[Emerging Ideas] Phonotonos: Through-Skin Ultrasonic Blood Flow Sensing Using Smartphones

Shirui Cao[†], Jie Xiong[‡], Rishav Gupta[§], Sunghoon Ivan Lee[†], Jeremy Gummeson[†*], Dong Li[§*]
[†]University of Massachusetts Amherst, [‡]Nanyang Technological University, [§]University of Maryland Baltimore County
shiruicao@cs.umass.edu, jie.xiong@ntu.edu.sg, rishavg1@umbc.edu, silee@cs.umass.edu, jgummeso@umass.edu, dli@umbc.edu

Abstract

Cardiovascular diseases remain the leading cause of death worldwide, and Doppler blood flow indices play a crucial role in their early detection and diagnosis. Existing ultrasound devices, though effective, are costly, bulky, and unsuitable for daily or at-home use. In this work, we introduce *Phonotonos*, a system that transforms commodity smartphones into portable and accessible tools for through-skin ultrasonic sensing of blood flow velocity. We demonstrate that despite smartphones' inherent limitations—including low ultrasound power, coarse spatial resolution, and lack of beamforming—blood flow velocity waveforms can still be recovered using novel phase-based signal processing, including nonlinearity cancellation and baseline drift removal. We further propose a triple-modality framework that fuses Doppler ultrasound, arterial sound, and IMU for robust artery localization and interfering motion (e.g., involuntary hand motion) rejection. Extensive simulations, phantom experiments, and IRB-approved human studies validate that our proposed system can measure four key Doppler indices (AT, S/D ratio, RI, PI) with accuracy comparable to dedicated medical devices. These results highlight the potential of smartphones to democratize vascular health monitoring and enable continuous cardiovascular screening in everyday environments.

CCS Concepts

• **Human-centered computing** → *Smartphones.*

ACM Reference Format:
Shirui Cao[†], Jie Xiong[‡], Rishav Gupta[§], Sunghoon Ivan Lee[†], Jeremy Gummeson[†*], Dong Li[§*]. 2026. [Emerging Ideas] Phonotonos: Through-Skin Ultrasonic Blood Flow Sensing Using Smartphones. In *The 24th Annual International Conference on Mobile Systems, Applications and Services (MobiSys '26), June 21–25, 2026, Cambridge, United Kingdom.* ACM, New York, NY, USA, 17 pages. https://doi.org/10.1145/3745756.3809246

1 Introduction

Cardiovascular diseases remain the leading cause of death worldwide, responsible for nearly 18 million deaths each year, according to the World Health Organization (WHO) [79]. Among the cardiovascular diseases, blood flow-related disorders are the most prevalent and critical contributors to global morbidity and mortality [14]. Key Doppler indices of blood flow, including *Acceleration Time (AT)*, *Systolic-to-Diastolic Ratio (S/D)*, *Resistance Index (RI)*, and *Pulsatility Index (PI)*, are widely used to detect vascular abnormalities such as carotid stenosis, cerebrovascular insufficiency, and peripheral vascular disease [76, 104]. Regular monitoring and early detection of these abnormalities are essential for reducing disease burden and preventing severe complications including heart attacks, strokes, and organ damage.

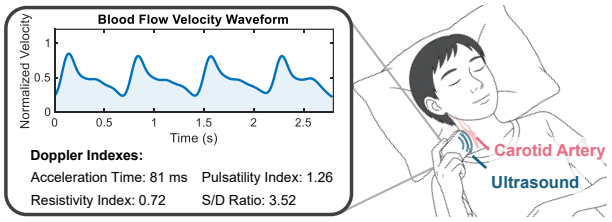


Figure 1: Phonotonos enables at-home blood flow monitoring using smartphones by extracting velocity waveforms and computing Doppler indices.

Medical ultrasound devices are commonly used for blood flow Doppler indices [6], but their high cost, bulky size, and dependence on clinical infrastructure make them impractical for routine or home-based monitoring. Handheld Doppler devices [27] offer a more accessible alternative, yet they remain relatively expensive and are not designed for daily, on-the-go use, limiting their practicality for continuous tracking.

In this paper, we propose Phonotonos[1] to explore the feasibility of transforming smartphones into portable, affordable, and user-friendly devices for estimating Doppler indices from blood flow velocity waveforms, as illustrated in Fig. 1. This approach enables seamless integration of cardiovascular monitoring into daily life for early detection and management, while the widespread use of smartphones [96] offers the potential to democratize healthcare and reach under-served populations.

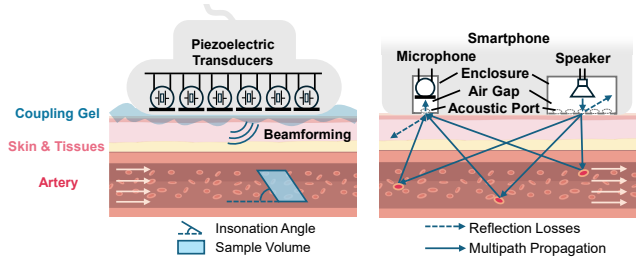
To realize this vision, we need to first understand how medical ultrasound devices compute blood flow Doppler indices. As illustrated in Fig. 2(a), these devices measure blood flow by transmitting ultrasound that propagates through human tissue, which scatters off red blood cells and is then received by the same transducer to extract velocity-dependent Doppler shifts [94]. They achieve reliable measurements with high transmission power (up to 2.3 MPa, 221 dB) [37] and coupling gel to overcome air-skin impedance mismatch, ensuring sufficient energy despite strong in-body attenuation [111, 112]. To suppress multipath, they employ high-frequency ultrasound (several MHz) for fine spatial resolution [73], Pulsed-Wave Doppler with short pulses (~ 5 μs) and time-gating to isolate

*Co-corresponding authors.

MobiSys '26, June 21–25, 2026, Cambridge, United Kingdom

ACM ISBN 979-8-4007-2027-7/2026/06
https://doi.org/10.1145/3745756.3809246

[1]Derived from the Ancient Greek roots for "sound" (phōnē) and "strain" (tonos), emphasizing its focus on cardiovascular sensing.



(a) Medical devices use coupling gel to reduce reflection losses, and transducer arrays with beamforming to confine the sample volume and suppress multipath.

(b) Smartphones have air gaps in their enclosures that cause reflection losses, and introduce multipath propagation.

Figure 2: Comparison of medical Doppler devices and smartphones for blood flow velocity sensing.

echoes from specific depths [66], and beamforming arrays to optimize the insonation angle and confine the sample volume [7, 66]. The resulting blood flow velocity waveforms are then analyzed to compute Doppler indices. Note that index computation relies only on the relative waveform shape, so a scaled (normalized) velocity waveform is sufficient.

However, when using smartphones to achieve the same function, we encounter three main limitations: (i) **Power and Coupling.** Smartphones transmit ultrasound at much lower power (typically < 120 dB) [26]. Due to the risk of leakage into speakers and microphones, coupling gels cannot be applied, resulting in a round-trip reflection loss of 60 dB [33, 48, 88]. (ii) **Resolution and Multipath.** Even with the latest techniques to improve the sampling rate [19], smartphones can only emit ultrasound below 96 kHz, resulting in coarse spatial resolution given the sound speed of 1540 m/s in soft tissue [92]. Smartphones also lack multipath-suppression mechanisms. For instance, while short pulses can separate multipath in the time domain, smartphone audio amplifiers cannot reliably produce them due to unstable transient responses [89]. Furthermore, as illustrated in Fig. 2(b), the absence of transducer arrays makes it impossible to separate signal paths from different directions. (iii) **Insonation Angle.** Accurate Doppler estimation requires aligning the ultrasound beam with the direction of blood flow, as errors increase with larger insonation angles [94]. Without imaging or beam-steering capabilities, smartphones cannot control the insonation angle.

In this paper, we demonstrate for the first time that all the aforementioned limitations of smartphones can be overcome, enabling the recovery of blood flow velocity waveforms even under conditions of low acoustic power and multipath. Our conclusions are supported through mathematical modeling, large-scale Monte Carlo simulations, phantom experiments, and experiments involving human participants. We model through-skin multipath using Rayleigh scattering [92] and show that the smartphone's asymmetric transmit–receive geometry suppresses positive Doppler components while retaining part of the negative ones, effectively creating a favorable insonation angle similar to beam steering in medical ultrasound. A key finding is that placing the smartphone in firm contact with the skin forms a sealed air cavity that mechanically couples the speaker to tissue and greatly increases acoustic penetration [75, 114], while medical ultrasound transducers must rely on gels for effective coupling. A key enabler of our approach is the use of the single-bin Discrete Fourier Transform (DFT), which allows super-resolution frequency estimation: by tracking phase variations in the transmitted frequency bin, we can resolve Doppler shifts with a precision far exceeding the limits of conventional time–frequency analysis. Building on this, we derive that blood flow velocities can be linearly approximated from the single-bin DFT phase and validate this through Monte Carlo simulations involving billions of red blood cells. Finally, we experimentally confirm these findings with a silicone phantom that mimics tissue and an alumina suspension with particle sizes comparable to red blood cells (6 μm). By varying flow velocity and tissue thickness, we evaluate the system's ability to detect and measure blood flow.

