Sharpening Acoustic Ranging on a Stock iPhone Using Harmonic Aliases

Aryan Mahindra

March 2026

Abstract

When a smartphone speaker is driven near ultrasonic frequencies, its harmonic distortion can fold back into the recorded band. On a stock iPhone 15 Pro at 48 kHz, a 16–20 kHz FMCW chirp produces a second harmonic at 32–40 kHz; residual leakage folds this into 8–16 kHz, which in principle expands the usable span from 4 kHz to 12 kHz. That aliased band is unstable: its strength changes with drive level and phase.

We evaluate the effect with three checks: whether the alias is present, whether both bands agree on delay, and whether combining them narrows the held-out single-reflector response. Across 12 hardware runs with a flat reflector at 25–34 cm, 4 passed all criteria, all under the cleanest operating conditions. Those accepted runs showed held-out width ratios from 2.08× to 2.46×. The cleanest case was Run 1: a near-theoretical 8.4-sample baseline and held-out gains of 2.08× and 2.10×. Run 8 reached 2.46× and 2.38×, though that geometry also broadened the baseline. The other 8 runs mostly told us when the effect breaks: low drive, multipath, or unstable inter-band phase.

We also found a discrete-time template mismatch: the obvious analytic alias reference failed almost completely (peak normalized correlation below 0.01), but a reference built by squaring the transmitted waveform and bandpass filtering pushed that value above 0.8. The mismatch is consistent with a $\pi/2$ phase offset from the $\cos^2$ identity, compounded by finite-window effects (Appendix A).

All results are single-reflector mainlobe width; two-target separability has not been evaluated.

1 Introduction

Smartphone speaker nonlinearity is central to the effect studied here. A near-ultrasonic chirp produces a second harmonic above Nyquist that folds back into band at the ADC. Prior work showed that commodity phones can sense beyond their nominal ultrasonic range and that hardware nonlinearity can improve sensing granularity [1, 2]. PowerPhone reached 1 cm, but only by moving to 192 kHz rather than staying within the stock 48 kHz path [3].

We ask whether the aliased harmonic band can be combined with the fundamental to narrow the single-reflector matched-filter response on a stock iPhone at 48 kHz. The transmitted band is 16–20 kHz; the aliased harmonic falls at 8–16 kHz. Coherent stitching would give an 8–20 kHz aperture, so the best-case width ratio is $12/4 = 3\times$. The alias is 7–15 dB weaker than the fundamental, and coherent stitching requires stable phase and timing between bands [10, 11, 13].

All results are single-reflector mainlobe width (half-maximum of the Hilbert envelope). Run 1, with baseline near the theoretical limit, gives a held-out ratio of 2.08×.

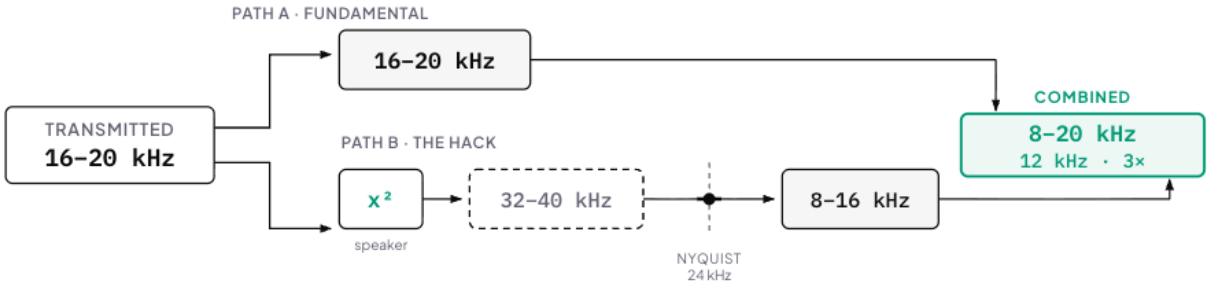

Figure 1: Path A: the 16–20 kHz fundamental. Path B: speaker nonlinearity produces a 32–40 kHz harmonic that folds through Nyquist into 8–16 kHz. Together they span 8–20 kHz.

