AliasProbe: Coherent Acoustic FMCW Range-Response Sharpening on a Stock iPhone via an Aliased Second-Harmonic Band

Anonymous draft for arXiv submission
Project AliasProbe

March 24, 2026

Abstract

Smartphone acoustic sensing is usually treated as bandwidth-limited by the nominal 48 kHz audio path, which leaves only a narrow near-ultrasonic band for inaudible frequency-modulated continuous-wave (FMCW) ranging. This paper studies a different operating regime on a stock iPhone: the built-in speaker is driven with a 16–20 kHz FMCW chirp, the speaker's nonlinear transfer generates a second harmonic at 32–40 kHz, the microphone path leaks a weak portion of that energy past the analog anti-alias filter, and the 48 kHz analog-to-digital converter folds it into an 8–16 kHz aliased down-chirp. We show how to treat the resulting waveform as a second sensing band rather than as a nuisance artifact. The paper provides a complete signal model, a three-stage estimator chain, a simulation environment, and an implementation history on stock iPhone hardware. The three estimator stages are: (i) alias existence testing by chirp-on/guard SNR plus split-half reproducibility; (ii) common-delay verification with a joint generalized likelihood ratio test (GLRT) over the expected target window; and (iii) coherent wideband combination using per-band GCC-β, direct-path calibration, and held-out validation on interleaved pulse-inversion halves. The ideal aperture extension from 4 kHz to 12 kHz implies a $3\times$ best-case ceiling in matched-filter mainlobe width. On real hardware, we report all-stage confirmed held-out improvements above $2\times$ and best-case held-out narrowing in the 2.1–$2.8\times$ range under favorable operating conditions, while also documenting why stable $3\times$ performance is difficult: the second band is weak, drive-dependent, and highly sensitive to pulse-inversion quality and inter-band phase errors. The resulting claim is deliberately narrow and defensible: on stock 48 kHz iPhone audio hardware, the aliased second-harmonic band is a real, verifiable source of additional ranging information that can reduce held-out single-reflector FMCW range-response width by more than $2\times$ relative to the fundamental-only baseline. We do not yet claim a formal two-target Rayleigh-style resolution result; that direct separability experiment is the next step.

1 Introduction

Acoustic sensing on phones is attractive because speakers and microphones are ubiquitous, the propagation speed of sound is low enough to support fine-grained delay measurements, and FMCW processing maps delay into frequency in a way that is computationally friendly on mobile hardware. The standard bottleneck is the fixed audio sample rate: with a stock 48 kHz path, a smartphone can only record up to 24 kHz in principle, and in practice the useful inaudible band is narrower due to roll-off in the speaker, microphone, and analog anti-alias path. Prior work has therefore either stayed within the 18–24 kHz band, exploited nonlinearity for other sensing goals, or reconfigured phones to higher sample rates when the operating system and codec permitted it [1–3].

AliasProbe asks a different question: if the stock iPhone hardware already leaks a weak above-Nyquist harmonic into the sampled signal, can that leakage be turned into useful ranging bandwidth rather than filtered away or ignored? The answer explored here is yes, but only with careful estimator design and strong guardrails against artifacts. The speaker's second-order distortion creates a second harmonic of a 16–20 kHz chirp at 32–40 kHz. Because that energy sits above the 24 kHz Nyquist frequency of the 48 kHz ADC, any residual harmonic that survives the microphone's analog filtering aliases to 8–16 kHz and appears as a reversed-sweep down-chirp. The alias and the fundamental traverse the same acoustic path and therefore carry the same time-of-flight information, even though their amplitudes and phase responses differ.

The key observation is that the stock iPhone effectively emits and records two separable bands from one transmitted waveform:

- the fundamental sensing band at 16–20 kHz; and
- an aliased second-harmonic band at 8–16 kHz.

If those bands can be validated, delay-aligned, and combined coherently, the usable aperture becomes 8–20 kHz instead of 16–20 kHz. In the ideal case, matched-filter width scales inversely with bandwidth, so the 12 kHz combined aperture implies a $3\times$ best-case reduction relative to the 4 kHz fundamental-only band.

That theory is straightforward. The difficulty is turning it into a claim that is scientifically defensible on commodity hardware. Three obstacles dominate in practice. First, the second band is weak because loudspeaker harmonic distortion is substantially below the linear response, even before microphone roll-off. Second, the band is drive-dependent because loudspeaker large-signal behavior changes with operating level [12]. Third, wideband coherent stitching is sensitive to timing and phase offsets between bands; the multiband sensing literature emphasizes that apparent bandwidth gains disappear unless hardware distortions and phase offsets are compensated before concatenation [10, 11].

This paper presents the resulting system and the evidence chain that supports it. The goal is not to claim a magical bypass of sampling theory. The goal is to show that a stock iPhone already contains a measurable physical mechanism—speaker nonlinearity plus weak anti-alias rejection—that can be turned into a real second sensing band. Our contributions are:

1. a complete signal model for why a 16–20 kHz chirp produces a usable 8–16 kHz aliased harmonic band on stock iPhone hardware;
2. a robust three-stage pipeline that separately proves alias existence, common-delay support, and held-out wideband narrowing;
3. a simulation environment and a debugging history that explain why naive implementations fail and why certain design choices are necessary; and
4. an honest, narrow claim: more-than-$2\times$ held-out reduction in single-reflector matched-filter range-response width on stock iPhone hardware, with $3\times$ as the best-case ceiling rather than the current guaranteed median.