The phase variations induced by blood flow are extremely subtle, often less than 0.1 rad. Capturing such small variations on smartphones is challenging due to hardware imperfections. We observe that microphone nonlinearity introduces intermodulation between low-frequency audible sounds and the ultrasonic sensing signal, resulting in phase noise. In addition, sampling clock inaccuracies and thermal effects in smartphone speakers cause phase drift over time. Unlike prior acoustic sensing systems where large phase variations masked these issues [51], blood flow estimation is far more sensitive due to the much higher sound speed in tissue (1540 m/s vs. 340 m/s in air).

To address the above issues, we first design an adaptive Least Mean Squares (LMS) cancellation method, inspired by active noise cancellation in modern earphones [56], to suppress nonlinear distortion. To address phase drift, we exploit the sparsity of blood flow velocity waveforms in both the signal and its first three derivatives. Leveraging this property, we adapt the Baseline Estimation and Denoising with Sparsity (BEADS) algorithm [74] to effectively separate the true waveform from slow-varying phase drift.

Arteries vary in size, depth, and structure across individuals and body locations. Conventional ultrasound systems rely on trained clinicians and imaging guidance (e.g., B-mode) for accurate probe placement [5, 7, 66]. In contrast, smartphone-based sensing lacks both expert oversight and imaging support, making it difficult for untrained users to obtain accurate and reliable measurements. Ensuring reliable measurements therefore requires automated artery localization and interfering motion rejection, since arteries cannot be visualized and interfering motions can corrupt signals.

To enable accurate artery localization, we introduce a triple-modality approach that integrates Doppler ultrasound, arterial sound, and motion sensed via IMU. The key idea is that when the smartphone is correctly positioned over an artery, the heart rate estimated from blood velocities aligns with that derived from arterial pulse sounds captured by the microphone. By comparing these signals, the system delivers real-time guidance through audio and haptic feedback, where stronger vibrations or clearer pulse sounds indicate optimal placement. During measurement, IMU-based motion detection further rejects velocity segments corrupted by interfering motions.

We implemented our prototype on four smartphones, including Motorola Edge 2020, Xiaomi Redmi Note 9 Pro, Samsung Galaxy

S9+, and Samsung Galaxy Fold7. We conducted extensive experiments across diverse conditions to evaluate its robustness and clinical utility. A demonstration video and an open-source end-to-end algorithm demo can be found at Appendix C. To summarize, this paper makes the following contributions:

- We take a first step toward enabling through-skin ultrasound sensing on smartphones, paving the way to transform smartphones into affordable, and portable cardiovascular monitoring tools. We demonstrate blood flow velocity measurements across major arteries, including the carotid, radial (wrist), temporal, and posterior tibial (ankle) sites.
- We investigate why smartphone-emitted ultrasound can penetrate skin and capture scattered Doppler information from red blood cells, validating this through theoretical modeling, Monte Carlo simulations, and phantom experiments. To address the inherent challenges of subtle Doppler extraction, we design a robust DSP pipeline that incorporates a single-bin DFT phase-based algorithm, an LMS-based nonlinearity cancellation method, and a baseline estimation and denoising algorithm for correcting phase drift.
- We introduce a triple-modality framework that combines ultrasound blood flow velocity, arterial pulse sounds, and IMU data. This fusion enables robust artery localization and real-time user guidance through audio and haptic feedback, detects interfering motion, and eliminates the need for expert operation.
- Through IRB-approved experiments, we comprehensively evaluate our system across 20 subjects, multiple arterial locations, and diverse environments. Extensive results show that our system achieves a performance comparable to a dedicated medical Doppler ultrasound device [27] in measuring four key Doppler indices. Among those, we achieved 87.48% accuracy in Acceleration Time, 80.39% accuracy in S/D ratio, 94.92% accuracy in Resistance Index, and 81.19% accuracy in Pulsatility Index.

2 Modeling and Simulation for Smartphone Blood Flow Velocity Sensing

In this section, we leverage the unique opportunities of smartphones to enable through-skin Doppler sensing of blood flow velocity. We begin with a porcine tissue experiment using two smartphones to validate through-skin acoustic penetration and quantify channel attenuation. All subsequent analysis, however, is based on a single-smartphone configuration: we build a mathematical model of continuous-wave (CW) scattering from Red Blood Cells (RBCs) based on Rayleigh scattering theory, and demonstrate that Doppler information is preserved in the received microphone signals due to the asymmetric TX–RX geometry. Building on this insight, we introduce a phase-based method to extract subtle Doppler drifts and validate our approach through Monte Carlo simulations with billions simulated RBCs.

2.1 Modeling Transmission Channel

We analyze propagation of a 21 kHz tone aimed at an artery 3 cm beneath skin and validate key terms of the link budget using

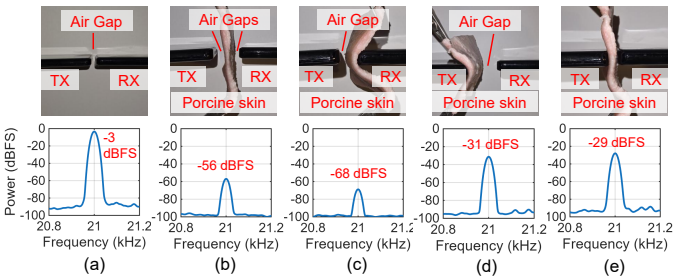


Figure 3: Porcine phantom penetration experiment configurations and measured attenuation: (a) baseline with no skin; (b) air gaps at both TX and RX; (c) TX with air gap, RX in contact; (d) TX in contact, RX with air gap; (e) both TX and RX in contact.

porcine-tissue experiments. Below, we characterize the attenuation encountered along the acoustic pathway.

Speaker Output and Air–Skin Reflection. Smartphone speakers generate about 120 dB SPL at 1 kHz [39] with a 10 dB reduction at 21 kHz [19], giving an estimated ultrasonic output of 110 dB. At the air–skin boundary, the intensity reflection coefficient is

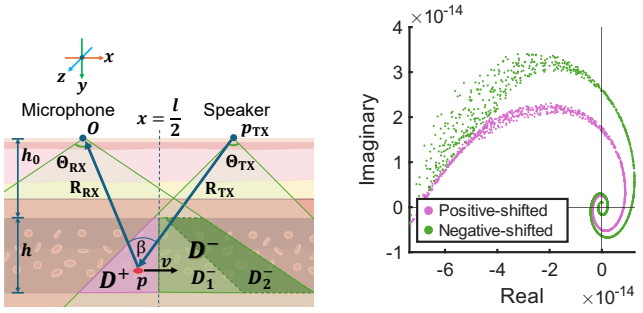
$$\Gamma_{\text{air-skin}} = \left(\frac{Z_{\text{skin}} - Z_{\text{air}}}{Z_{\text{air}} + Z_{\text{skin}}} \right)^2, \tag{1}$$

with $Z_{\text{air}} = 428$ Rayl and $Z_{\text{skin}} = 1.68$ MRayl. This mismatch reflects 99.9% incident energy, corresponding to 30 dB one-way loss (60 dB round trip).

We validated this effect using porcine skin as shown in Fig. 3. With TX smartphone and RX smartphone facing each other across a small air gap, we record a baseline with minimal loss (Fig. 3(a)). Inserting porcine skin between them results in 53 dB attenuation (Fig. 3(b)), slightly lower than the expected 60 dB. This discrepancy is due to the additional environmental reflections. To suppress these environmental contributions, we place the RX microphone directly against the skin surface, forcing the acoustic path to pass through it. Under this controlled geometry, the measured attenuation increases to 65 dB (Fig. 3(c)), consistent with the theoretical round-trip loss.

Mechanical Coupling Transmission. Interestingly, we find that when the smartphone is placed in firm contact with the skin, the speaker enclosure forms a sealed air cavity that facilitates mechanical coupling between the device chassis and the tissue [75, 114]. As shown in Fig. 3(d), contacting only the TX side dramatically reduces total attenuation to 28 dB, despite the RX still experiencing air-to-skin reflection losses. When both TX and RX are placed in close contact with the tissue (Fig. 3(e)), attenuation is further reduced to 26 dB—approaching the one-way 30 dB prediction. This behavior arises because mechanical coupling greatly reduces losses at the TX side; however, reflection at the RX interface cannot be fully eliminated, as air remains within the microphone housing before sound reaches the microphone element.

Tissue absorption and Link Budget. Tissue absorption at 21 kHz is < 1 dB over the 6 cm round trip and is negligible compared with interface losses. Combining the 110 dB source level, 10 dB rolloff, and 26 dB mechanically coupled attenuation yields a received SPL of 74 dB. With smartphone MEMS microphones providing 68 dB SNR (26 dB SPL noise floor) [23, 38, 61], the system retains a 48 dB



(a) Negative and positive Doppler shifts around $x = \frac{l}{2}$.

(b) Negatively and positively shifted signals on complex plane.