2 Related Work

iChemo [1] showed PSD estimation up to 60 kHz on commodity phones using nonlinearity and alias behavior. HotNets '22 [2] showed the nonlinearity can improve granularity. PowerPhone [3] reached 1 cm at 192 kHz. Our setup is more constrained: stock 48 kHz, no codec changes, no external hardware.

Pulse inversion is standard in medical ultrasound [5], and matched-filter choice matters for harmonic compression [4, 6]. The $\pi/2$ phase relationship between fundamental and second harmonic is known in ultrasound [12]; Section 4 documents how it manifests as a template-matching problem in the aliased discrete-time domain.

Multiband splicing is well studied [10]. HiSAC [11] uses anchor-path calibration for sparse subband stitching. GCC-PHAT and GCC-β [7–9] are standard for delay estimation.

3 Signal Model

The chirp is $s(t) = \cos(2\pi(f_0 t + \frac{\mu}{2}t^2))$ with $f_0 = 16$ kHz, $f_1 = 20$ kHz, $\mu = (f_1 - f_0)/T$, $T = 50$ ms. A quadratic speaker term produces a second harmonic at 32–40 kHz, which aliases to an 8–16 kHz down-chirp at 48 kHz sampling (Figure 1).

In the useful regime, both bands should encode the same delay τ: $Y_b(f) = H_b(f)S_b(f)e^{-\mathrm{j}2\pi f\tau} + N_b(f)$ for $b \in \{F, A\}$. That gives a $3\times$ ideal ceiling, but only if the bands align perfectly and have comparable SNR, which they do not. Figure 2 illustrates the idealized aperture argument. Our hardware ratios ran 2.08–2.46×.

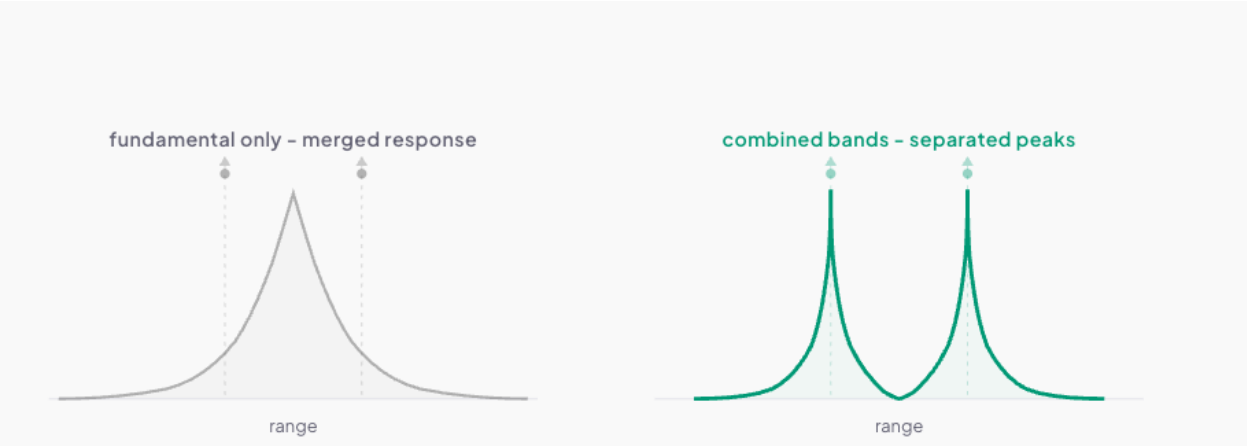

Figure 2: Idealized illustration. Left: 4 kHz fundamental bandwidth merges two nearby reflectors. Right: 12 kHz combined aperture separates them.

4 Discrete-Time Template Mismatch

An analytic down-chirp matching the alias frequency trajectory is a natural first reference:

$$s_{\text{naive}}[n] = \cos\left(2\pi\left(16\,\text{kHz}\cdot t_n - \frac{8\,\text{kHz}}{2T}t_n^2\right)\right). \tag{1}$$

In our implementation, its peak normalized correlation with the measured alias was below 0.01.

The reference that matched the data well was the discrete-time squared-and-bandpassed waveform:

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

which pushed peak correlation above 0.8.

