Analog and Power IC Test Development: A Field Guide to Pitfalls and Good Practice

A SiliconEng white paper · October 2026 · Version 1.8

1. Why analog and power test is different

Digital test is mostly a question of logic: did the right bit come out at the right time? Analog and power test is a question of measurement: is this voltage, this current, this frequency really what the datasheet says it should be — and can you prove it with the instrument you are measuring it with?

That distinction changes everything about test development. In digital test, the tester is a stimulus and a judge. In analog and power test, the tester is part of the measurement itself. Every source, digitizer, relay, cable, socket contact, and inch of loadboard copper between the tester's resources and the device under test (DUT) adds error to the number you record. If you do not budget that error, you are not testing the part — you are testing your setup, and calling the result the part's performance.

This paper is a companion to our earlier paper, *From Datasheet to ATE* (referred to below as Paper 1), which covered the method: extraction, normalization, test plan, limits. This one covers the traps that method runs into in practice, using two running examples: a step-down (buck) switching regulator, the Texas Instruments TPS54360, and a precision instrumentation amplifier, the Texas Instruments INA333. Both are ordinary, well-documented parts. The pitfalls below are ordinary too — which is exactly why they keep catching experienced engineers.

2. The first rule: the tester is in the measurement

2.1 The 10:1 rule of thumb

Metrology has a long-standing rule of thumb: the accuracy of your measurement should be at least ten times better than the tolerance you are verifying. This is the test uncertainty ratio (TUR), and 10:1 is the target. A 4:1 ratio is generally considered the minimum for a defensible production test; below that, you are making decisions with an instrument nearly as uncertain as the thing it judges.

Put numbers on it with the INA333. Its input offset voltage is specified with a 1/G term: ±(10 + 25/G) µV typical, ±(25 + 75/G) µV maximum. At a gain of 100 the maximum is 25.75 µV — and because the INA333 is itself the gain stage, that offset appears at its output already multiplied: 25.75 µV × 100 = 2.575 mV. A 10:1 TUR on the 25.75 µV maximum means an input-referred uncertainty of about 2.6 µV, or equivalently 0.26 mV at the output at a gain of 100. A digitizer specified at ±0.5 mV (an illustrative figure) measuring that 2.575 mV gives a TUR of about 5:1 — short of the 10:1 target, which would need ±0.26 mV. The problem is real, but it is a 5:1 problem, not the far worse ratio you would get by trying to resolve the 25 µV directly at the input: nobody measures it that way, because the DUT has already amplified it. The test is hard not because the number is small, but because the digitizer's error is a large fraction of the amplified number.

The TPS54360 makes the same point from the other direction. Its internal voltage reference is 0.8 V, specified from 0.792 V to 0.808 V over TJ = −40 to 150 °C — a ±1% band, 16 mV wide. A 10:1 TUR asks for about 0.8 mV of measurement uncertainty — the band is ±8 mV, and the convention in both papers is the half-width (see §2.2). That is achievable on a good digitizer — until you remember the reference is not what production tests. Production tests the output voltage, which is set by a resistor divider off that reference, and the divider's tolerance, the load regulation, and the contact resistance in the output path all join the error budget. The datasheet number is never the number you measure. The number you measure is the datasheet number plus everything between the die and your digitizer.

2.2 Build the uncertainty stack before you write the test

For every precision test, write down the uncertainty stack before you commit to a test method. The entries, in the order they usually bite:

1. **Source accuracy.** The voltage or current you force is only as accurate as the source that forces it. A forced input that is off by 1 mV is a 1 mV error in every measurement downstream of it.

2. **Digitizer accuracy and resolution.** Accuracy is not resolution. A digitizer can resolve microvolts and still be inaccurate by millivolts if its gain and offset errors are uncorrected. Read the accuracy specification at the range, temperature, and aperture you will actually use — not the headline resolution.

3. **Noise.** Random noise can be averaged down (Section 3.4); it cannot be ignored, and it sets the floor for how long a test must take to reach a given uncertainty.

4. **Contact and path resistance.** Every relay contact, socket pin, and trace between the tester and the DUT drops voltage when current flows. At 3.5 A of buck output current, even 10 mΩ of path resistance is 35 mV — more than twice the entire TPS54360 reference band. Kelvin (force-and-sense) connections exist for exactly this reason, and they only work if the sense lines carry no current and connect at the right point (Section 5).

5. **Thermal effects.** Offsets drift with temperature, and so do your instruments, your loadboard, and your socket. A measurement made at a different temperature than the calibration that supports it carries an error nobody wrote down.

6. **The loadboard itself.** Divider resistors, sense resistors, filter capacitors, and any active circuitry you add to the board all have tolerances and temperature coefficients. They are part of the instrument now.

