radio
TCXO temperature-compensated crystal oscillator

Further reading

- J. Mitola, "The software radio architecture," *IEEE Communications Magazine*, 1995.
- The SoapySDR project, for the abstraction layer described in section 7.
- ISO/IEC 17025:2017, on the competence of testing and calibration laboratories.
- ICX-FieldHawk handheld, rugged and USB datasheets, berkeleynucleonics.com.

If you are weighing an open instrument against building on bare software-defined radio hardware, our application engineers would be glad to talk the measurement through, including the cases where the bare hardware is the right answer.

VERIFICATION NOTE

The following values in this paper are not yet confirmed against a published Berkeley Nucleonics datasheet and are marked verify in the text. They must be confirmed before this paper is released.

- The SoapySDR and GNU Radio integration in section 7 is documented in a platform integration guide and demonstrated with captured results, and in an August 2026 review the platform partner confirmed the integration working on ICX series units. It does not yet appear on a datasheet, so treat it as demonstrated and partner-confirmed rather than specified. The distribution path for the driver package is not settled, so no package name, repository or driver string is printed here.
- Tracking generator coverage per model, and IQ file and bit format details, are not currently published.
- The section 7 demonstrations are recorded in a platform integration guide. The field evidence in section 8 reports buyer feedback from account history rather than a published source, and no end user is named.

Berkeley Nucleonics Corp.
Test, Measurement and Nuclear Instrumentation since 1963
For details: info@berkeleynucleonics.com or 800-234-7858

BNC-AN-101 Rev B
August 2026