Throughout the paper, "resolution gain" is used informally to refer to reduction in matched-filter mainlobe or envelope width for a single reflector. The main real-hardware claim is therefore stated as *range-response width reduction*. A direct two-reflector separability experiment is planned but not yet part of the present evidence.

2 Related Work and Positioning

Three strands of prior work matter most.

Smartphone acoustic sensing beyond nominal audio limits. iChemo showed that off-the-shelf mobile devices can exploit nonlinearity and alias-like phenomena to sense high-frequency ultrasound beyond the devices' nominal capture range, reporting PSD estimation up to 60 kHz on commodity devices [1]. HotNets 2022 showed that hardware nonlinearity can improve sensing granularity on commodity devices including smartphones [2]. PowerPhone went further by reconfiguring smartphone audio hardware and software to operate at 192 kHz, improving sensing resolution from 7 cm to 1 cm and enabling broader ultrasonic applications [3]. AliasProbe is narrower and more conservative than PowerPhone: it does not reconfigure the operating system or codec; instead it exploits a weak but measurable second band already present in the stock 48 kHz path.

Harmonic imaging, pulse inversion, and matched filtering. In ultrasound imaging, second-harmonic extraction is well established, and pulse inversion (PI) is a classical way to cancel the linear component while preserving even-order terms [5]. Harmonic chirp imaging and nonlinear chirp coding show that second-harmonic chirps can be pulse-compressed effectively, but also that matched-filter design and pulse-inversion quality strongly affect mainlobe width and sidelobes [4, 6]. AliasProbe borrows this logic directly: it uses PI to isolate the second-harmonic branch and learns a discrete-time alias template instead of relying on an ideal analytic down-chirp.

Coherent multiband splicing and delay estimation. The radio sensing literature treats coherent multiband splicing as a principled way to extend effective bandwidth without changing per-band sample rate. The benefit is well known: delay-domain resolution scales as $1/B$, but the stitching only works after compensation for phase offsets and hardware distortion [10]. HiSAC provides a recent and especially relevant example, using an anchor path to initialize timing and phase before combining sparse subbands into a super-resolved ranging estimate [11]. In parallel, generalized cross-correlation (GCC) and GCC-PHAT form the classical maximum-likelihood family for time-delay estimation [7, 8]. Weighted GCC-PHAT with a tunable exponent β has also been shown useful in low-SNR ultrasonic localization [9]. AliasProbe combines these ideas: common-delay GLRT for the target lag, anchor-path calibration from the direct path, and per-band GCC-β for controlled sharpening.

Position of this work. The claim here is not that smartphone nonlinearity has never been exploited before. That broader claim would be false in light of iChemo, HotNets 2022, and PowerPhone. The narrower and likely novel claim is that on a stock iPhone running at its stock 48 kHz audio rate, an aliased second-harmonic band can be coherently combined with a fundamental FMCW band to produce more-than-2 $\times$ held-out single-reflector range-response narrowing without external hardware or OS modification.

3 Signal Model and Why a 3 $\times$ Ceiling Exists

3.1 Transmit chirp and second-harmonic generation

Let the transmitted real chirp be

$$s(t) = \cos\!\left(2\pi\left(f_0 t + \frac{\mu}{2}t^2\right)\right), \qquad 0 \le t < T, \tag{1}$$

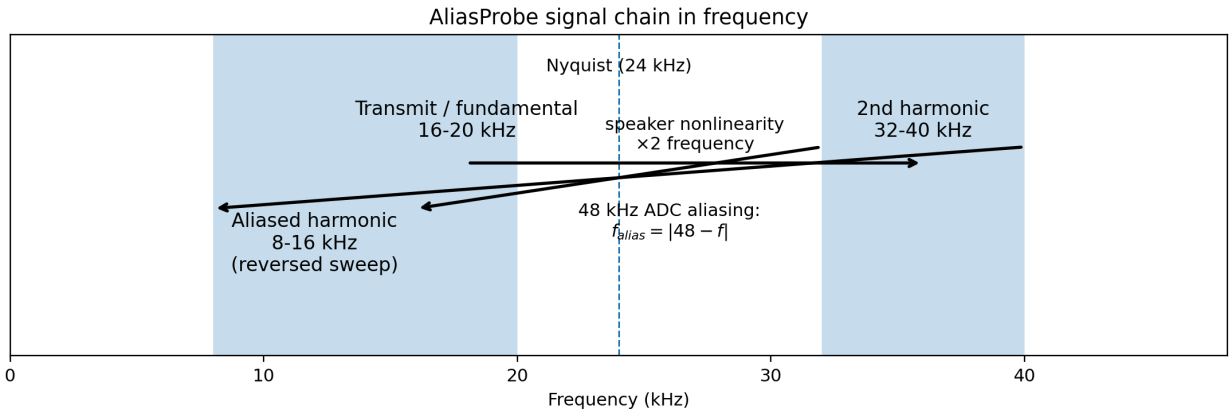

Figure 1: Frequency-domain view of AliasProbe. A 16–20 kHz transmitted chirp produces a second harmonic at 32–40 kHz. Residual energy that leaks past the microphone's analog anti-alias roll-off folds through the 48 kHz ADC into an 8–16 kHz aliased down-chirp.