Two definitions are used consistently in this paper and its companion. TUR (test uncertainty ratio) = tolerance ÷ expanded measurement uncertainty U, with coverage factor k = 2 (approximately 95% confidence). (For a worst-case linear sum, the TUR is more conservative than the k = 2 figure.) Here 'tolerance' means the half-width: the ± figure for a two-sided specification, the limit value itself for a one-sided specification. (A 10:1 TUR on the INA333's ±25.75 µV maximum therefore allows 2.6 µV of uncertainty, and a 10:1 TUR on the TPS54360's ±8 mV reference half-band allows 0.8 mV.) The target is 10:1; the minimum for a defensible production test is 4:1, with a documented guardband. If the stack adds up to more than a tenth of the tolerance, you are below the 10:1 target: either improve the measurement (Section 3) or accept a documented, guardbanded test at no worse than 4:1 (Section 3.7). A 4:1 TUR consumes a quarter of the tolerance in measurement uncertainty, which is why it is a minimum, not a target.

2.3 What happens when you skip this

Testing with an inadequate TUR does not produce random failures. It produces confident wrong answers. Parts near a limit get passed or failed depending on the tester's error that day, not on their own performance. Yield looks noisy for no reason anyone can find, because the noise is in the measurement, not the silicon. Worse, the guardband you apply to protect against escapes (see Paper 1 §5) gets set against the wrong uncertainty, so you either ship marginal parts or throw

away good ones — and the data will not tell you which, because the data is the problem.

3. When you cannot get 10:1: ways around it

Sometimes the honest answer to the uncertainty stack is that the tester, as configured, is not good enough. That is not a reason to ship the test anyway. It is a reason to change the measurement. These are the workarounds that actually work, in rough order of cost.

3.1 Spot calibration: calibrate at the test point, not at the factory

A tester's accuracy specification covers its whole range, over temperature, over time since its last calibration. Your test uses one point. Spot calibration exploits that gap: before (or interleaved with) the production test, you measure a known reference at exactly the voltage, current, and temperature the test will use, and you correct the channel's gain and offset error at that point. What remains is drift since the spot calibration, noise, and repeatability — often substantially smaller than the factory accuracy specification.

Spot calibration is powerful and it has rules. Calibrate at the operating point, not at a convenient one: a channel calibrated at 5 V does not stay calibrated at 50 mV. Re-do it often enough that drift cannot grow back — per insertion, per lot, or per temperature change, depending on what your data shows. And record it: a spot calibration is part of the test method, and a test method that depends on an undocumented calibration step will silently break the first time someone "simplifies" the program.

3.2 Bring in an external source or instrument

When the tester's own resources cannot reach the required accuracy, bring a better instrument to the loadboard. A laboratory-grade digital multimeter or a precision voltage source, switched into the measurement path through the relay matrix, can serve as the reference for the tests that need it — the INA333 offset test, the TPS54360 reference test — while the tester's own resources handle everything else.

This costs test time (relay switching, settling, instrument communication) and it costs discipline. The external instrument must itself be calibrated and its calibration kept current. Its connection to the loadboard must be Kelvin where it matters, and its ground must be the DUT's ground, not a convenient chassis point — a ground offset between the external instrument and the tester is an error

injected straight into the measurement. And the correlation between the external instrument's reading and the tester's reading on the same part must be established during development, so that production does not depend on an instrument that disappears when the bench unit goes back to the lab. Used with those disciplines, an external instrument is the fastest honest fix for a TUR problem.

3.3 Gain it up: a precision amplifier on the loadboard

The INA333 offset problem is sometimes described as 'gain it up,' but the description needs care: the INA333 is itself the gain stage, so its offset can only be observed at its output, already multiplied by G. What a loadboard stage can do is post-amplify: at G = 100 the DUT presents up to 2.575 mV of offset, and a further ×10 post-amp presents about 25.8 mV to the digitizer, so a ±0.5 mV digitizer error becomes about 0.5 µV input-referred instead of 5 µV. The generic principle still applies wherever the signal has no on-chip gain: current-sense voltages, error-amplifier outputs, the difference between two nearly-equal voltages — there, a first gain stage on the loadboard is what makes the measurement possible at all.

The post-amp is not free accuracy. Its own offset, drift, noise, and gain error join the budget, so it must be a precision part in its own right, and its gain must be known — measured, not assumed from resistor values. Its bandwidth and settling must cover the signal: at G = 100 the INA333 itself is only 3.5 kHz and needs 400 µs to settle to 0.01% (500 µs to 0.001%, 4 V step), so a post-amp stage needs to be comfortably faster than that, and the wait before measuring must respect the slower of the two. Its stability must be verified in the actual layout, with the actual source impedance driving it — an amplifier that is stable on the bench can oscillate when it sees the capacitance of a socket and a relay matrix. Simulate it first (Section 6), characterize its gain and offset across temperature, and treat it as part of the instrument: it gets calibrated, and its calibration gets checked.