Figure 4: Asymmetry of Θ_{RX} and Θ_{TX} creates a summed signal that still preserve Doppler information.

margin for detecting Doppler-scattered components. This budget demonstrates that, despite limited transmit power and significant interface reflection, we still have sufficient link budget for through-skin blood flow velocity sensing.

2.2 Modeling Multipath Propagation

We mathematically model through-skin ultrasound propagation, starting from a single red blood cell and extending to billions of red blood cells. We show that the asymmetric speaker radiation and microphone reception patterns create an effect analogous to beamforming in medical ultrasound, effectively providing a favorable insonation angle.

2.2.1 Scattering from a Single Red Blood Cell (RBC). As shown in Fig. 4(a), the smartphone speaker (TX) emits ultrasound that propagates through tissue, scatters off RBCs, and is then received by the microphone (RX). Taking the microphone as the origin of a Cartesian coordinate system, the speaker is located at $\mathbf{p}_{TX} = (l, 0, 0)$. We now model how the ultrasound changes along this propagation path.

Transmitted Signal. We use a single-frequency continuous wave as the transmitted signal, since smartphones cannot generate precise short pulses like medical devices. With separate speakers and microphones, smartphones naturally support simultaneous transmission and reception. The transmitted signal can be denoted as $s_{TX}(t) = e^{j(\omega t+\theta_0)}$, where $\omega = 2\pi f_0$ is the angular frequency, f_0 is the transmitted frequency, and θ_0 the initial phase.

TX-to-RBC Propagation. Given an RBC at $\mathbf{p} = (x, y, z)$, the ultrasound emitted by the speaker is modeled as a spherical wave using Green's function [77]. Incorporating tissue absorption, the transfer function from the speaker at $\mathbf{p}_{TX}$ to the RBC at $\mathbf{p}$ is [45]:

$$H_{TX}(\mathbf{p}) = \frac{e^{-\alpha R_{TX}}}{4\pi c R_{TX}} e^{-jkR_{TX}}, \tag{2}$$

where α is the absorption coefficient in m^{-1}, c is the sound speed in tissue, $k = \omega/c$ is the wave number, and $R_{TX} = \|\mathbf{p} - \mathbf{p}_{TX}\|$ is the distance from the speaker to the RBC.

RBC Scattering. Since RBCs are much smaller than the ultrasound wavelength, ultrasound interacts with them mainly via Rayleigh scattering rather than reflection [92]. The corresponding scattering transfer function is given by [28, 30, 94]:

$$\begin{aligned} H_{RBC}(\mathbf{p}) &= \frac{k^2 V_{RBC}}{4\pi} (\kappa - \rho\cos(\beta))\, e^{-j\theta_{RBC}}, \\ \kappa &= \frac{\kappa_0 - \kappa_s}{\kappa_0}, \quad \rho = \frac{\rho_s - \rho_0}{\rho_s}, \end{aligned} \tag{3}$$

where V_{RBC} is the volume of a single RBC, κ_s and κ_0 are the adiabatic compressibilities of the scatterer and medium, ρ_s and ρ_0 are their densities, θ_{RBC} is a constant scattering phase shift [93]. β is the scattering angle, which can be defined as:

$$\beta = \arccos\left(\frac{(\mathbf{p} - \mathbf{p}_{TX}) \cdot \mathbf{p}}{R_{TX} R_{RX}}\right), \tag{4}$$

where $R_{RX} = \|\mathbf{p}\|$ is the distance from RBC to microphone.

RBC-to-RX Propagation. When an RBC scatters ultrasound, it can be modeled as a point source emitting a spherical wave. Using Green's function [77], the transfer function from the RBC to the microphone is given by:

$$H_{RX}(\mathbf{p}) = \frac{e^{-\alpha R_{RX}}}{4\pi c R_{RX}} e^{-jkR_{RX}}. \tag{5}$$

Doppler Shift Induced by Path Change. The RBC motion introduces a Doppler shift $\omega_D(\mathbf{p})$ that manifests as a time-varying phase term in the received signal:

$$H_D(\mathbf{p}, t) = e^{-j\omega_D(\mathbf{p})t}. \tag{6}$$

With separated and stationary transceiver, the Doppler frequency $\omega_D(\mathbf{p})$ can be derived from the bistatic Doppler model as the time derivative of the TX–RBC–RX path length [115]:

$$\omega_D(\mathbf{p}) = -\frac{1}{2\pi\lambda}\frac{d}{dt}\left(R_{TX} + R_{RX}\right), \tag{7}$$

where λ is the ultrasound wavelength. For an RBC moving with velocity $\mathbf{v} = \frac{d}{dt}\mathbf{p} = (v_{flow}, 0, 0)$, this reduces to:

$$\omega_D(\mathbf{p}) = -\frac{v_{flow}}{2\pi\lambda}\left(\frac{x}{R_{RX}} + \frac{x-l}{R_{TX}}\right), \tag{8}$$

where v_{flow} is the blood flow velocity along the x-axis.

Complete Received Signal from a Single RBC. The signal captured by the microphone can be expressed as:

$$\begin{aligned} s_{RX}(\mathbf{p}, t) &= s_{TX}(t)\, H_{TX}(\mathbf{p})\, H_{RBC}(\mathbf{p})\, H_{RX}(\mathbf{p})\, H_D(\mathbf{p}, t) \\ &= \frac{e^{-\alpha(R_{RX}+R_{TX})} k^2 V_{RBC}(\kappa - \rho\cos(\beta))}{64\pi^3 c^2 R_{RX} R_{TX}} \\ &\quad \cdot e^{j(\omega t + \theta_0 + \theta_{RBC} - k(R_{TX}+R_{RX}) + \omega_D(\mathbf{p})t)} \\ &= A_{RX}(\mathbf{p}) \cdot e^{j\theta_{RX}(\mathbf{p},t)}, \end{aligned} \tag{9}$$

where $\theta_{RX}(\mathbf{p}, t)$ is the instantaneous phase of the received signal. It can be rewritten as:

$$\theta_{RX}(\mathbf{p}, t) = \underbrace{(\omega t + \theta_0 + \theta_{RBC})}_{\text{Static Phase}} + \underbrace{k(R_{TX} + R_{RX})}_{\text{Path Phase}} + \underbrace{\omega_D t}_{\text{Doppler}}. \tag{10}$$

$A_{RX}(\mathbf{p})$ is the position-dependent amplitude of the received signal. It can be rewritten as:

$$A_{RX}(\mathbf{p}) = \underbrace{\frac{e^{-\alpha(R_{RX}+R_{TX})}}{R_{TX} R_{RX}}}_{\text{Path Attenuation}} \cdot \underbrace{(\kappa - \rho\cos(\beta))}_{\text{Angular Attenuation}} \cdot C, \tag{11}$$

where C contains constants from Equation 9. Note that for blood plasma and RBC, $\kappa > \rho$, so the impact of β will only change the amplitude, not the sign.

2.2.2 Scattering from Billions of RBCs. The received signal is modeled as the integral of scattered contributions from billions of RBCs within the overlap region D of the TX radiation and RX reception zones, as shown in Fig. 4(a). From Equation (8), we can obtain that the Doppler frequency is zero for RBCs at $x = \frac{l}{2}$, positive for RBCs in region D^+ where $x < \frac{l}{2}$ (denoted ω_D^+), and negative for RBCs in region D^- where $x > \frac{l}{2}$ (denoted ω_D^-). Accordingly, the received signal integral can be separated into two parts:

$$\begin{aligned} s_{\mathrm{RX}}(t) &= \iiint_D s_{\mathrm{RX}}(\mathbf{p}, t)\, dS \\ &= \iiint_{D^+} s_{\mathrm{RX}}(\mathbf{p}, t)\, dS \;+\; \iiint_{D^-} s_{\mathrm{RX}}(\mathbf{p}, t)\, dS, \end{aligned} \tag{12}$$

where dS is the differential volume element. When $\Theta_{\mathrm{RX}} = \Theta_{\mathrm{TX}}$, regions D^+ and D^- are symmetric about $x = \frac{l}{2}$ in terms of path attenuation, angular attenuation, and Doppler shift. Their contributions thus cancel in the complex plane, leaving only the static phase θ_S and amplitude A_S.

In smartphones, however, $\Theta_{\mathrm{RX}} > \Theta_{\mathrm{TX}}$ typically holds, creating an additional region D_2^- that preserves Doppler information after cancellation. This asymmetry arises because smartphone speakers, embedded in the chassis, radiate sound narrowly forward with steep off-axis roll-off at higher frequencies [8, 78], while MEMS microphones are nearly omnidirectional [9, 22, 90]. The mismatch between the narrow Θ_{TX} and broad Θ_{RX} creates a directional bias that separates positive and negative Doppler components, effectively mimicking directional sampling in medical ultrasound.