with $f_0 = 16$ kHz, $f_1 = 20$ kHz, $\mu = (f_1 - f_0)/T$, and $T = 50$ ms. A memoryless second-order loudspeaker nonlinearity is well approximated by

$$y(t) \approx a_1 s(t) + a_2 s^2(t) + a_3 s^3(t) + \cdots . \tag{2}$$

Using $\cos^2\theta = \frac{1}{2}(1 + \cos 2\theta)$, the second-order term contains a second harmonic with phase

$$\phi_2(t) = 2 \cdot 2\pi \left(f_0 t + \frac{\mu}{2} t^2 \right) = 2\pi \left(2 f_0 t + \mu t^2 \right). \tag{3}$$

Its instantaneous frequency is therefore

$$f_2(t) = \frac{1}{2\pi} \frac{d\phi_2(t)}{dt} = 2 f_0 + 2\mu t, \tag{4}$$

which sweeps from 32 kHz to 40 kHz.

3.2 Aliasing into 8–16 kHz

When sampled at $f_s = 48$ kHz, any frequency f above 24 kHz aliases to $|f_s - f|$ if it lies in the first image above Nyquist. The 32–40 kHz harmonic therefore aliases to 16–8 kHz, i.e.

$$f_{\text{alias}}(t) = 48 \text{ kHz} - f_2(t) = 16 \text{ kHz} - 2\mu t, \tag{5}$$

which is a down-chirp sweeping from 16 kHz to 8 kHz.

Figure 1 summarizes the frequency-domain chain.

Proposition 1 (Shared time of flight)**.** *Let the direct or reflected acoustic path introduce a path delay τ that is common to both the fundamental and second-harmonic components. In the frequency domain, the received spectra can be written as*

$$Y_F(f) = H_F(f) S_F(f) e^{-\mathrm{j}2\pi f \tau} + N_F(f), \tag{6}$$

$$Y_A(f) = H_A(f) S_A(f) e^{-\mathrm{j}2\pi f \tau} + N_A(f), \tag{7}$$

where F denotes the fundamental band and A the aliased second-harmonic band. Then both bands carry the same linear delay term $e^{-\mathrm{j}2\pi f\tau}$, even though the complex templates $H_F S_F$ and $H_A S_A$ differ.

Proof sketch. The acoustic propagation delay enters as a phase ramp in frequency, independent of carrier frequency. The band-specific terms $H_F S_F$ and $H_A S_A$ capture loudspeaker, microphone, aliasing, and matched-filter differences, but not the geometric path delay. Therefore, once the band-specific templates are learned or calibrated, both bands are observations of the same delay parameter. □

3.3 Why the ideal ceiling is 3 ×

For a matched or whitened pulse-compression system, mainlobe width is inversely related to effective bandwidth. The fundamental-only system uses 16–20 kHz, i.e. 4 kHz. If the aliased second harmonic from 8–16 kHz can be phase-aligned and combined coherently, the effective contiguous aperture is 8–20 kHz, i.e. 12 kHz. The best-case width ratio is therefore

$$G_{\max} = \frac{12 \text{ kHz}}{4 \text{ kHz}} = 3. \tag{8}$$

The same logic appears throughout multiband splicing: delay-domain resolution improves approximately as $1/B$, where B is the effective spliced bandwidth [10, 11]. In practice, the real gain is lower because the second band is weaker, noisier, and not perfectly stationary.

3.4 Why the alias reference cannot be an ideal analytic down-chirp

A critical debugging result in AliasProbe was that the naive ideal reference

$$s_{\text{ideal}}[n] = \cos\!\left(2\pi \left(16\,\text{kHz} \cdot t_n - \frac{8\,\text{kHz}}{2T} t_n^2\right)\right) \tag{9}$$

had almost no correlation with the actual aliased waveform. The correct reference is not the analytic down-chirp alone; it is the *discrete-time* waveform produced by squaring the transmitted chirp, band-limiting the result to 8–16 kHz, and normalizing it. In other words,

$$s_A[n] = \mathcal{N}\Big(\text{BP}_{8\text{–}16\text{kHz}}\{s[n]^2\}\Big), \tag{10}$$

not a hand-written formula for a down-chirp. This discrete-time template captures the exact phase structure introduced by squaring, finite windowing, and alias folding. That bug fix increased correct-template correlation by orders of magnitude in the implementation and is essential to the claimed result.