3.4 Average down the noise — and know what averaging cannot fix

Random noise falls with the square root of the number of samples averaged: 100 samples buy a factor of 10. For a noise-limited measurement, averaging is nearly free accuracy, paid for in test time. For a systematic error — a gain error, an offset, a contact drop, a ground shift — averaging buys nothing. A thousand samples of a wrong number is still a wrong number, now with a deceptively tight standard deviation.

Averaging also has a time cost that matters on a production tester, and a subtle failure mode: if the quantity being measured drifts during the averaging window (self-heating is the usual suspect on power parts), the average is not the value at any defined condition. Average over a window short enough that the DUT is thermally stable, or control the thermal condition first and average second.

3.5 Kelvin sensing, nulling, and bracketing

Three classical techniques, each worth having in the toolbox:

- **Kelvin (four-wire) sensing.** Force current through one pair of contacts, measure voltage through a separate pair that carries no current. The contact and trace resistance drops out of the measurement. This is non-negotiable for the buck's output voltage and current paths, and for any low-voltage measurement at real current. The sense connection must be made at the point whose voltage you actually care about — at the DUT pin, not at the socket, not at the board edge. The rule is: sense at the point the specification is defined at. For most pins that is the DUT pin. For the buck's output voltage, the specified point is the output node — the output capacitor in the application circuit — so sense there, and state the sense point in the test definition.
- **Nulling.** Measure the setup's own offset with zero stimulus (inputs shorted, source set to zero), then subtract it from the measurement. This removes the fixed part of the error — the tester's offset, the preamp's offset, the thermal EMF of one junction — and leaves only drift and noise. Null often, because offsets drift.
- **Bracketing (binary search).** For threshold tests — the TPS54360's enable threshold (1.2 V typical, 1.1–1.3 V limits), undervoltage lockout, overvoltage protection — do not measure the threshold with one coarse step. Binary-search it with a DAC: each step halves the uncertainty, and ten steps buy a factor of a thousand. Threshold tests are where a crude method wastes the most accuracy.

3.6 Correlate, don't duplicate

Not every production test has to reproduce the datasheet measurement. For parameters where the bench can measure precisely and the tester cannot, the honest production strategy is correlation: characterize the parameter on the bench with precision instruments, find a production-screenable proxy that tracks it (a related voltage, a go/no-go test at a guardbanded limit, a room-temperature measurement that predicts the hot result), prove the correlation on real parts

across lots and temperatures, and document it. The datasheet parameter is then covered by a documented argument, not by a bad measurement pretending to be a good one. What you must never do is run a bad measurement, get a number, and file it as if it meant something. In the disposition vocabulary of Paper 1 §3.3, this is 'tested via correlated proxy.'

3.7 If the TUR is still short, say so and guardband for it

When every workaround is spent and the ratio is still short of 10:1, the remaining honest move is to widen the guardband by the full measurement uncertainty and accept the yield cost — and to write down that the test is measurement-limited. A test that passes parts only when they clear the limit by more than the uncertainty is a test you can defend. A test that pretends to 10:1 with a 2:1 instrument is a test that will embarrass you in front of a customer, usually at the worst possible time. In Paper 1 §3.3 this is the 'tested, measurement-limited (guardbanded)' disposition, and the guardband is GB as defined in Paper 1 §5: the measurement uncertainty U, which that paper's worked example combines as a worst-case linear sum (more conservative than a k = 2 RSS).

4. Two worked examples: where the traps actually are

4.1 The buck regulator (TPS54360): power makes every error bigger

The TPS54360 is a 60 V, 3.5 A step-down converter with an integrated high-side MOSFET, a 0.8 V ±1% reference (specified over TJ = −40 to 150 °C), cycle-by-cycle current limit (5.2–5.85 A at 25 °C and 12 V input; 4.5–6.8 A over all input voltages and temperatures, 4.5–6.25 A at 12 V), and a switching frequency set by a resistor (450–550 kHz with a 200 kΩ timing resistor). Each of its headline tests has a characteristic trap:

Output voltage accuracy: test the chain, not the reference. The output is the reference times the divider ratio, minus load-regulation effects. The divider resistors on the loadboard are part of the measurement: 1% application resistors contribute up to ~2% worst-case ratio error (about 1.5% at the output for a 3.3 V output), which alone exceeds the reference's own ±1% band. For the production test, use 0.1% divider resistors on the loadboard, Kelvin-sense the output at the output node (the DUT's output capacitor — the point the specification is defined at; see §3.5), and measure at defined load currents — output voltage at no load and at 3.5 A are different tests with different error budgets, because contact resistance and self-heating scale with current.