To visualize this effect, we treat the TX and RX as the foci of an ellipsoid and plot the complex-valued scattered contribution of each RBC inside the ellipsoid (Fig. 4(b)). Using Eq. 9 with the complex origin shifted by the static phase in Eq. 10, the positive- and negative-shifted components form concentric spirals. Near the origin, each positively shifted component has a corresponding negatively shifted counterpart, representing cancellation between D^+ and D_1^-. However, the asymmetric causes the remaining negative-Doppler contributions to dominate in both amplitude and number, so the summed signal still retains negative Doppler components. Thus, the received signal becomes:

$$s_{\mathrm{RX}}(t) = A_S \cdot e^{j(\omega t + \theta_S)} + \iiint_{D_2^-} s_{\mathrm{RX}}(\mathbf{p}, t)\, dS. \tag{13}$$

Direct evaluation of the integral over D_2^- is intractable, as it involves Fresnel integrals [13]. However, because the effective scattering region is restricted to a narrow directional sector of the artery, the resulting Doppler signals belong to a class that has been extensively analyzed in medical ultrasonography. By approximating D_2^- as a parallelogram, the received signal can be expressed as [16, 72, 86]:

$$\begin{aligned} &s_{\mathrm{RX}}(t) = A_S e^{j(\omega t + \theta_S)} + A_D e^{j(\omega t + \theta_0 + \omega_D' t)}, \\ &\omega_D' \propto \frac{v_{\mathrm{flow}} \sin(\Theta_{\mathrm{RX}} - \Theta_{\mathrm{TX}})}{c}\, \omega, \\ &A_D \propto \mathrm{Area}(D_2^-), \end{aligned} \tag{14}$$

where ω_D' is the effective Doppler shift after integration.

2.2.3 Incorporating Other Multipath. In practice, the received signal also includes a strong direct path between the speaker and microphone, as well as additional multipath components such as reverberation within the smartphone enclosure. Because these paths can be modeled as linear time-invariant transfers, they manifest as a static phase offset in the resulting signal. We can therefore rewrite the expression in the following form (see Appendix A), representing the final signal observed at the microphone:

$$s_{\mathrm{RX}}(t) = A e^{j(\omega t + \theta_S + \Delta\omega t)}, \tag{15}$$

where A is the effective amplitude after superposition, θ_S absorbs constant phase terms from both direct and scattered paths, and $\Delta\omega$ represents the resulting Doppler component. Note that scattered signals from RBCs are much weaker than static paths such as the direct path, i.e., $A_S \gg A_D$. This leads to $\Delta\omega \propto \omega_D'$, and from Equation 14 we obtain

$$\Delta\omega \propto v_{\mathrm{flow}}. \tag{16}$$

Thus, the aggregated multipath signal received at the microphone preserves a Doppler frequency shift $\Delta f = \frac{\Delta\omega}{2\pi}$ proportional to blood flow velocity.

It should be noted that this analysis models blood flow as the superposition of scattered signals from individual RBCs [86]. In practice, blood scattering is stochastic, as signals scattered by one cell may undergo further scattering by others. Nevertheless, prior studies [3, 93] have shown that the individual-scatterer model provides results equivalent to full stochastic analyses.

2.3 Phase-based Method To Extract Subtle Frequency Shift

We have shown that smartphones can capture Doppler information from RBCs, but effective extraction is challenging. Arterial blood flow ranges from 10–200 cm/s, and resolving 10 cm/s—a clinically meaningful threshold [44]—requires frequency resolution of about 0.5 Hz. Conventional time–frequency analysis needs windows longer than a cardiac systole, making it unsuitable.

We address this with a single-bin DFT centered on the transmit frequency. By exploiting the fact that Doppler shifts are extremely small, we avoid full-spectrum search and extract phase directly from the target bin. This improves sensitivity, lowers complexity to $O(N)$, supporting real-time processing. Let the DFT input signal be the signal received in Equation (15), with a length-N discrete form:

$$x[n] = A e^{j\left(2\pi \frac{f}{F_s} n + \theta_S\right)}, \quad f = f_0 + \Delta f, \quad 0 \le n < N, \tag{17}$$

where f is the input frequency and F_s is the sampling rate. The phase of the i-th bin DFT when the input frequency is slightly off from the bin-centered frequency is (Appendix B):

$$\phi = \mathrm{Arg}(X[i]) = \pi d \frac{N-1}{N} + \theta_S, \quad d = \frac{fN}{F_s} - i \tag{18}$$

When the transmitted signal is aligned with a DFT bin center such that $f_0 = k f_{\mathrm{bin}}$, the Doppler shift can be directly calculated using Equations (18):

$$\Delta f = (\phi - \theta_S) \cdot \frac{F_s}{\pi(N-1)}. \tag{19}$$

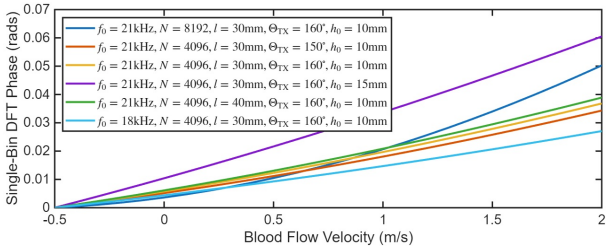


Figure 5: Simulation shows near-linear relation between ϕ and v_{flow}.

This expression shows that the Doppler shift is linearly positively correlated to the phase angle of the single-bin DFT.

Although such phase-based method improves frequency resolution, it cannot fully separate Doppler shifts from phase changes caused by variations in propagation path length—a limitation rooted in the Heisenberg uncertainty principle [32]. Arterial wall motion introduces two phase components: a Doppler shift from wall velocity and a geometric phase shift from diameter changes. The Doppler term is negligible because wall velocities are small (e.g., 3 mm/s in the CCA [100]). The displacement term, however, is proportional to flow velocity: the CCA expands by 0.5–1.5 mm per cardiac cycle [4, 98], and prior work shows

$$\Delta h = \frac{8\, v_{\text{flow}}\, \mu}{\tau_0}, \tag{20}$$

where μ is blood viscosity and τ_0 is the shear-stress constant [42]. Because path-length changes translate linearly to phase shifts, and the smartphone direction in Fig. 4(a) preserves this mapping, the measured phase remains proportional to flow velocity. Thus, combining Equations. (14) and (19), we obtain:

$$v_{\text{flow}} = a(\phi - \theta_S), \tag{21}$$

where a and θ_S are constants. In other word, the blood flow velocity can be linearly approximated using the single-bin DFT phase. Note that frequency resolution depends on both the DFT size N and the phase noise of the system.

2.4 Model Verification by Simulation

In the previous subsection, we derived a linear approximation model for blood velocity. In this subsection, we will verify our approximation model using an end-to-end Monte Carlo simulation. Specifically, the received signal is calculated using the following integral:

$$s_{\text{RX}}(t) = \iiint_D s_{\text{RX}}(\mathbf{p}, t)\, dS \approx \sum_{i=0}^{M-1} s_{\text{RX}}(\mathbf{p}[i], t), \tag{22}$$

The integration domain D corresponds to the intersection of a cylindrical artery, the speaker radiation cone, and the microphone reception cone. RBCs are assumed to be homogeneously distributed within the artery and M represents their total number. The final simulated received signal is $s_{\text{RX}}(t)$ plus reflected signal from upper artery wall.

For constants involved in the simulation, we will use values from the previous literature: $c = 1540$ m/s, $\alpha_{\text{absorb}} = -1$ dB/cm/MHz [111], $V_{\text{RBC}} = 8.7{\times}10^{-11}\text{cm}^3$, $\rho_0 = 1029\,\text{kg/m}^3$, $\kappa_0 = 4.09{\times}10^{-10}\,\text{m}^2/\text{N}$ [120], $\rho_s = 1108\,\text{kg/m}^3$, and $\kappa_s = 3.48 \times 10^{-10}\,\text{m}^2/\text{N}$ [91]. The artery is modeled with diameter $(h + \Delta h)$ and depth $(h_0 - \Delta h/2)$ beneath the skin, where Δh is modeled as $\Delta h = 0.0005 \times v_{\text{flow}}$[97]. The reflection coefficient of artery wall is derived using impedance-mismatch reflection model as in Section 2.1 and acoustic impedance of $Z_{\text{artery}} = 1.70$ MRayl, $Z_{\text{fat}} = 1.34$ MRayl [7]. A total of $M = 10$ billion scattered signals are summed in the time domain using the definitions in Equation 8 and Equation 9. An N-point single-bin DFT, as defined in Appendix B, is then applied to the aggregated signal. Fig. 5 shows simulation results with fixed parameters $F_S = 48$ kHz, $\theta_{\text{RBC}} = \frac{\pi}{2}$, $\theta_0 = \frac{\pi}{4}$, $\Theta_{\text{RX}} = 170°$, $h = 5$ mm, while varying $f_0, N, l, \Theta_{\text{TX}}$ and h_0. The blood flow velocity v_{flow} is simulated over the range -0.5 m/s to 2.0 m/s, where the negative velocity represents possible reverse blood flow in human arteries [43]. For clarity, the y-axis is shifted so that the phase ϕ at $v_{\text{flow}} = -0.5$ m/s is aligned to zero. The simulation results show that ϕ maintains a nearly linear relationship with v_{flow} under various conditions, confirming the conclusion in Equation 21.