4 System Design

4.1 Hardware and waveform

AliasProbe runs on a stock iPhone using the built-in speaker and microphone at 48 kHz sample rate. The transmitted waveform is a 50 ms 16–20 kHz FMCW chirp followed by a 25 ms guard interval. The app transmits 200 cycles per measurement, optionally with a leading silence to simplify onset discovery. The complete analysis runs on-device in Swift using Accelerate/vDSP.

The paper deliberately separates the physical mechanism from the estimator. The mechanism is:

1. speaker nonlinearity creates a second harmonic from the transmitted chirp;

2. the microphone path leaks a weak portion of that harmonic past the anti-alias filter; and
3. the ADC aliases the leaked harmonic into a lower-frequency band that is still time-of-flight consistent with the original path.

The estimator's job is to prove that this lower band is real, deterministic, direction-consistent, and useful.

4.2 Pulse inversion and pair registration

AliasProbe transmits alternating $+s$ and $-s$ chirps. If the loudspeaker output is approximated by a polynomial, then for consecutive cycles

$$y_+[n] \approx a_1 s[n] + a_2 s[n]^2 + a_3 s[n]^3 + \cdots , \tag{11}$$

$$y_-[n] \approx -a_1 s[n] + a_2 s[n]^2 - a_3 s[n]^3 + \cdots . \tag{12}$$

Summing adjacent cycles suppresses odd-order terms and preserves even-order terms; differencing does the opposite. In practice, adjacent cycles are not perfectly aligned, so AliasProbe first registers each pair before forming sum and difference:

$$(a_k, d_k) = \arg\min_{a,d} \left\| w \odot \left(x_k^+ + a\, D_d\{x_k^-\} \right) \right\|_2^2 , \tag{13}$$

where D_d is a fractional delay operator and w is a short window over the direct-path segment. The pairwise channels are then

$$r_{F,k} = \tfrac{1}{2} \left(r_k^+ - a_k D_{d_k}\{r_k^-\} \right) , \tag{14}$$

$$r_{A,k} = \tfrac{1}{2} \left(r_k^+ + a_k D_{d_k}\{r_k^-\} \right) . \tag{15}$$

Averaging 100 such registered pairs gives the fundamental and alias channels used downstream.

4.3 Same-session direct-path calibration and learned templates

The calibration strategy that transferred best in practice was not a rigid external template subtraction. Instead, AliasProbe learns the direct-path complex templates from the same session whenever possible. Let $S_F(f)$ and $S_A(f)$ be the ideal discrete-time references for the fundamental and alias bands. Let $Y_{F,\mathrm{dir}}(f)$ and $Y_{A,\mathrm{dir}}(f)$ be the direct-path spectra extracted from a short low-lag window. The effective templates are

$$T_F(f) = H_F(f) S_F(f), \tag{16}$$

$$T_A(f) = H_A(f) S_A(f), \tag{17}$$

with H_F and H_A learned from direct-path-gated deconvolution and smoothed across the active bins. This structure is directly aligned with harmonic pulse-compression practice, where template mismatch is known to broaden the mainlobe and raise sidelobes [6].

4.4 Three-stage pipeline

Figure 2 shows the full processing chain.

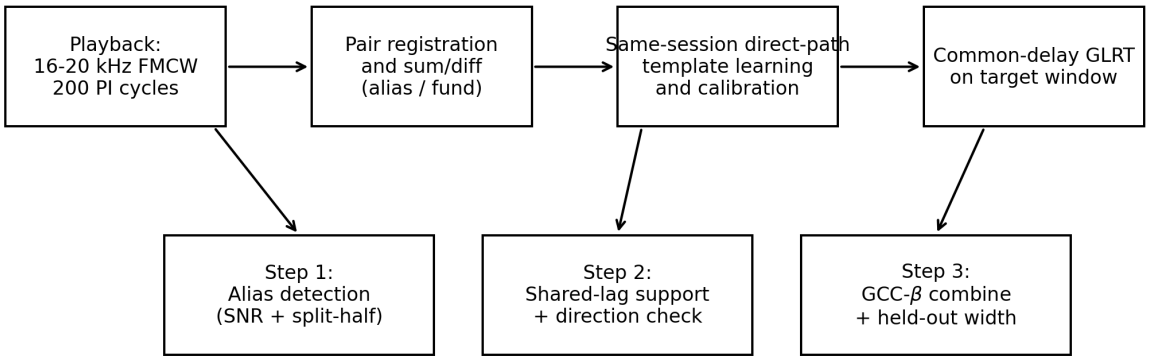

Figure 2: Processing chain used in the current AliasProbe implementation. A run is claimed only if it passes independent alias-existence, common-delay, and held-out narrowing checks.

Step 1: alias existence. The alias channel is band-limited to 8–16 kHz and tested in two ways:

1. chirp-on versus guard-interval power, reported as an alias SNR; and
2. split-half reproducibility, in which odd-indexed PI pairs and even-indexed PI pairs are averaged separately and then correlated.

A passing run must show both nontrivial alias SNR and positive split-half reproducibility.