Current limit: replicate the condition, or correlate to it. The datasheet specifies the current limit under open-loop conditions — the error amplifier driven so the converter is not regulating. A production test that simply loads the output until the converter gives up is measuring something else: inductor saturation, thermal limiting, or the load's own behavior. Either replicate the open-loop condition (force the feedback node, measure the peak switch current with a fast shunt and differential amplifier), or characterize the open-loop limit on the bench and correlate it to a closed-loop production screen — and document which you chose. The datasheet's footnote to the current-limit rows states that the open-loop limit is measured directly at the SW terminal and is independent of inductor value and slope compensation, so the threshold itself does not depend on the inductor. In the application, though, the current-limit comparator has a 60 ns delay, and during that delay the inductor current keeps rising (by (VIN − VOUT)/L × 60 ns — about 52 mA for a 12 V input, 3.3 V output, and a 10 µH inductor, for illustration), so the peak current the test sees depends on the inductor value. Note the range as well: 5.2–5.85 A applies at 25 °C and 12 V only; over all input voltages and temperatures the datasheet allows 4.5–6.8 A, so a production test run hot must use the window that matches its input voltage: 4.5–6.25 A at 12 V over temperature, or the full 4.5–6.8 A if VIN is varied over the datasheet range.

Switching frequency and timing: bandwidth is an accuracy spec. A 500 kHz switching waveform with a 135 ns minimum on-time (typical, VIN = 12 V, 25 °C) needs a measurement path with real bandwidth: the digitizer, the relay path, and the probing all have to pass the edges you are timing. A slow channel does not measure frequency wrong by a little; it measures a filtered, slew-limited version of the waveform, and the error shows up as frequency and duty-cycle shifts that look like part variation. Measure frequency with a counter or a fast digitizer channel, and verify the path bandwidth with a known signal before trusting it on the DUT.

Efficiency: four measurements, compounding errors. Efficiency is output power divided by input power — two voltages and two currents. If each measurement carries 0.5% error, the efficiency result carries roughly 1% (errors add in quadrature for a ratio of products, and systematic errors add directly). A 1% efficiency uncertainty is large against the 1–2% efficiency spread that typically separates a good part from a mediocre one, and it is almost entirely a measurement story — with one qualification: errors from a shared shunt or instrument are partly correlated and can cancel, so the quadrature figure is a

planning estimate, not a guarantee. The budget must also cover shunt tolerances, Kelvin discipline, and simultaneous sampling (input and output must be measured at the same operating point, not sequentially across a drifting thermal state). If efficiency is a production screen, budget it as a four-measurement problem and say so.

Thermal: the part heats itself while you watch. At 3.5 A, the TPS54360 dissipates real power, and its reference, RDS(on), and current limit all move with junction temperature. A test that measures a cold part, then a hot part, and calls the difference "variation" is measuring its own test sequence. Define the thermal condition (soak time, load duration, handler chuck temperature), keep it consistent, and remember the socket's thermal resistance is part of the condition — a socket that insulates the package is a different thermal environment than the application board.

4.2 The instrumentation amplifier (INA333): small signals, hostile environment

The INA333 is a low-power precision instrumentation amplifier: input offset VOSI ±(10 + 25/G) µV typical, ±(25 + 75/G) µV maximum; drift ±(0.1 + 0.5/G) µV/°C typical; CMRR 100 dB minimum at gain ≥ 10 (80 dB at gain 1); input bias current ±70 pA typical, ±200 pA maximum; gain set by a single external resistor from 1 to 1000. In this paper, 'typical' means the datasheet typical value (usually at 25 °C), not a guarantee; ranges quoted without qualification are min/max. Every one of those numbers is a measurement challenge, and the challenges are different from the buck's:

Offset: measure at gain, refer back to the input. The offset specification is input-referred with a 1/G term: VOSI is ±(10 + 25/G) µV typical and ±(25 + 75/G) µV maximum, so at a gain of 100 the maximum is 25.75 µV, appearing as 2.575 mV at the output (the /G term is the output stage's offset referred back to the input — it is already inside the specification, not an addition to it). Report offset input-referred (output divided by measured gain, not nominal gain — the gain resistor's tolerance is otherwise silently folded into the offset result; at G = 100 the datasheet gain error is ±0.07% typical, ±0.25% maximum, so using nominal gain misstates the offset by up to 0.25% of reading — small for offset itself, material for larger signals). And short the inputs properly for the offset test: a "short" made of a relay contact with thermal EMF and a few ohms of resistance is not a short at the microvolt level.

Drift: the setup must drift less than the part. Drift is specified as ±(0.1 + 0.5/G) µV/°C typical — about 0.105 µV/°C at a gain of 100 — so over the −40 °C to +125 °C range the part moves about 17.3 µV typical input-referred. That is smaller than the offset itself, and smaller than the thermal EMF of a single dissimilar-metal junction in your relay matrix if that junction sees a temperature gradient. Drift testing demands an isothermal setup: keep relay and socket junctions at the same temperature, minimize the number of junctions in the measurement path, use copper-to-copper connections where possible, and measure the setup's own drift with a zeroed input before believing the part's drift. If your setup drifts 0.5 µV/°C, you cannot measure a 0.105 µV/°C part, no matter how good the digitizer is.