3 System Design

Our system integrates three tightly coupled components, shown in Fig. 6. First, smartphone-specific signal processing mitigates distortions introduced by microphones and speakers, enabling extraction of a clean blood velocity waveform. Second, a real-time feedback loop assists users in positioning the phone and flags poor-quality signals, after which a 15-second measurement is recorded. Finally, validated signals are used to compute Doppler indices. This design bridges commodity hardware limitations with user-guided software, providing reliable measurements under practical conditions.

3.1 Measuring Blood Velocity Using Commodity Smartphone Hardware

Commodity smartphones suffer from hardware limitations that degrade signal fidelity. To address this, we design signal-processing methods tailored to these devices that extract a reliable blood velocity waveform. The following subsections describe techniques to mitigate microphone nonlinearity and speaker phase noise, enabling robust velocity estimation on low-cost hardware.

3.1.1 Microphone Nonlinearity Cancellation. Smartphone microphones exhibit nonlinear responses that introduce phase noise and intermodulation, degrading velocity estimation. In preliminary experiments, we observed broadened spectral lobes and interference from ambient noise in the ultrasonic band, indicating nonlinear behavior. For example, on the Motorola Edge (2020), we found that artificially induced band-limited Gaussian noise between 20–200 Hz was modulated into the carrier frequency $f_0 = 21$ kHz, producing spurious components in the measurement (Fig. 7(a)).

This kind of nonlinearity is common for MEMS microphones used on smartphones across all frequency bands [84]. The effect can be modeled as a quadratic nonlinearity [123]:

$$s'_{\text{RX}}(t) = A s_{\text{RX}}(t) + B(s_{\text{RX}}(t))^2, \tag{23}$$

where A and B represent linear and quadratic contributions. The quadratic term produces intermodulation (IM) between ultrasonic and audible signals, injecting unwanted phase components into the bin at f_0.

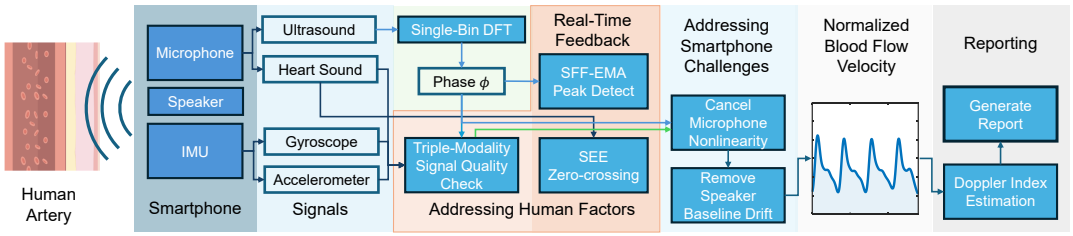
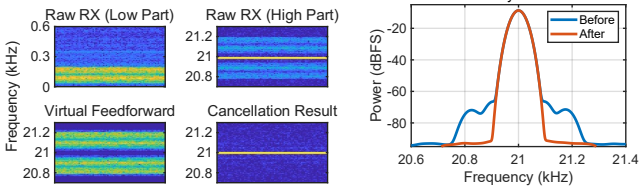


Figure 6: System Overview.



(a) Spectrograms showing the intermodulation between audible sound and ultrasound, and can be canceled using our virtual feedforward microphone.

(b) Our nonlinearity cancellation can cancel the intermodulation by −20 dB, effectively reducing phase noise.

Figure 7: Nonlinearity cancellation.

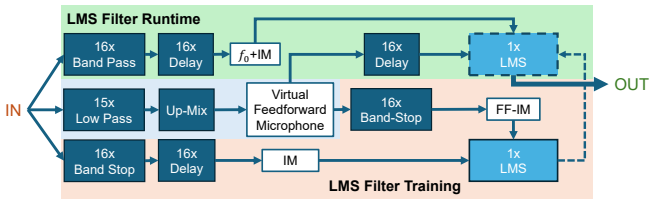


Figure 8: DSP pipeline of nonlinearity cancellation

To address this challenge, we developed an adaptive nonlinearity cancellation pipeline inspired by active noise cancellation in modern earphones [56]. The idea is based on Mergelyan's theorem [20], which infers that any stable linear time-invariant system can be approximated by FIR filters [12]. Modeling the microphone nonlinearity as in Equation 23, the system can be expressed as a linear combination of two known signals, enabling the intermodulation to be predicted and canceled using a feedforward strategy. Our method employs a Least Mean Squares (LMS) adaptive filter [36] with feedforward closed-loop control. By adjusting the filter order to introduce delays and up-mixing low-frequency components to f_0, we construct a virtual feedforward microphone, which provides an early reference signal that can effectively suppress intermodulation with lower delay than the received signal.

Fig. 8 illustrates the proposed DSP pipeline. All filters are FIR, and the numbers above them denote their relative orders with respect to the FIR filter inside the LMS block. The pipeline operates in two stages: training and runtime. During training, we solve the Wiener–Hopf equation [12] to determine optimal FIR weights. Band-stop filters are applied to both the virtual feedforward microphone and the received signal to remove the base frequency f_0, ensuring that the LMS optimization targets only the intermodulated components. Once convergence is achieved, the LMS-optimized FIR filter weights are fixed and reused during runtime. In runtime, the band-stop filters are replaced with appropriate delays to preserve the delay relationship between the virtual feedforward microphone and the delayed received signal. Fig. 7(b) shows results on the smartphone. Our nonlinearity cancellation reduces the intermodulated components by −20 dB, effectively suppressing them to the noise floor and reducing the phase noise.

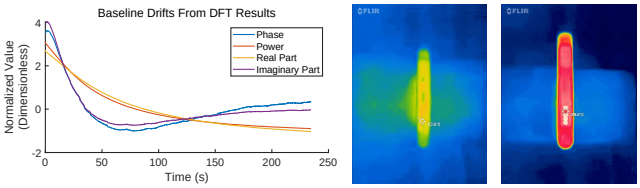
3.1.2 Speaker Phase Drift Compensation. After addressing microphone nonlinearities, we next consider distortions from smartphone speakers. Even on a static silicone pad with no motion, the single-bin DFT phase (Section 2.3) showed substantial baseline drift (Fig. 9(a)). Validation with a calibrated reference microphone [67] confirmed the drift originates in the speaker, and was consistent across frequencies and devices. This behavior is consistent with prior observations in related acoustic systems [11, 19, 68] and with studies of loudspeaker behavior [2, 10].

To better understand this effect, we examined the drift phenomena in detail. It exhibited two stages: an initial rapid decrease followed by a slower, near-linear trend. Phasor analysis revealed the imaginary component tracked the drift. Modeling the speaker as a passive element with complex impedance [89], this decline in the imaginary part indicates reduced inductance.

These observations suggested that heating drives the initial drift. To mitigate this, we introduced a warm-up procedure where the speaker plays audio before measurement, which effectively suppresses the first stage. Infrared imaging [62] confirmed a 5 °C rise in temperature after 100 s of playback (Figs. 9(c), 9(b)). Prior studies likewise show that rising temperature lowers coil inductance [2, 10], explaining the observed baseline drift.

To address the residual slow drift, we implemented BEADS (baseline estimation and drift suppression). BEADS is well-suited to blood flow velocity signals, which combine a smooth baseline with sparse pulsatile features. It enforces baseline smoothness through high-order difference penalties while preserving sharp systolic upstrokes and peaks via asymmetric sparsity constraints. With $F_S = 48$ kHz and $N = 4096$, we set the baseline cutoff frequency to 0.5 Hz and use an asymmetry penalty factor of 2.

A key trade-off is that BEADS removes all low-frequency components, including the static path phase θ_S from Equation 14. To restore θ_S, we record a short segment in which no blood flow is present (e.g., over a region without an artery). This additional step ensures that BEADS removes drift without erasing essential



(a) Baseline drift in single-bin DFT results. (b) IR image when idling. (c) IR image when playing.

Figure 9: Speaker heat-up causes baseline drift.

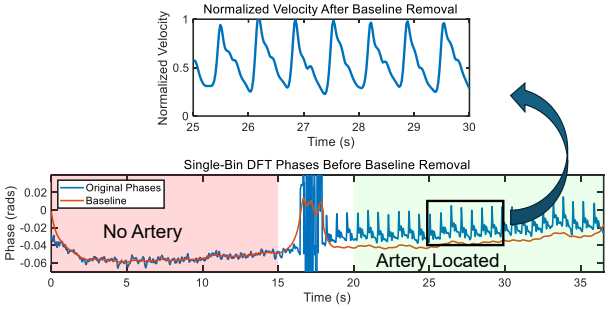
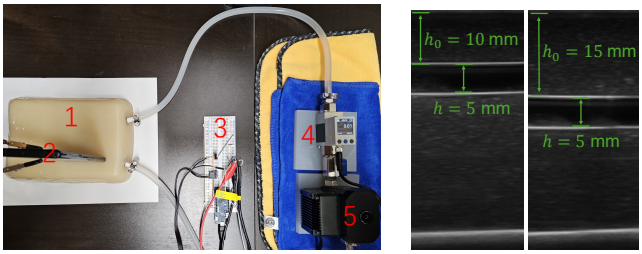


Figure 10: Baseline removal and denoising successfully remove baseline while preserve useful DC information.