Step 2: common-delay support. Instead of relying on separate peak picking, AliasProbe computes a shared-lag GLRT over the expected target window $\tau \in [70, 100]$ samples:

$$m_F(\tau) = \sum_{f \in B_F} W_F(f) Y_F(f) T_F^*(f) e^{\mathrm{j}2\pi f \tau}, \tag{18}$$

$$m_A(\tau) = \sum_{f \in B_A} W_A(f) Y_A(f) T_A^*(f) e^{\mathrm{j}2\pi f \tau}, \tag{19}$$

$$J(\tau) = \frac{|m_F(\tau)|^2}{E_F} + \frac{|m_A(\tau)|^2}{E_A}, \tag{20}$$

with E_F and E_A the corresponding normalization terms. The common lag estimate is $\hat{\tau} = \arg\max J(\tau)$. A run passes Step 2 when both bands support $\hat{\tau}$ strongly enough and the alias still prefers the physically correct sweep direction in a local check.

Step 3: wideband GCC-β combining. For each band, AliasProbe forms a GCC-β spectrum

$$G_b(f) = \frac{Y_b(f) T_b^*(f)}{\left(|Y_b(f) T_b^*(f)| + \varepsilon\right)^{\beta}}, \qquad b \in \{F, A\}, \tag{21}$$

optionally after direct-path deflation and calibration rotation. The final coherent profile is

$$P(\tau) = \left| Z_F(\tau) + e^{-\mathrm{j}\Delta\phi}\, Z_A(\tau, \delta\tau) \right|, \tag{22}$$

Table 1: Major implementation failures discovered during development.

Issue	Symptom	Resolution
Wrong alias reference	Correct and wrong sweep directions were nearly indistinguishable; matched-filter energy was tiny.	Generate the alias template by discrete-time squaring of the transmitted chirp, then band-limit to 8–16 kHz.
Onset search in silence	All detected cycle starts landed before the first chirp; sum/diff channels collapsed toward zero.	Expand onset search to include the leading-silence offset.
Inconsistent bin ranges	Different functions used different rounding conventions for FFT bin masks, causing incoherent boundary behavior.	Replace all range calculations with a single canonical integer bin-range routine.
Wrap-around in FFT correlation	Peaks reappeared near the end of the FFT frame, creating ghost lags and impossible Step 3 results.	Increase the FFT size so that linear correlation support fits without circular overlap.
Cross-session calibration drift	A stored external calibration produced wrong phase compensation when phone geometry or volume changed.	Move to same-session direct-path templates and anchor-path phase updates.
Reflector geometry artifacts	Angled phone placement caused the table and room to dominate the target window, pulling the lag to implausible values.	Flatten the phone, face the bottom edge toward the reflector, and keep a fixed geometry.

where Z_F and Z_A are inverse transforms of the band-limited GCC-β spectra, and $(\delta\tau, \Delta\phi)$ are a tightly constrained residual delay and phase correction centered on the same-session anchor-path estimate.

Held-out validation. The decisive safeguard is an interleaved odd/even split. The system fits the residual correction on one half and measures the width reduction on the other, then swaps the roles. A run is only accepted if both held-out directions still show improvement. This prevents the claim from depending on tuning and scoring the same data.

5 Implementation History and Major Debugging Results

The final pipeline emerged only after a long debugging cycle. The most important bugs are summarized in Table 1. These are worth documenting because several early "wins" were revealed to be artifacts once the bugs were fixed.

Table 2: Evolution of the Monte Carlo simulation success rate.

Stage	Step 1	Step 2	Step 3	Overall
Pre-fix baseline	44.9%	3.2%	5.6%	0.0%
After alias-reference and onset fix	100%	29.2%	0.0%	0.0%
After GLRT target estimator	100%	16.7%	71.3%	12.5%
After removing direction as a hard gate	100%	58.3%	70.4%	41.2%
Final simulated configuration	100%	100%	70.4%	70.4%

Two lessons stand out. First, the alias branch is extremely sensitive to reference mismatch. The discrete-time alias template was not a cosmetic improvement; it changed a nonfunctional matched filter into a useful one. Second, the claimed bandwidth gain depends more on calibration correctness than on any single line of correlation code. Any method that lets calibration drift or phase-wrap artifacts into the combination will produce apparently narrow peaks at the wrong delay.

6 Simulation Environment

AliasProbe includes both a Python and a Swift simulation environment. The Python simulator is a single-file implementation that synthesizes the entire acquisition chain: waveform generation, pulse inversion, polynomial loudspeaker nonlinearity, fractional-delay propagation, direct and reflected paths, per-cycle jitter, and additive noise. The Swift package mirrors the same DSP stack used on the phone so that algorithm changes can be tested offline before deployment.

For each transmitted cycle, the simulator applies the PI sign, evaluates a polynomial nonlinearity of the form $a_1x + a_2x^2 + a_3x^3$, propagates the resulting signal through one or more delayed paths, and adds noise. The simulation can therefore answer two distinct questions: whether the estimator is logically capable of recovering the alias under the assumed physics, and which failures are likely due to hardware path instability rather than a broken algorithm.