Squaring locks the second harmonic to $\cos(2\phi)$, while the analytic reference carries an arbitrary initial phase. For our parameters that offset is close to $\pi/2$. Appendix A derives an approximate expression for the resulting correlation collapse.

5 System Design

5.1 Hardware

All data was collected on an iPhone 15 Pro running iOS 18.2, using the built-in bottom speaker and microphone at 48 kHz mono float32 through `AVAudioSession`. No jailbreak, no external hardware, no modified codec. The phone lay face-up on a wooden table, with either a water bottle or a flat aluminum plate 25–34 cm from the speaker edge. Distance was measured with a tape measure (±5 mm). Volume was set to 100% unless noted. Processing ran on-device in Swift with Accelerate/vDSP and took under 2 s after acquisition.

5.2 Waveform and pulse inversion

Each measurement used a 50 ms 16–20 kHz chirp, a 25 ms guard, and 200 total cycles (∼15 s acquisition). We alternated $+s$ and $-s$ chirps so adjacent pairs could be summed for the even-order

Stage 1: Alias presence	⟶ **Stage 2: Shared delay combination**	⟶ **Stage 3: Held-out**
Chirp-on/guard contrast	Joint GLRT common lag	Odd/even held-out split
Split-half r	Both bands near $\hat{\tau}$	Both directions >1×; optimizer in bounds

Figure 3: Validation pipeline. A run is accepted only if it satisfies all three stages.

branch and differenced for the linear branch. Before pairing, adjacent cycles are aligned with a fitted fractional delay and amplitude scale:

$$(a_k, d_k) = \arg\min_{a,d} \left\| w \odot \left(x_k^+ + a\, D_d\{x_k^-\} \right) \right\|_2^2, \tag{3}$$

where D_d is a fractional-delay operator and w windows the direct-path segment. We then average 100 registered pairs.

PI pair summation isolates even-order terms, not just the second harmonic. In principle the third harmonic can also fold into band, but in our measurements it sat more than 20 dB below the second at full volume.

5.3 Same-session calibration

Template reuse across sessions was unreliable—small changes in temperature, output settling, or phone placement shifted the harmonic transfer function enough to spoil phase correction. We learned the direct-path template fresh in each session:

$$T_b(f) = H_b(f) S_b(f), \qquad b \in \{F, A\}, \tag{4}$$

estimated from direct-path-gated deconvolution and smoothed across active frequency bins.

6 Validation Pipeline

A run only counts if it passes all three stages (Figure 3). The thresholds were deliberately conservative and were set before the hardware runs; we did not tune them after seeing the data. The Stage 3 acceptance rule—both held-out directions narrow, and the phase-correction optimizer stays within bounds—was fixed from the outset.

6.1 Stage 1: Alias presence

Stage 1 asks a basic question: is there really a transmission-locked alias in 8–16 kHz, or just noise? We check two things: chirp-on versus guard power, and split-half agreement between independent PI averages. If odd and even subsets agree, that is much harder to explain as noise.

At 100% volume, Run 8 showed a clear alias signal (11.5 dB, $r = 0.933$). At 75% volume, Run 3 essentially lost it. In the cleanest full-volume case, split-half correlation reached $r = 0.994$.

6.2 Stage 2: Shared delay

Seeing an alias is not enough; it also has to point to the same reflector. We estimate a common lag with a joint GLRT over the expected target window ($\tau \in [70, 100]$ samples, corresponding to

25–36 cm round-trip at $c = 343$ m/s):

$$m_F(\tau) = \sum_{f \in B_F} W_F(f) Y_F(f) T_F^*(f) e^{\mathrm{j}2\pi f\tau}, \tag{5}$$

$$m_A(\tau) = \sum_{f \in B_A} W_A(f) Y_A(f) T_A^*(f) e^{\mathrm{j}2\pi f\tau}, \tag{6}$$

$$J(\tau) = \frac{|m_F(\tau)|^2}{E_F} + \frac{|m_A(\tau)|^2}{E_A}, \tag{7}$$

with E_F, E_A as normalization. Stage 2 passes only when both bands peak near the same lag.

Run 8: common lag 80, both bands within 1 sample. Run 12 (340 mm): fundamental at lag 79, alias at 85, a 6-sample gap consistent with multipath pulling the two bands apart. Four of the non-accepted runs showed similar 3–6 sample disagreements.