CMRR: 100 dB is a ratio, and ratios expose mismatch. The specification is DC to 60 Hz, with the common-mode voltage swung from (V−) + 0.1 V to (V+) − 0.1 V — at a 5.5 V supply the maximum swing is about 5.3 V, so a 5 V common-mode step fits only at the top supply and is impossible at lower supplies. And the minimum is 100 dB only at gain ≥ 10; at gain 1 it is 80 dB. 100 dB means rejection by a factor of 100,000 — a 5 V common-mode step is allowed to produce only 50 µV of input-referred error. The test's enemy is not the amplifier; it is asymmetry in the test circuit. A fraction of a percent of mismatch in the source resistors, a ground offset between the common-mode source and the measurement reference, or a relay contact difference between the two input paths all masquerade as CMRR failure. One asymmetry deserves its own warning: the REF pin must be driven by a low impedance or grounded — the datasheet tolerates 15 Ω or less of stray resistance in the REF path while maintaining specified CMRR, but tens of ohms in series with REF cause noticeable CMRR degradation, and a relay contact or a long trace in the REF path is exactly such a resistance. Drive both inputs from a single, well-matched source, Kelvin-reference everything to one ground point, and prove the setup's own CMRR with a shorted-input or known-good device before testing production parts.

Bias current: leakage is the measurement. Bias current is ±70 pA typical, ±200 pA maximum — and 200 pA is 200×10^{-12} amps. The leakage of a dirty socket, a flux residue on the loadboard, a humid relay, or an unguarded PCB trace can exceed it by orders of magnitude. Bias-current testing needs a clean board (cleaned, and kept clean), guard traces around the input nodes driven to the input potential, low-leakage relays in the input path, and a settling time that respects the RC time constants of the guarded node. If the bias current numbers look noisy or lot-dependent for no reason, suspect the board and the socket before

suspecting the silicon. The test circuit must also provide a bias-current return path: without one, the inputs float beyond the common-mode range and the amplifiers saturate. The return path (a resistor to a defined potential, or the source itself if its resistance is low) is part of the test circuit design, not an afterthought.

Noise: 50 nV/√Hz needs a quiet room, electrically. Noise testing needs shielding, a quiet supply, bandwidth limiting to the band the datasheet specifies, and an understanding that the tester's own noise floor is part of the result. Measure the setup's noise floor with the inputs shorted and the DUT removed — if the floor is above the part's noise, the test measures the setup. The 50 nV/√Hz figure is specified at 1 kHz, G = 100, with RS = 0 Ω, plus 1 µV peak-to-peak from 0.1 Hz to 10 Hz. A source resistance above zero adds its own thermal noise and a bias-current offset (RS × IB — 10 kΩ × 200 pA = 2 µV, already a tenth of the offset limit), so the noise test's source impedance is part of the specification.

Settling and bandwidth: wait for the part. At a gain of 100 the INA333's −3 dB bandwidth is 3.5 kHz, and it needs 400 µs to settle to 0.01% (500 µs to 0.001%) after a 4 V step; on power-up it needs about 100 µs to reach specified offset accuracy. A test that measures before the part has settled measures a moving target — and a post-amp or digitizer faster than the DUT does not fix a DUT that has not settled. Build the wait into the test timing, and let the settling time, not the handler speed, set the pace.

5. Socket selection and layout: the part of the instrument you can touch

Test engineers argue about tester resources. They should argue more about the socket and the loadboard, because that is where most analog and power test programs quietly lose their accuracy.

Contact resistance and current. A buck delivering 3.5 A through a socket contact rated for 1 A is not a test setup; it is a heater with a measurement problem. Parallel contacts for high-current paths, derate for temperature, and measure the contact resistance — then measure it again after a thousand insertions, because contacts wear, oxidize, and loosen, and a test program qualified on a fresh socket will drift as the socket ages. Put socket inspection and replacement on a schedule, and treat a contact-resistance excursion as a test-program event, not a maintenance footnote.

Kelvin discipline in copper. Kelvin sensing only works if the sense traces carry no current and connect exactly where the voltage matters. On the layout, that means separate force and sense traces, routed together so they see the same environment, connected at the DUT pin — not at the socket body, not at a via three centimeters away. A Kelvin connection made at the wrong point is worse than no Kelvin connection, because it reports a precise measurement of the wrong node. The rule, as in §3.5: sense at the point the specification is defined at — normally the DUT pin; for the buck's output, the output node.

Inductance in fast, high-current loops. The buck's switch node slews amps in tens of nanoseconds. Every nanohenry of loop inductance in the switch path, the gate-drive path, or the ground return rings, overshoots, and radiates — and the ringing shows up in the measurements as noise, as apparent instability, or as outright damage (Section 8). Keep power loops short and tight, put the input and output capacitors close to the DUT, give the switch current a continuous ground return under the loop, and simulate the loop inductance before committing the layout (Section 6).