A representative Monte Carlo sweep used 216 trials: 6 distances, 4 reflector amplitudes, 3 noise levels, and 3 random seeds. The evolution of the success rate over the main bug fixes is shown in Table 2. The final simulated configuration reaches a 70.4% overall pass rate, with a mean width ratio of 2.80 × among passing trials and 88.9% of passing trials exceeding 2.0 × on both held-out splits.

The simulator does *not* prove the physical iPhone result by itself. What it does prove is that once the alias template, onset logic, and coherent combination are implemented correctly, the intended signal model does support 2 × to near-3 × mainlobe narrowing under realistic SNR and jitter. The remaining gap between simulation and hardware therefore needs to be explained by hardware path variability, not by missing math.

7 Real Hardware Evaluation

7.1 Experimental setup

The real-hardware prototype is an iOS app that runs entirely on-device. The stock iPhone audio path is used at 48 kHz, with no jailbreak, no external sensors, and no modified codec driver. The

main target experiment uses a single flat reflector placed roughly 25–34 cm from the bottom edge of the phone. The expected reflection lag at 48 kHz lies in the neighborhood of 70–100 samples depending on exact geometry. Good runs were found to require a stable phone pose, a hard surface, strong pulse-inversion cancellation, and a clean same-session calibration.

The strongest operating zone discovered so far is a reflector distance of roughly 31–34 cm with the phone flat and the reflector facing the bottom speaker edge. In that regime, Step 1 becomes much more reliable and the common target lag moves to a physically plausible region near 79–81 samples instead of earlier parasitic locks around 65–71 samples.

7.2 Metrics and guardrails

A result is only claimed if it survives all of the following:

- Step 1 shows nontrivial alias SNR and positive split-half reproducibility;
- Step 2 finds a physically plausible common lag with support from both bands;
- Step 3 yields held-out narrowing on odd-fit/even-test and even-fit/odd-test splits; and
- the run does *not* show a residual-search boundary hit, impossible gain, or migration to an implausible lag.

This matters because Step 3 alone can produce sharp peaks even when the target lag is wrong. The committee-facing claim therefore depends on the full evidence chain, not on a single narrowing metric.

7.3 Representative real-phone results

Table 3 summarizes the clearest real-hardware results documented to date.

The most conservative interpretation is therefore:

- all-stage confirmed, held-out reductions above $2\times$ have been observed on stock hardware; and
- additional runs in the 2.1–2.8 $\times$ range exist but are not all accepted by the strictest current gates.

That is why the claim in this draft is written as "more than $2\times$" rather than "stable $3\times$" or even "reproducible $2.8\times$ median."

7.4 Why stable $3\times$ remains difficult

Four practical factors dominate.

Weak and drive-dependent second band. Loudspeaker large-signal behavior changes with operating amplitude, and the harmonic branch is inherently weaker than the linear branch [12, 13]. Mixed calibration/probe volume combinations produced attractive but unstable results; matched-volume operation is more honest and more transferable.

Pulse-inversion sensitivity. PI cancellation varied from roughly 4 dB to more than 30 dB across runs. Real improvements only appeared consistently when PI cancellation was high. This is consistent with the harmonic imaging literature, where incomplete PI extraction degrades second-harmonic compression [5, 6].

Table 3: Representative real-hardware results on stock iPhone. "Confirmed" means all stages passed under the current guardrails.

Setup	Step 1 evidence	Step 2 status	Held-out ratio	Status
Reflector at 332 mm	Alias SNR 11.5 dB, split-half 0.933, direction ratio 1.8 ×	Common lag in target window	2.46 ×	Confirmed
Reflector near 253 mm	All stages passed; lag near 71 samples	Common lag at edge of target window	2.08 × / 2.10 ×	Confirmed
Reflector near 253 mm	Step 3 robust but Step 1 SNR marginal	Common lag plausible	2.23 × / 2.79 ×	Suggestive, not counted
Reflector at 31–34 cm, 100% volume	Alias SNR 15.0–15.3 dB, split-half 0.63–0.80	Common lag near 79–80, but separate-band peaks differ by 5–6 samples	1.71 × to 2.51 ×	Plausible, Step 2 not yet accepted

Template drift. Even after the main alias-reference bug was fixed, cross-session template reuse was unreliable. Same-session direct-path updates produced much more stable results.

Target not perfectly point-like. The single-reflector baseline is not always a clean delta reflector at 4 kHz bandwidth. Local clusters or parasitic geometry can inflate the fundamental-only width, which is one reason why some runs showed large apparent ratios that were subsequently rejected.

8 Why the Claim is Defensible

The strongest committee question is not "can you make a narrow peak?" It is "why is the result not smoke and mirrors?" The answer is that AliasProbe uses multiple independent tests that would have to fail in a coordinated way for a false positive to survive.