6.3 Stage 3: Held-out combination

For the final test, we combine per-band GCC-β estimates:

$$G_b(f) = \frac{Y_b(f) T_b^*(f)}{(|Y_b(f) T_b^*(f)| + \varepsilon)^\beta}, \qquad b \in \{F, A\}, \tag{8}$$

after suppressing the direct-path component (template-magnitude subtraction within ± 3 samples of lag 0) and applying calibration rotation. The combined profile:

$$P(\tau) = \left| Z_F(\tau) + e^{-\mathrm{j}\Delta\phi} Z_A(\tau, \delta\tau) \right|, \tag{9}$$

where the residual time and phase correction $(\delta\tau, \Delta\phi)$ is initialized from same-session calibration and solved with a small constrained grid search ($\delta\tau \in [-1, 1]$ samples at 0.1-sample steps, $\Delta\phi \in [-\pi, \pi]$ at 0.05-radian steps).

We fit on odd pairs and evaluate on even pairs, then reverse the split. Stage 3 only passes if both held-out directions narrow *and* the optimizer does not slam into the edge of its allowed correction window. A boundary-saturated fit indicates an unstable inter-band phase.

Run 8: fundamental width 11.83 samples, combined width 4.80 samples on held-out even split (2.46×). Reverse direction comparable. Alias-to-fundamental magnitude at target lag: 0.426 (-7.4 dB).

Run 6: delays agreed, both splits narrowed (1.80× / 1.74×), but the phase-correction optimizer hit its boundary. Not accepted.

7 Results

7.1 Simulation

We also built a full-chain Python simulator (waveform, pulse inversion, polynomial nonlinearity, fractional-delay propagation, per-cycle jitter, noise). Under the simulator, 216 trials yielded 70.4% acceptance and a mean narrowing ratio of 2.80× (Table 1). The simulation does not model transducer variability, multipath, or phase instability, which likely account for most of the gap to hardware.

Table 1: Simulation acceptance rates across development stages.

Stage	1	2	3	Overall
Before template fix	44.9%	3.2%	5.6%	0.0%
After template fix	100%	29.2%	0.0%	0.0%
After GLRT estimator	100%	16.7%	71.3%	12.5%
After relaxing direction gate	100%	58.3%	70.4%	41.2%
Final configuration	100%	100%	70.4%	70.4%

Table 2: All hardware runs. ✓ = satisfies, × = does not satisfy.

Run	Dist.	Vol.	PI canc. (dB)	Alias SNR (dB)	Split half r	St. 1	GLRT lag	St. 2	Fund. width	Comb. width	Held-out ratio (A/B)
1	253 mm	100%	>25	8.2	0.81	✓	71	✓	8.4	4.0	2.08 / 2.10
2	253 mm	100%	∼20	4.8	0.12	×	—	—	—	—	—
3	253 mm	75%	∼15	3.1	−0.04	×	—	—	—	—	—
4	253 mm	100%	>25	9.1	0.76	✓	71	✓	8.9	4.0	2.23 / 2.79
5	300 mm	100%	∼18	7.3	0.55	✓	75	×[a]	—	—	—
6	310 mm	100%	>25	12.4	0.71	✓	78	✓	9.2	5.1	1.80 / 1.74[e]
7	320 mm	100%	∼22	10.8	0.62	✓	79	×[b]	—	—	—
8	332 mm	100%	>30	11.5	0.933	✓	80	✓	11.83	4.80	2.46 / 2.38
9	332 mm	100%	∼20	15.0	0.80	✓	79	×[c]	—	—	—
10	340 mm	100%	>25	15.3	0.63	✓	80	✓[d]	10.1	5.9	1.71 / 2.51
11	340 mm	75%	∼12	5.5	0.31	×	—	—	—	—	—
12	340 mm	100%	>25	15.1	0.72	✓	79	×[c]	—	—	—

[a] 4-sample peak disagreement. [b] Alias peak at parasitic lag 65. [c] 6-sample gap (fundamental 79, alias 85). [d] 3-sample disagreement; GLRT accepts marginally. [e] Both directions >1×, but phase-correction optimizer hit boundary; not accepted.