(a) Experimental setup: 1) Silicone phantom; 2) A smartphone with its microphone and speaker in firm contact with the phantom; 3) Arduino controller; 4) Flow meter; 5) Pump and reservoir.

(b) Phantom arteries (d = 5 mm) located at depths of 10 mm and 15 mm.

Figure 11: Silicone phantom model experiment setup.

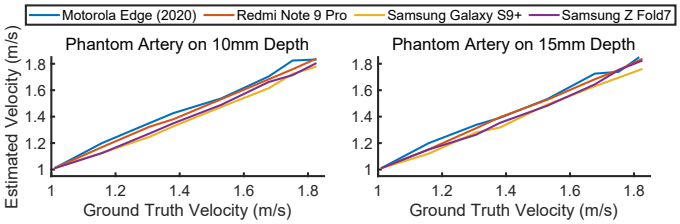


Figure 12: Flow velocities of phantom model experiment.

phase information, providing a stable reference for Doppler analysis. Fig. 10 illustrates an example waveform before and after BEADS baseline removal, highlighting how the additional recording step enables recovery of θ_S. Together, the warm-up procedure and BEADS correction ensure that commodity smartphone speakers provide a reliable phase reference for Doppler measurements.

3.1.3 Method Verification by Silicone Phantom. We validated our velocity measurement method using a silicone phantom model (Fig. 11). The phantom contained simulated arteries of $d = 5$ *mm* diameter at depths of 10 mm and 15 mm, as shown in the ultrasound cross-section (Fig. 11(b)). A closed-loop fluid system was constructed to circulate an alumina suspension with 6 μm particles, mimicking the acoustic scattering of blood. Flow was driven by a variable-speed pump (Bykski CP-SR-PMD3-X [17]) and measured by a vortex flow meter (ASAIR AFD5004A10 [64]). The Arduino board controlled pump speed and captured the flow meter's analog output. Smartphone recordings were synchronized with ground-truth measurements via Android Debug Bridge (ADB).

Calibration of the linear bias term θ_S and slope a in Equation 21 was performed using 0 m/s and 1 m/s flows in the phantom. All smartphones were tested with $F_s = 48$ kHz, $f_0 = 21$ kHz, $N = 4096$. This setup provided a ground-truth reference for validating smartphone velocity estimates, demonstrating strong agreement with flow meter measurements (Fig. 12).

3.2 Blood Flow Velocity Measurement in Human Subjects

Building on the phantom validation, we next evaluated our method in human subjects. In vivo measurements introduce two challenges: arteries are concealed beneath the skin, making localization difficult, and motion during measurement can introduce artifacts that must be detected and rejected. To address these issues, we developed an interactive interface with audio and haptic feedback to guide users to the correct artery position. The system computes a signal quality score to confirm alignment and an IMU-based score to detect interfering motion. Together, these mechanisms form a comprehensive signal quality check, ensuring that only reliable measurements are used in human-subject experiments.

3.2.1 Real-Time Feedback for Locating Arteries. In human-subject measurements, accurate localization of arteries is essential. Because arteries are soft and can shift when pressed by the smartphone, misalignment can reduce signal quality. We observed that proper alignment yields both Doppler velocity signals and low-frequency arterial pulse sounds, with heart rates from the two sources matching. This motivates a real-time feedback mechanism to guide users during localization.

For the Doppler signal, we detect peaks using the Slope Sum Function (SFF) [85], with an adaptive threshold based on the Exponential Moving Average (EMA):

$$\begin{aligned} SFF(t) &= \sum_n \max(\phi(n) - \phi(n-1), 0), \\ T_{EMA}(t) &= \delta \cdot SFF(t) + (1-\delta) T_{EMA}(t-1). \end{aligned} \tag{24}$$

For arterial pulse sounds, we apply the Shannon Energy Envelope (SEE) [53] after band-pass filtering the microphone signal between 20–200 Hz. Peaks are then identified using a zero-crossing method:

$$SEE(t) = -s_{LOW}(t)^2 \log(s_{LOW}(t)^2). \tag{25}$$

Real-time feedback is triggered when an SFF peak and SEE zero-crossing align within $\tau = 100$ ms. In that case, the smartphone

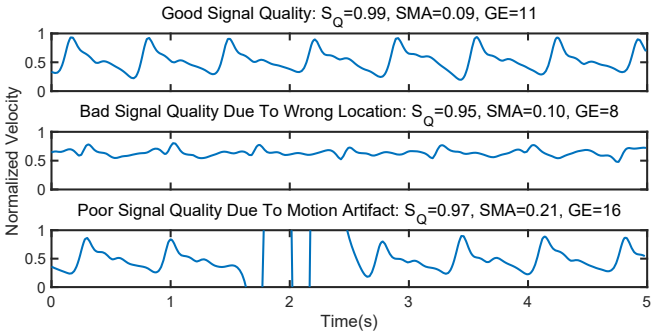


Figure 13: Our triple-modality signal quality check rejects poor signals caused by incorrect artery location or interfering motion.

provides both audio (beep) and haptic (vibration) cues to confirm correct placement. Together, this feedback loop improves usability and ensures reliable localization of arteries during measurement.

3.2.2 Triple-Modality Signal Quality Check. Real-time feedback guides the user toward arteries, but instantaneous cues may produce false positives. To ensure reliable measurements, we add a triple-modality signal quality check that combines Doppler flow, arterial sounds, and IMU data.

When alignment and contacting force are appropriate, the heart rate from the Doppler velocity waveform closely matches that from arterial pulse sounds. We define a Doppler–arterial consistency score S_Q:

$$S_Q = 1 - \frac{|BPM_{\text{Doppler}} - BPM_{\text{Arterial}}|}{BPM_{\text{Doppler}}}, \tag{26}$$

where BPM_{Doppler} and BPM_{Arterial} are estimated by autocorrelation and peak-finding.

To capture interfering motion, we use accelerometer and gyroscope signals. For the accelerometer, we compute the Signal Magnitude Area (SMA):

$$SMA(t) = \frac{1}{N}\sum_{i=0}^{N-1}(|\mathbf{a}_x| + |\mathbf{a}_y| + |\mathbf{a}_z|), \tag{27}$$

where $\mathbf{a}_x, \mathbf{a}_y, \mathbf{a}_z$ are linear accelerations (excluding gravity via the Android API [82]). For the gyroscope, we compute the Gyroscopic Energy (GE):

$$GE(t) = \sum_{i=0}^{N-1}\sqrt{\omega_x^2 + \omega_y^2 + \omega_z^2}, \tag{28}$$

where $\omega_x, \omega_y, \omega_z$ are angular velocities.

In practice, S_Q is computed with a 15 s window and 1 s hop. Signals are accepted if $S_Q \geq 0.96$, $SMA < 0.15$ m/s^2, and $GE < 15$ rad/s. Fig. 13 illustrates how this check rejects degraded recordings while retaining high-quality signals, ensuring that only reliable measurements are used in velocity analysis.

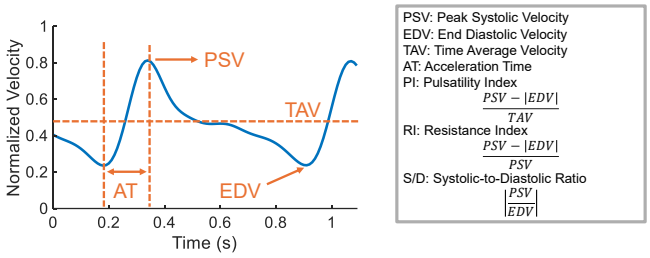


Figure 14: Definition of Doppler indices.

3.3 Computing Doppler Indices

From the reconstructed velocity waveform, we can observe clear physiological patterns (Fig. 14). The peak of the waveform corresponds to the Peak Systolic Velocity (PSV), while the trough corresponds to the End Diastolic Velocity (EDV). Both values can be reliably extracted using standard peak-finding methods. The mean velocity between PSV and EDV defines the Time-Averaged Velocity (TAV), and the interval from EDV to the next PSV gives the Acceleration Time (AT). Several Doppler indices are derived from these measurements. The Systolic-to-Diastolic Ratio (SD) is PSV/EDV. The Pulsatility Index (PI) is $(PSV - EDV)/TAV$, and the Resistance Index (RI) is $(PSV - EDV)/PSV$. As shown in Equation 14, the linear slope term a scales absolute velocity values. Because these indices are defined as ratios, they remain invariant to scaling. Thus, even a normalized waveform is sufficient for computing accurate Doppler indices, enabling reliable analysis with commodity smartphones.

3.4 System Implementation

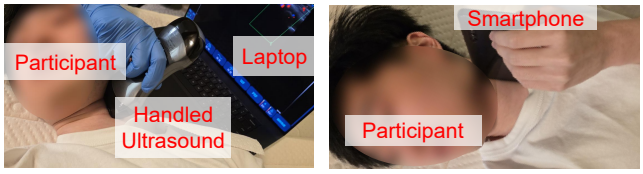
Fig. 6 shows the workflow. At launch, the bottom speaker transmits a continuous-wave (CW) ultrasound at f_0, while the microphone records in parallel to train the LMS filter in a background thread. To avoid system-level effects such as automatic gain control and equalization, audio I/O is implemented in C++ with the Android AAudio API [81].