Proposition 2 (The 8–16 kHz band is not arbitrary noise)**.** *The aliased band is transmission-locked, because it appears preferentially during chirp-on intervals rather than guard intervals and survives coherent averaging. It is also deterministic, because odd and even PI halves reproduce it with positive correlation. Random environmental noise would not satisfy both conditions simultaneously.*

Proposition 3 (The aliased band is not merely an unrelated spectral artifact)**.** *The aliased band prefers the physically correct 16→8 kHz sweep direction over the wrong 8→16 kHz direction, and in accepted runs it supports the same common target lag as the fundamental. An unrelated artifact would not be expected to be both direction-consistent and delay-consistent.*

Proposition 4 (The claimed narrowing is not a pure fitting artifact)**.** *AliasProbe fits the residual inter-band correction on one half of the data and measures narrowing on the other, then swaps the halves. A result only counts if both held-out directions still narrow. Therefore, the claim does not depend on tuning and scoring the same data.*

These safeguards are deliberately stricter than needed to produce pretty plots. In several runs, Step 3 showed visually narrow peaks but the run was rejected because the lag moved outside the physical window, the search saturated at its boundary, or the alias evidence failed reproducibility tests. The existence of meaningful failures is itself evidence of a real physical effect rather than an over-permissive post-processing trick.

9 Limitations and Scope of the Current Claim

This draft is intentionally conservative about what is and is not established.

What is established. On stock 48 kHz iPhone audio hardware, the built-in speaker and microphone can generate and record a usable aliased second-harmonic band from a 16–20 kHz FMCW chirp. That band is detectable, direction-consistent, and capable of producing more-than-2 $\times$ held-out single-reflector range-response narrowing relative to the fundamental-only baseline in controlled conditions.

What is not yet established. We do not yet claim a two-reflector separability experiment in which the combined method resolves targets that the 4 kHz fundamental alone cannot. That direct resolution experiment is the natural next step. We also do not yet claim stable median 3 $\times$ performance across arbitrary phone poses, volumes, and reflector geometries.

Why this matters for wording. The safest and most defensible wording today is therefore "single-reflector range-response width reduction" or "held-out matched-filter narrowing." "Resolution gain" can still be used informally, but only with the caveat that the direct two-target test is future work.

10 Future Work

Three directions are most likely to strengthen the result.

Per-volume template banks and source shaping. The harmonic pulse-compression literature suggests that matching the excitation spectrum to the transducer improves second-harmonic SNR and sidelobes [6]. AliasProbe should therefore maintain per-volume calibration banks and test transmit pre-emphasis designed to flatten the alias branch.

Hardware routing characterization. Apple documents multiple built-in microphone data sources and system stereo/mono controls on iPhone [14, 15]. That makes it plausible that some speaker/microphone geometries are substantially better than others. A structured characterization of per-speaker and per-microphone transfer functions is likely to improve repeatability.

Direct separability experiment. The most important scientific next step is a preregistered two-reflector experiment with known spacing. The test should show that the fundamental-only processor cannot separate the reflectors while the combined processor can. That would convert the current strong proxy into a direct resolution result.

11 Conclusion

AliasProbe shows that a stock iPhone's own hardware imperfections can be turned into sensing bandwidth. A 16–20 kHz FMCW chirp excites a second harmonic at 32–40 kHz through the built-in speaker's nonlinearity. Residual leakage of that harmonic through the microphone path aliases into 8–16 kHz at the stock 48 kHz ADC. With pulse inversion, same-session direct-path calibration, common-delay GLRT, GCC-β combination, and held-out validation, that aliased band becomes a real second observation of the same acoustic path.

The theoretical ceiling remains $3\times$, corresponding to coherent use of the full 8–20 kHz aperture. The current honest real-hardware claim is narrower and stronger: more-than-$2\times$ held-out single-reflector range-response narrowing has already been observed on stock iPhone hardware, and the remaining gap to stable $3\times$ is explained by identifiable physical factors rather than by missing theory. In short, the effect is real; the engineering problem that remains is making the second band stationary enough for the full ceiling to appear consistently.

A Detailed Derivation of the Discrete-Time Alias Template

A subtle but important point is that the recorded alias should not be modeled as an arbitrary analytic down-chirp. Let the transmitted discrete-time waveform be $s[n] = \cos\phi[n]$ with finite support. Then

$$s[n]^2 = \frac{1}{2} + \frac{1}{2}\cos(2\phi[n]). \tag{23}$$

The second term has twice the phase of the original chirp, but its discrete-time support, onset, and windowing are inherited from the finite-length transmitted waveform. The actual aliased harmonic therefore follows this pipeline:

1. square the transmitted chirp in discrete time;
2. band-limit the result to the harmonic or alias band of interest;
3. account for the sign reversal induced by folding through the 48 kHz Nyquist frequency; and
4. normalize the resulting discrete-time template.

Any analytic formula that skips these steps can mismatch both phase and windowing, which is precisely what happened in early versions of AliasProbe.

B Step-by-Step Estimator Summary

For completeness, the current production pipeline is:

1. generate 200 chirps, optionally with pulse inversion, at 16–20 kHz and 48 kHz sample rate;
2. detect and refine cycle onsets by local cross-correlation around the expected stride;
3. register adjacent PI pairs in the direct-path window by fractional delay and complex gain fitting;
4. form pair-sum (alias) and pair-difference (fundamental) channels;
5. average the registered pairs;

6. learn or update same-session direct-path templates in both bands;
7. Step 1: compute alias chirp-on/guard SNR and split-half reproducibility;
8. Step 2: compute the shared-lag GLRT over the target window and verify both bands support the same lag;
9. Step 3: perform per-band GCC-β, direct-path deflation, anchored inter-band correction, and held-out width measurement; and
10. reject any run that shows implausible lag migration, correction-boundary saturation, or an impossible gain.