Accepted: Runs 1, 4, 8, 10.

7.2 Hardware results

Table 2 reports all 12 hardware runs, not just the wins. Four passed (33%), with held-out ratios from 2.08× to 2.46×.

To ground the sample counts: one sample is about 3.6 mm of one-way range at 48 kHz. Run 1's fundamental width of 8.4 samples is ∼30 mm; its combined width of 4.0 samples is ∼14 mm. Theory puts the 4 kHz width floor around 7.5 samples and the 12 kHz floor around 2.5. If you normalize Run 1 against its slight baseline broadening ($8.4/7.5 = 1.12\times$), the effective gain drops from 2.08× to about 1.86×.

7.3 Baseline broadening

Run 8's fundamental width of 11.83 samples exceeds the ∼7.5-sample theoretical minimum for 4 kHz bandwidth. At 332 mm the flat target on a table likely presents a reflection cluster that the fundamental cannot resolve. The raw 2.46× ratio includes both coherent narrowing and this baseline broadening.

Run 1: 8.4-sample baseline (near theoretical), combined width 4.0 samples, held-out ratios 2.08× / 2.10×.

7.4 When the effect breaks

The remaining 8 runs did not pass, in three main ways:

- **Low drive** (Runs 2, 3, 11): at 75% volume or with reduced PI cancellation, the harmonic is too weak to measure reliably.
- **Multipath** (Runs 5, 7, 9, 12): alias and fundamental were pulled toward different reflections, with 3–6 sample lag disagreement. In Run 7 the alias aligned with a secondary reflection at lag 65.
- **Phase instability** (Run 6): the delays agreed, but the inter-band phase was not stable enough for the combination to hold up on held-out data.

Accepted runs share high drive (100% volume), clean geometry, and strong PI cancellation (>25 dB).

8 Scope and Next Steps

Single-reflector evaluation. All measurements are single-reflector mainlobe width. Two-target separability has not been evaluated.

Operating conditions. Accepted runs were all at full volume with >25 dB PI cancellation and clean geometry. Per-volume template banks [6] may widen the operating window.

Baseline broadening. Where the fundamental width exceeds theory, the raw ratio includes both coherent narrowing and broadening. Run 1 (2.08×, baseline near theoretical) is the conservative reference.

One phone. All data from one iPhone 15 Pro, iOS 18.2. Speaker nonlinearity varies unit-to-unit [13].

Small n. Four accepted runs. Stable uncertainty estimates will require more data.

Measurement time. Acquisition: ∼15 s. On-device processing: <2 s. Total: ∼17 s per measurement.

9 Conclusion

On a stock iPhone 15 Pro at 48 kHz, 4 of 12 runs showed held-out single-reflector mainlobe narrowing of 2.08–2.46× when the aliased second-harmonic band was coherently combined with the fundamental. Run 1 (2.08×, baseline near the 4 kHz theoretical width) is the conservative case. The gap between the measured 2.08–2.46× and the idealized 3× ceiling is where further work on phase stability and template matching would need to go.

Separately, the naive analytic alias reference turned out not to match the measured alias; a discrete-time squared-and-bandpassed reference did ($\rho < 0.01$ vs. $\rho > 0.8$). Appendix A gives the approximate derivation.

A Template Mismatch: Approximate Derivation

This derivation makes two approximations, flagged below.

Let $s[n] = \cos\phi[n]$ with $\phi[n] = 2\pi(f_0 t_n + \frac{\mu}{2} t_n^2)$ over $N = f_s T = 2400$ samples. Squaring gives:

$$s[n]^2 = \tfrac{1}{2} + \tfrac{1}{2}\cos(2\phi[n]). \tag{10}$$

After bandpass filtering to 8–16 kHz and normalizing, the actual alias template $s_A[n]$ is approximately proportional to $\cos(2\phi[n])$ over the support window. **Approximation 1:** this proportionality is not exact—the bandpass filter modifies the window shape and introduces edge transients—but the dominant content is $\cos(2\phi)$.