Leakage and capacitance on precision nodes. The INA333's inputs are high-impedance, low-current nodes. A guard ring driven to the input potential, no solder mask or flux residue bridging the guarded node, minimal dielectric near the inputs, and short, shielded routing are not refinements — they are what makes a 200 pA bias current measurable at all. The same layout care applies to capacitance: stray capacitance on a high-impedance input, combined with the source resistance, is a low-pass filter that changes settling time and noise, and it differs from socket to socket if the layout is sloppy.

Thermal path. The socket is the part's heatsink during test, whether you designed it to be or not. A socket with poor thermal contact lets a power part self-heat beyond its application condition, shifting every temperature-sensitive result; a socket with a good thermal path but no thermal control makes the result depend on the handler, the room, and the previous part. Decide what the thermal condition is supposed to be, design the socket and handler interface to produce it, and verify it with a thermal measurement on a real part — not with an assumption.

Mechanical reality. Pin sequencing (ground mates first), insertion force, alignment tolerance, and the wear life of the contact system are test-program parameters, not purchasing details. A socket that damages pins, or that requires

an operator to force a part, will produce a failure signature that looks exactly like a silicon problem. Qualify the socket the way you qualify any other instrument: with data, on the production setup, before release.

6. Simulate before you cut copper

Simulation tools — LTspice (Analog Devices), TINA-TI (Texas Instruments), PSpice, and SIMPLIS for switching converters — are free or free-tier, fast, and embarrassingly underused during loadboard and test-hardware design. LTspice and TINA-TI are free; PSpice for TI is free in a size-limited Lite edition; SIMPLIS offers a restricted free version. They should be used at three points:

First, simulate the test circuit, not just the DUT circuit. The vendor's model of the TPS54360 or INA333 tells you how the part behaves. What you need to know is how the part plus your loadboard behaves: the preamp stage's gain, offset, and stability with the real source impedance; the Kelvin loop's behavior with real trace resistance; the switch-node ringing with real loop inductance; the CMRR test's sensitivity to resistor mismatch. Put the parasitics in the model — socket inductance, relay capacitance, trace resistance — because the parasitics are the measurement.

Second, simulate the transients you are afraid of. The spike check of Section 8 should not be the first time you see what relay sequencing does to a pin. Simulate power-up, power-down, relay switching under load, and fault recovery, and look at every pin's voltage against the absolute maximum ratings. A simulation that shows a 20 V overshoot on a 6 V pin is a cheap lesson; the same overshoot on a real part is a damaged part and a mystery failure.

Third, use simulation to choose, not just to confirm. When two test methods are on the table — direct measurement versus gain-up preamp, shunt versus current probe, one socket versus another — simulate both and compare their error budgets. Simulation is how you spend accuracy wisely, before the board exists.

Simulation has limits, and they must be stated plainly. Vendor models are typically typical-value models, not min/max corners — you must add your own corners and tolerances. Models do not include your socket's wear, your relay's thermal EMF, or your operator's habits. Simulation tells you where measurement will hurt; only measurement tells you how much. Use it to design the test, then verify the test with data.

7. Gauge R&R and the deep dive: proving the numbers mean something

7.1 Gauge repeatability and reproducibility

A test program is a measurement system, and measurement systems get qualified with gauge repeatability and reproducibility (GR&R) studies. Repeatability is the variation when the same part is measured repeatedly on the same setup; reproducibility is the variation across different setups, operators, handlers, or sockets. Together they answer the only question that matters before release: how much of the spread in my data is the parts, and how much is my measurement?

The standard design is simple: a handful of parts spanning the range of interest, measured several times each, on each setup you intend to use, by each operator or handler configuration you intend to run. Analyze by ANOVA, not by eyeballing. The accepted yardsticks (% of tolerance, per the AIAG Measurement Systems Analysis manual, 4th edition) are: %GR&R under 10% is good; 10–30% is conditional — acceptable only with a documented reason and, usually, a tighter guardband; over 30% is not acceptable, and no amount of guardbanding fixes a measurement system that cannot distinguish parts from itself. Note the relationship to TUR (§2): TUR compares the half-width tolerance to expanded uncertainty (k = 2), while %GR&R compares the full tolerance span to the measurement-system variation (6σ here). A 10:1 TUR corresponds to roughly 15% GR&R — inside the conditional band — so a GR&R under 10% implies a better TUR (roughly 15:1), and a setup at the 4:1 TUR minimum (about 37% GR&R) will generally not pass a GR&R study. (Some references and older editions use 5.15σ, which would give about 13% and 32% for the same two cases. State which multiplier your GR&R software uses.) The two checks must agree before release; neither replaces the other. A second yardstick, the number of distinct categories (ndc), should be at least 5 — the system must be able to resolve at least five distinct levels of the measured quantity, or it is a go/no-go gauge wearing a measurement costume.