After the user presses Start, the real-time peak detection algorithm (Section 3.2.1) runs while a 10 s buffer is maintained to compute the signal quality score S_Q. Once S_Q exceeds the threshold, the system prompts the user to hold the device steady for 15 s. During this period, feedback is paused and IMU data is recorded.

At the end of the measurement, S_Q and IMU motion scores (SMA and GE, Section 3.2.2) are checked. If thresholds are exceeded, the user re-locates the artery and repeats; otherwise Doppler index estimation begins. A 10 s pre-recording before the measurement is retained to estimate the θ_S.

Processing starts with nonlinearity cancellation using the virtual feedforward microphone (Section 3.1.1) on the band-passed CW ultrasound. Single-bin DFT (Equation 45) is then applied to segments of size N to extract phase, followed by BEADS baseline removal (Section 3.1.2) using the pre-recording to recover θ_S. Subtracting θ_S yields the normalized velocity waveform. From this, PSV and EDV are detected and Doppler indices are computed.

Finally, the system generates a report with the velocity waveform and Doppler indices, completing the pipeline on commodity smartphones.



(a) Ground-truth measurement. (b) Smartphone measurement.

Figure 15: Our human experiment setup.

4 Evaluation

After implementing our system on smartphones, we conducted a series of IRB-approved experiments with human subjects to evaluate its performance (Fig. 15). Unless otherwise noted, experiments were carried out using a Motorola Edge (2020) smartphone, configured with $f_s = 48$ kHz and $f_0 = 21$ kHz. For the single-bin DFT, we set $N = 1024$ when extracting Acceleration Time (AT), and $N = 4096$ for all other Doppler indices. To obtain ground-truth measurements, we used a commercial medical Doppler device [27].

We followed established medical practice guidelines [87]: each subject lay supine and rested for five minutes before measurements began. During testing, subjects held the smartphones themselves, placing the speaker and bottom microphone directly on the skin. Performance was quantified using Mean Absolute Error (MAE) and Accuracy (ACC), defined as

$$\text{MAE} = \frac{1}{n}\sum_{i}^{n}\left|\frac{y[i]-\hat{y}[i]}{y[i]}\right|, \quad \text{ACC} = (1-\text{MAE}), \tag{29}$$

where y is the ground truth and $\hat{y}$ is the estimated value.

4.1 Overall Doppler Index Accuracies

We evaluated the system on the left external carotid artery in 20 healthy participants recruited externally via email lists, independent of the authors' team. The participants include 14 self-identified males and 6 self-identified females, with ages ranging from 21 to 42 years. Body Mass Index (BMI) values ranged from 17.1 to 35.0, with a mean of 23.5. For each subject, ground-truth Doppler indices were collected with a medical device, then smartphone measurement was performed. The procedure was repeated three times, yielding 60 trials in total.

As shown in Figures 16(a)–16(d), our method achieved accuracies of 87.48% for Acceleration Time (AT), 81.19% for Pulsatility Index (PI), 80.39% for Systolic/Diastolic Ratio (S/D), and 94.92% for Resistance Index (RI), demonstrating that smartphones can recover clinically meaningful Doppler indices comparable to medical devices.

4.2 Different Impact Factors

Beyond overall performance, we further examined how impact factors influence system accuracy. Unless otherwise specified, each condition was tested with 5 subjects and 2 repetitions per subject, yielding 10 samples for computing MAE.

4.2.1 Impact of Arteries. To test generalizability across anatomical sites, we measured Doppler indices on several arteries. In addition to the left carotid artery (common in cardiovascular assessment), we evaluated the right superficial temporal artery (linked to cerebrovascular disease), the left radial artery (wrist), and the posterior tibial artery (ankle), both used in peripheral artery disease diagnosis. As shown in Fig. 16(e), the MAE remains within a reasonable range across all four arteries, demonstrating that the system can adapt to different vascular locations.

4.2.2 Impact of Environmental Noise. Section 3.1.1 showed that microphone nonlinearity makes our Doppler method sensitive to audible sounds. To evaluate robustness, we tested performance under varying noise levels. Experiments were conducted in a room with a TV playing human conversation. Sound Pressure Levels (SPLs) were calibrated using a reference microphone [67], ranging from 30 dBA (quiet room) to 80 dBA (vacuum cleaner). Each condition was repeated 10 times for one subject. Fig. 16(f) reports the MAE for four Doppler indices. Without nonlinearity cancellation, errors rose sharply above 50 dBA. With cancellation enabled, the system remained stable, achieving MAE $< 20\%$ even at 80 dBA.

4.2.3 Impact of Sampling Rate. Prior work [19] suggested that higher sampling rates improve acoustic sensing. We tested this by reconfiguring the Motorola Edge (2020) to operate at $F_S = 96$ kHz and $F_S = 192$ kHz. To preserve frequency resolution (Equation 19), the single-bin DFT size N was scaled proportionally with F_S. Fig. 16(i) shows that higher sampling rates enhance performance, with 192 kHz achieving less than 6% MAE across all four Doppler indices.

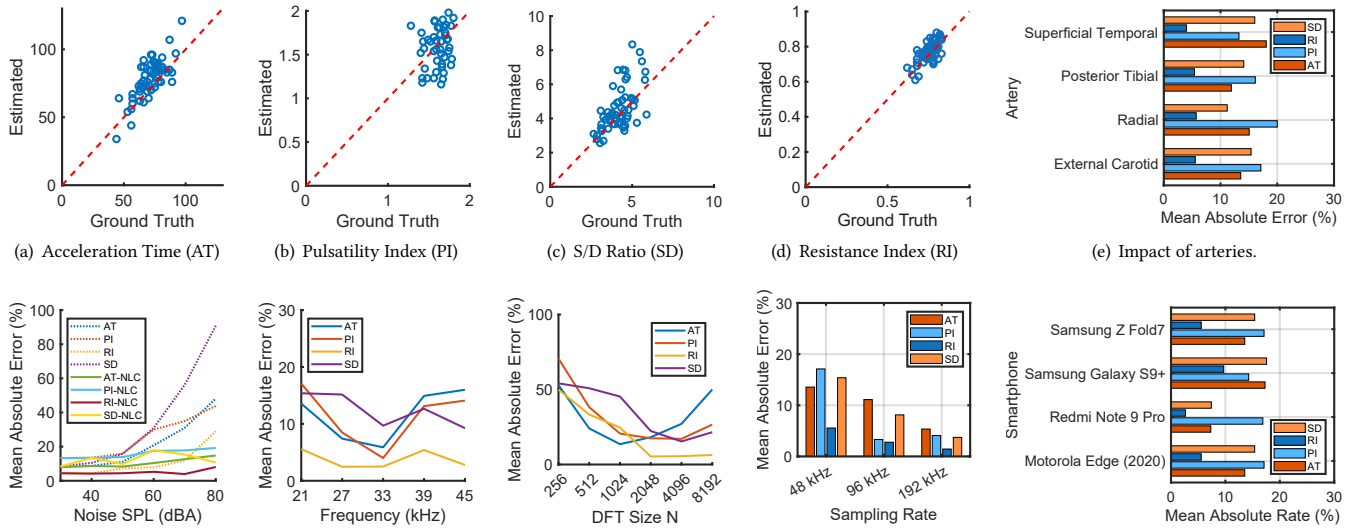
4.2.4 Impact of Transmitting Frequency. Because Doppler shift is proportional to the transmit frequency f_0 (Equation 14), increasing f_0 may improve performance. However, transmitting above 24 kHz requires higher sampling rates to satisfy the Nyquist criterion. Using the Motorola Edge (2020) at $F_S = 192$ kHz, we varied f_0 from 21 kHz to 45 kHz in 6 kHz steps, repeating each condition 5 times for 2 subjects. As shown in Fig. 16(g), error decreased until $f_0 = 33$ kHz, then rose at higher frequencies. This trend matches the device's frequency response, which begins to degrade above 35 kHz [19].

4.2.5 Impact of Single-bin DFT Size. Equation 19 shows that increasing the single-bin DFT size N improves frequency resolution but reduces temporal resolution. If N is too large, waveforms can alias across cardiac cycles. We varied N from 256 to 8192 and repeated the measurements. Fig. 16(h) shows the MAE across four Doppler indices: Acceleration Time (AT) performed best at $N = 1024$, while the other indices reached optimal accuracy at $N = 4096$.

4.2.6 Impact of Smartphones. We also evaluated performance across different smartphones. Four smartphones were tested: Motorola Edge (2020), Redmi Note 9 Pro, Samsung Galaxy S9+, and Samsung Z Fold7. Each device was tested with 2 subjects and 5 repetitions. As shown in Fig. 16(j), the MAE remains consistently low across devices, indicating that the system is robust to differences in smartphone hardware.