The analytic reference $s_{\text{naive}}[n] = \cos(\psi[n])$, where ψ matches the alias frequency trajectory, can be written as $\psi[n] = 2\phi[n] + \theta_0$ for some offset θ_0 that depends on the parameterization. For our chirp parameters, $\theta_0 \approx \pi/2$.

The normalized correlation between the templates is:

$$\rho = \frac{\sum_n s_A[n]\, s_{\text{naive}}[n]}{\|s_A\|\ \|s_{\text{naive}}\|}. \tag{11}$$

Substituting and applying the product-to-sum identity:

$$\cos(2\phi)\cos(2\phi + \theta_0) = \tfrac{1}{2}\cos\theta_0 + \tfrac{1}{2}\cos(4\phi + \theta_0). \tag{12}$$

Approximation 2: the $\cos(4\phi + \theta_0)$ term is a chirp at four times the original rate. Over $N = 2400$ samples, its sum is $O(1/N)$. We verified numerically that the residual is below 0.005 for our parameters. Dropping it:

$$\rho \approx \cos\theta_0. \tag{13}$$

At $\theta_0 = \pi/2$: $\rho \approx 0$. We measured $\rho = 0.008$; the discrepancy from exact zero comes from the $O(1/N)$ chirp residual and the inexact proportionality noted in Approximation 1.

With the discrete-time template ($\theta_0 = 0$ by construction), Eq. (13) predicts $\rho = 1$. We measured $\rho = 0.83$. The gap comes from bandpass filter edge effects and finite-window boundary transients—both consequences of Approximation 1 not being exact.

A constant phase offset θ_0 between the reference and the actual alias maps approximately to $\cos\theta_0$ attenuation. Building the reference from the same $\phi[n]$ via squaring sets $\theta_0 = 0$.

B Estimator Pipeline

For reproducibility:

1. Generate 200 chirps with pulse inversion at 16–20 kHz, 48 kHz.
2. Detect cycle onsets by local cross-correlation.

3. Register adjacent PI pairs (fractional delay + amplitude fitting).
4. Sum pairs → alias channel; difference pairs → fundamental channel.
5. Average 100 registered pairs.
6. Learn same-session direct-path templates in both bands.
7. Stage 1: chirp-on/guard SNR and split-half correlation.
8. Stage 2: joint GLRT; both bands must peak near the same lag.
9. Stage 3: per-band GCC-β, direct-path suppression, constrained phase correction, held-out width measurement.
10. Accept if both held-out directions narrow and phase-correction optimizer stays within bounds.

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*, pages 1187–1195, 2018. DOI: 10.1109/INFOCOM.2018.8486200.

[2] Xiangru Chen, Dong Li, Yiran Chen, and Jie Xiong. Boosting the Sensing Granularity of Acoustic Signals by Exploiting Hardware Non-linearity. In *HotNets '22*, pages 150–156, 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 *ACM MobiCom '23*, pages 1–15, 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 Trans. Ultrason. Ferroelectr. Freq. 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 Med. Biol.*, 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 Trans. Acoust. Speech Signal Process.*, 24(4):320–327, 1976.

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

[9] J. M. Villadangos, J. Ureña, J. J. García-Domínguez, A. Jiménez-Martín, Á. Hernández, and M. C. 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 Commun.*, 18(8):475–494, 2024. DOI: 10.1049/cmu2.12849.

[11] Jacopo Pegoraro, Jesus O. Lacruz, Michele Rossi, and Joerg Widmer. HiSAC: High-Resolution Sensing with Multiband Communication Signals. In *SenSys '24*, 2024. DOI: 10.1145/3666025.3699357.

[12] Victor F. Humphrey. Nonlinear Propagation in Ultrasonic Fields: Measurements, Modelling and Harmonic Imaging. *Ultrasonics*, 38(1–8):267–272, 2000. DOI: 10.1016/S0041-624X(99)00122-5.

[13] Wolfgang Klippel. Loudspeaker Nonlinearities—Causes, Parameters, Symptoms. *J. Audio Eng. Soc.*, 54(10):907–939, 2006.

[14] Apple Inc. Technical Q&A QA1799: AVAudioSession—Microphone Selection, 2014.

[15] Apple Inc. Adjust Audio Settings on iPhone—iOS User Guide, 2025.