C Reproducibility Checklist and Suggested Claim Language

For a claim-ready result, the following checklist is recommended:

- matched-volume calibration and probe;
- high pulse-inversion cancellation and low pair-registration jitter;
- same-session direct-path template update;
- target lag within the physically plausible window;
- positive held-out narrowing on both odd/even split directions; and
- no boundary hit in the residual delay search.

Suggested claim language for talks or papers:

> "To our knowledge, this is the first stock-iPhone, stock-48 kHz demonstration of more-than-2 $\times$ held-out acoustic FMCW single-reflector range-response narrowing by coherently exploiting an aliased second-harmonic band generated by the built-in speaker."

If the two-reflector separability experiment is added later, "single-reflector range-response narrowing" can be replaced by "range-resolution improvement."

References

[1] Yuchi Chen, Wei Gong, Jiangchuan Liu, and Yong Cui. I Can Hear More: Pushing the Limit of Ultrasound Sensing on Off-the-Shelf Mobile Devices. In *IEEE INFOCOM 2018 – IEEE Conference on Computer Communications*, 2018.

[2] Xiangru Chen, Dong Li, Yiran Chen, and Jie Xiong. Boosting the Sensing Granularity of Acoustic Signals by Exploiting Hardware Non-linearity. In *Proceedings of the 21st ACM Workshop on Hot Topics in Networks (HotNets '22)*, 2022. DOI: 10.1145/3563766.3564091.

[3] Shirui Cao, Dong Li, Sunghoon Ivan Lee, and Jie Xiong. PowerPhone: Unleashing the Acoustic Sensing Capability of Smartphones. In *Proceedings of the Annual International Conference on Mobile Computing and Networking (MobiCom)*, 2023. DOI: 10.1145/3570361.3613270.

[4] J. M. G. Borsboom, C. T. Chin, A. Bouakaz, M. Versluis, and N. de Jong. Harmonic Chirp Imaging Method for Ultrasound Contrast Agent. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 52(2):241–249, 2005. DOI: 10.1109/TUFFC.2005.1406550.

[5] Jaehee Song, Sangwon Kim, Hak-Yeol Sohn, Tai-Kyong Song, and Yang Mo Yoo. Coded Excitation for Ultrasound Tissue Harmonic Imaging. *Ultrasonics*, 50(6):613–619, 2010. DOI: 10.1016/j.ultras.2010.01.001.

[6] Muhammad Arif, Muhammad Asim Ali, Muhammad Mujtaba Shaikh, and Steven Freear. Investigation of Non-linear Chirp Coding for Improved Second Harmonic Pulse Compression. *Ultrasound in Medicine & Biology*, 43(8):1690–1702, 2017. DOI: 10.1016/j.ultrasmedbio.2017.03.005.

[7] C. H. Knapp and G. C. Carter. The Generalized Correlation Method for Estimation of Time Delay. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 24(4):320–327, 1976.

[8] G. C. Carter. Coherence and Time Delay Estimation. *Proceedings of the IEEE*, 75(2):236–255, 1987.

[9] José Manuel Villadangos, Jesús Ureña, Juan Jesús García-Domínguez, Ana Jiménez-Martín, Àlvaro Hernández, and M. Carmen Pérez-Rubio. Dynamic Adjustment of Weighted GCC-PHAT for Position Estimation in an Ultrasonic Local Positioning System. *Sensors*, 21(21):7051, 2021. DOI: 10.3390/s21217051.

[10] Sigrid Dimce and Falko Dressler. Survey on Coherent Multiband Splicing Techniques for Wideband Channel Characterization. *IET Research Journals / IET Communications preprint*, 2024.

[11] Jacopo Pegoraro, Jesus O. Lacruz, Michele Rossi, and Joerg Widmer. HiSAC: High-Resolution Sensing with Multiband Communication Signals. In *Proceedings of the 22nd ACM Conference on Embedded Networked Sensor Systems (SenSys '24)*, 2024. DOI: 10.1145/3666025.3699357.

[12] Wolfgang Klippel. Loudspeaker Nonlinearities – Causes, Parameters, Symptoms. Technical paper, Klippel GmbH, 2005.

[13] Lalitha K. Gudupudi, Adam Wilson, and Mike Brookes. Characterisation and Modelling of Non-linear Loudspeakers. In *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing Workshops*, 2014.

[14] Apple. Technical Q&A QA1799: AVAudioSession – Microphone Selection. 2014. `https://developer.apple.com/library/archive/qa/qa1799/_index.html`.

[15] Apple. Adjust Audio Settings on iPhone. 2025. `https://support.apple.com/guide/iphone/adjust-audio-settings-iphb80ab7516/ios`.