Run GR&R on the production setup — handler, socket, program, and all — not on the bench. A bench GR&R that passes and a production GR&R that fails is the normal case, because the handler and socket are where the variation lives. Run it before release, and re-run it after anything that touches the measurement: a socket change, a handler change, a program revision, a loadboard revision. A

GR&R study is cheap. A production release with an unqualified measurement system is expensive, and the bill arrives as yield arguments nobody can resolve.

7.2 The deep dive: reading data like a test engineer

Once the program runs, the data itself is a diagnostic instrument. Before declaring that "the numbers look good," look for these things:

- **Distributions, not just pass/fail.** Plot histograms. A clean, single-peaked, well-centered distribution is what good looks like. A bimodal distribution is two populations — two sites, two lots, two handlers, a program that changed mid-run — and the average of two populations describes neither of them.
- **Clipping.** A distribution piled up against a test limit is not a tight process; it is a measurement or a limit that is cutting off real variation. Find out which.
- **Site, handler, and temperature deltas.** The same part measured on different sites or at different temperatures should agree within the measurement uncertainty. A systematic site delta is a setup difference — a contact, a ground, a thermal condition — and it will become a yield difference the moment a lot leans on one site.
- **Correlation with the bench.** Production results should track bench characterization within an agreed delta, established during development on the same parts. A production-to-bench gap that grows over time is a setup drifting, and the drift is usually in the socket or the calibration.
- **Stability over time.** Control charts, not snapshots. A mean that walks, a spread that grows, or a retest rate that creeps up are early warnings — of socket wear, of a calibration interval that is too long, of a handler that needs maintenance.
- **Outliers, handled honestly.** Investigate every outlier; never delete one without a cause. An outlier with a cause (a damaged socket pin, a misloaded part) is a setup finding. An outlier without a cause is information about the tails of your distribution, and the tails are where escapes live.
- **Retest behavior.** Parts that fail, then pass on retest, are telling you the measurement is marginal at that limit. Track the retest-pass rate per test, not just per lot: it points straight at the tests whose TUR is too low or whose guardband is too tight.

"Numbers look good" is a conclusion, not an impression. It means: distributions are clean and centered, variation is dominated by the parts rather than the measurement, sites and temperatures agree, bench correlation holds, trends are

flat, and the outliers have explanations. If any of that is missing, the numbers do not look good yet — they look unexamined.

8. The spike check: every pin, every transition

Of all the practices in this paper, this is the one most often skipped and most often regretted.

What it is. During power-up, power-down, relay switching, handler insertion and removal, enable toggling, mode changes, and fault recovery, pins see transient voltages and currents that no steady-state test will ever show: inductive kickback, relay bounce, charge injection from a switching relay, a supply rail that rises before the ground is connected, a signal driven into an unpowered part. These spikes can exceed absolute maximum ratings for microseconds — long enough to damage a junction, trigger latchup, or degrade a part that then passes every functional test and fails in the field.

How to do it. Put an oscilloscope on the DUT pins — at the DUT, not at the tester — and capture single-shot transients through every transition the program and the handler can produce. Check all pins, or at minimum every supply pin, every high-impedance input, every enable and control pin, and every output during the transitions that touch it. Compare every captured transient against the absolute maximum ratings with margin, in both polarities. Do this during development, and repeat it after any program, loadboard, socket, or handler change that touches sequencing or relay control. A spike check is not a one-time event; it is part of the change control for the test program.

Two examples make the point concrete — these are the datasheet absolute maximums the spike check is protecting:

Pin	Absolute maximum (datasheet)	Typical threat
TPS54360 EN	−0.3 to 8.4 V	Enable driven from a higher rail, or a relay transient
TPS54360 SW	−0.6 V (−2 V for a 10 ns transient) to 65 V	Switch-node undershoot from loop inductance
TPS54360 FB, COMP	−0.3 to 3 V	Low-voltage analog pins exposed to relay transients
INA333 inputs	(V−) − 0.3 V to (V+) + 0.3 V; beyond-rails current limited to 10 mA	Signal driven into an unpowered or differently-powered part
INA333 supply	7 V max	Supply overshoot on power-up

How to fix what you find. The fixes are simple, which is why there is no excuse for skipping the check:

- **Sequence power before signals, and ground first.** Supplies up before signal sources are connected; ground connected before anything else, via socket pin lengths or explicit sequencing. A signal driven into an unpowered part forward-biases protection diodes and is the single most common spike mechanism.
- **Never hot-switch relays.** Switch relays only with the path de-energized; use break-before-make sequencing, and add a delay between relay action and the next stimulus or measurement so contacts settle and bounce ends.
- **Add series resistance on signal lines.** A small series resistor limits the current a transient can drive into a pin, and combined with the pin capacitance it forms a filter that rounds off the fastest edges. On precision inputs, size the resistor carefully: it adds a bias-current offset (R × IB — 100 kΩ × 200 pA = 20 μV on the INA333, nearly the offset limit itself) and noise, so a resistor that protects the pin can corrupt the measurement it was meant to protect.
- **Clamp where it is safe.** Schottky diodes to the rails clamp overshoot — but watch their leakage on precision nodes, where a clamp diode's leakage can exceed the bias current you are trying to measure. Clamp the power and control pins freely; clamp precision inputs only with leakage in mind.
- **Snub inductive paths.** RC snubbers across inductive loads and relay coils absorb the kickback that would otherwise appear at the pin.
- **Control slew and soft-start.** Ramp supplies instead of slamming them, use the part's own soft-start where it exists, and ramp the tester's sources at a defined rate rather than stepping them.
- **Provide discharge paths.** Capacitors and inductors store energy; give it somewhere to go before a relay connects a charged node to a fresh pin.
- **Read the relay table line by line.** Most spike problems are visible in the program's relay sequence before any hardware runs. Review it as a sequence of electrical events, not as a list of settings.

A part damaged by a spike does not always fail. Sometimes it degrades — a little more leakage, a little less margin — and passes into a shipment. The spike check is the only practice in this paper that protects against damage you cannot see in the test data. Do it on every pin, on every transition, before release.

9. Good practice during test development: the checklist

- Budget the measurement uncertainty before writing the test: target 10:1 TUR, minimum 4:1 with a documented guardband (TUR = tolerance ÷ expanded uncertainty, k = 2; tolerance = half-width).
- Calibrate at the test point (spot calibration), not just at the factory, and record the calibration as part of the method.
- When the tester is not accurate enough, bring in an external instrument, gain the signal up with a precision amplifier, or average — and know which errors each technique can and cannot remove.
- Kelvin-sense every low-voltage and high-current measurement, at the point the specification is defined at (normally the DUT pin).
- Use 0.1% or better loadboard components wherever the loadboard is part of the measurement.
- Replicate the datasheet's test conditions exactly, or document a correlation to them — never silently substitute a different condition.
- Qualify the socket as an instrument: contact resistance, current rating, inductance, leakage, thermal path, wear life, and pin sequencing.
- Simulate the test circuit — DUT, loadboard parasitics, socket, and tester channel — in LTspice, TINA-TI, PSpice, or SIMPLIS before cutting copper, and simulate the transients you are afraid of.
- Run gauge R&R on the production setup before release (% of tolerance, AIAG MSA 4th edition): under 10% target, 10–30% conditional, ndc ≥ 5; re-run after any measurement-affecting change.
- Read the data as distributions: check for bimodality, clipping, site and temperature deltas, bench correlation, drift over time, and honest outlier handling.
- Spike-check every pin through every transition — power-up, power-down, relay switching, insertion, mode changes, fault recovery — with a scope at the DUT, against absolute maximum ratings with margin.
- Sequence power before signals and ground first; never hot-switch relays; clamp, snub, and slew-control where the spike check says you must.
- Guardband by the full measurement uncertainty when the TUR is short, and write down that the test is measurement-limited.
- Document every decision — the workaround, the correlation, the guardband, the spike fix — so the next engineer can disagree with it productively instead of rediscovering it painfully.

10. About SiliconEng

SiliconEng is a silicon engineering services company in California, focused on analog, power management, RF, and mixed-signal ICs. We develop ATE test programs (wafer probe and package), run product qualification (AEC-Q100, MIL-STD-883), and support silicon validation, characterization, and datasheet creation. Our work is built on three principles: competence, customer care, and continuous improvement.

For questions about this paper or our services: www.SiliconEng.com, info@siliconeng.com

Sources

Texas Instruments, TPS54360 datasheet, SLVSBB4G (June 2018); Texas Instruments, INA333 datasheet, SBOS445C (December 2015). Values in the worked examples are from these revisions; later revisions may change limits — re-verify before reuse.

Trademarks and Copyright Notice

© 2026 SiliconEng. All rights reserved.

TPS54360 and INA333 are products of Texas Instruments Incorporated; Texas Instruments and TI are trademarks of Texas Instruments Incorporated. LTspice is a trademark of Analog Devices, Inc. TINA-TI is a trademark of Texas Instruments Incorporated. PSpice is a trademark of Cadence Design Systems, Inc. SIMPLIS is a trademark of SIMPLIS Technologies, Inc. AIAG is a trademark of the Automotive Industry Action Group. AEC-Q100 is a standard of the Automotive Electronics Council; MIL-STD-883 is a standard of the United States Department of Defense. Device specifications cited in this paper are from the manufacturers' published datasheets; refer to the current datasheet for the authoritative specifications. All other trademarks are the property of their respective owners.