4.3 User Study

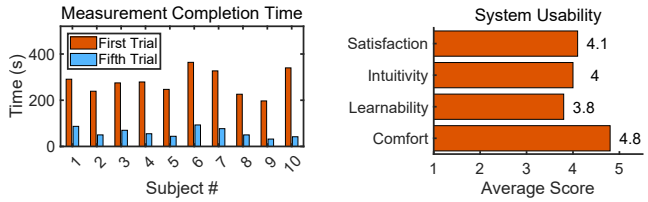
To assess usability, we conducted a study with 10 first-time users. After a brief introduction to the app, participants performed five Doppler measurements each. For every trial, we recorded the time



(a) Acceleration Time (AT) (b) Pulsatility Index (PI) (c) S/D Ratio (SD) (d) Resistance Index (RI) (e) Impact of arteries.

(f) Impact of ambient noise levels. (g) Impact of transmitting frequencies. (h) Impact of single-bin DFT size N. (i) Impact of sampling rates F_S. (j) Impact of smartphones.

Figure 16: Evaluation Results.



(a) Time required to complete a Doppler indices measurement decreases significantly from the first to the fifth trial.

(b) Usability scores from the user study indicate overall positive feedback on the system.

Figure 17: User study results.

from pressing *Start* until the recording passed the signal quality check (Section 3.2.2) and Doppler indices were displayed. This included both artery localization and any retries. Afterward, participants rated their experience on a 5-point Likert scale, covering overall satisfaction, intuitiveness of feedback, ease of learning, and comfort. Figure 17(b) shows consistently high usability scores, while Fig. 17(a) demonstrates a strong learning effect, with measurement time dropping sharply between the first and fifth trial. These results confirm that our system supports reliable at-home Doppler measurements without professional assistance.

5 Discussion

Clinical context of measured waveforms and Doppler indices: To contextualize the practical significance of our results, it is important to consider how Doppler ultrasound measurements are interpreted in clinical practice. Clinicians do not rely on isolated numerical values; instead, diagnosis is based on an integrated assessment of waveform morphology and Doppler indices, interpreted alongside anatomical and clinical information. This perspective is reflected in regulatory frameworks such as FDA-2017-D-5372 and IEC-60601-2-37, which do not impose strict accuracy thresholds on individual parameters. Rather, device performance is typically validated against benchmarks, often using flow phantoms, as in Section 3.1.3. As shown by the representative conditions summarized in Table 1, variations in waveform shape, resistance patterns, and timing characteristics are strongly associated with a range of vascular pathologies across arterial sites. This emphasizes that clinical utility depends not only on numerical accuracy but also on faithful preservation of waveform shapes. Our results indicate that Phonotonos can reliably capture these characteristics, suggesting its potential to support diagnostic assessment in real-world settings, particularly where flexible, scalable, or resource-efficient solutions are needed.

Comparison with existing ultrasound technologies: We further contextualize Phonotonos by comparing it with conventional in-hospital ultrasound systems and handheld Doppler ultrasound devices (Table 2). Hospital systems provide comprehensive imaging and clinically validated diagnostics but require bulky, expensive equipment and trained operators, restricting use to clinical environments; handheld devices improve portability and cost but still rely on dedicated hardware, gel, and trained users. In contrast, Phonotonos adopts a fundamentally different paradigm as a software-only system running on commodity smartphones, eliminating additional hardware while enabling low-cost, highly portable deployment in home and telehealth settings. Although it does not offer full imaging and currently focuses on estimating scaled blood flow velocity and Doppler indices, Phonotonos prioritizes accessibility and ease of use for untrained users. As such, it complements rather than replaces ultrasound technologies, being well suited for pre-screening and longitudinal monitoring, while traditional ultrasound remains essential for definitive diagnosis and detailed clinical assessment.

Artery	Condition / Disease	Diagnosis guideline	References
External carotid	Intracranial dural arteriovenous fistulae	Low-resistance waveform, continuous forward diastolic flow, reduced pulsatility, RI ≤ 0.7	[43, 103]
External carotid	Internalization of the external carotid artery	Increased diastolic forward flow, loss of the usual high-resistance contour, and lower PI.	[43, 101]
Superficial temporal	Giant cell arteritis	Distal severe stenosis/occlusion may produce damped/tardus–parvus waveforms.	[15, 25, 43]
Radial	Systemic sclerosis w/ Raynaud's phenomenon	Elevated RI with a higher-resistance spectral pattern, RI ≥ 0.85	[24, 43]
Radial	Proximal upper-extremity inflow disease	Tardus–parvus waveform with prolonged AT (>140 ms) and rounded systolic peak.	[1, 43, 102]
Radial	Hand ischemia	Rretrograde or bidirectional distal radial flow.	[43]
Posterior tibial	Peripheral arterial disease	fProlonged upstroke, damped monophasic waveform, AT $\geq$ 121 ms	[43, 113]
Posterior tibial	Diabetic foot ischemia	Monophasic or absent waveform. Intermediate: AT $\geq$ 120 ms; Severe: AT $\geq$ 160 ms	[40, 41, 102, 105]

Table 1: Example diseases and conditions that can be diagnosed using scaled velocity waveform and Doppler indices.

	Hospital Ultrasound	Handheld Ultrasound	**Phonotonos**
Functionalities	Full functionalities	Imaging, velocity, Doppler indices	Scaled velocity, Doppler indices
Hardware	Bulky equipment	Handled scanner	Smartphone only
Cost	$20,000+	$3,000+	$0 (software-only)
Portability	Low	Medium	High
Deployment	Hospital	Clinic, point-of-care	Home, telehealth
Diagnosis Capabilities	Historically proven	FDA-approved	Pre-screening (in development)
Gel Needed	Yes	Yes	No
Operator	Trained sonographer	Trained sonographer	End user
Ease of Use	Operator-dependent	Operator-dependent	Simple

Table 2: Comparison of Phonotonos vs. existing technologies.

Limitations: Several limitations remain. First, because the effective insonation angle created by speaker–microphone asymmetry is unknown, our system reconstructs only scaled blood flow velocity waveforms. Still, Doppler indices and waveform morphology are sufficient for diagnosing many cardiovascular conditions [76]. Second, we rely on CW rather than PW Doppler, since smartphones cannot generate short pulses or precise time-gating. This prevents depth discrimination, meaning venous signals may mix with arterial ones. Fortunately, venous flow is slower and smoother [43], and can be suppressed using our baseline removal method (Section 3.1.2). Finally, higher sampling rates and transmit frequencies improve accuracy (Sections 4.2.3, 4.2.4), and modern hardware often supports up to 96 kHz [19]. However, most manufacturers do not expose these sampling rates in user APIs, limiting consistency across devices until broader software support becomes available.

Future Work: This work demonstrates, for the first time, that smartphones can measure scaled blood flow velocity, introducing a new sensing modality alongside camera-based PPG [63] and microphone-based PCG [21]. Together, these modalities enable more comprehensive cardiovascular monitoring using a single smartphone. However, results from a sample of 20 healthy participants are insufficient to establish clinical value. Future work will focus on clinical validation with larger patient populations with known conditions to evaluate diagnostic utility. Beyond diagnosis, integrating Doppler sensing with other modalities to train machine learning models could further extend the capabilities of smartphone-based health monitoring.

6 Related Work

Commodity speakers and microphones have been widely repurposed for physiological sensing, including respiration and sleep monitoring [70, 109], overdose detection [69], pulmonary assessment [31, 46, 95], and cardiac monitoring [83, 107]. These systems demonstrate that acoustics can track breathing, lung function, and heart-induced vibrations. Acoustoscillogram [54] further shows that earphones can transmit ultrasound to measure Pulse Wave Velocity (PWV) by detecting skin-surface vibrations caused by arterial displacement. However, none of these approaches perform *through-skin* sensing of subsurface motion, which has traditionally required MHz ultrasound. Our work is the first to recover blood-flow velocity beneath the skin using only smartphone acoustics.

Beyond physiology, acoustic sensing has enabled diverse applications, including object imaging [55], ice thickness measurement [57], ambient temperature estimation [18, 106], robust speech interaction [99, 124, 125], indoor mapping [65, 80], gesture recognition [29, 35, 50, 52, 116, 119], finger tracking [34, 35, 71, 108, 117], eye blink detection [58] and multipath-resilient localization [121]. Other modalities—such as Wi-Fi, UWB [47, 60, 118, 122], cameras [63], or visible light PPG [110] can support physiological sensing but typically require additional hardware, higher power, or constrained operating conditions. In contrast, acoustic sensing is lightweight and universally available on commodity devices, albeit limited to sub-24 kHz bandwidth. Recent advances have improved this bandwidth [19] and system robustness [49, 59]. Our work builds on this foundation and extends acoustic sensing to the Doppler regime, bridging consumer audio hardware with capabilities traditionally reserved for medical ultrasound.

7 Conclusion

We presented *Phonotonos*, the first system to achieve through-skin Doppler sensing of blood-flow velocity on commodity smartphones. By combining single-bin DFT phase analysis with adaptive nonlinearity cancellation, baseline drift suppression, and triple-modality artery localization, the system overcomes hardware constraints and user variability. Human-subject experiments show that Phonotonos can measure clinically relevant Doppler indices with promising accuracies. Transforming smartphones into accessible vascular monitors, Phonotonos moves us toward scalable, low-cost cardiovascular screening and everyday health monitoring.

Acknowledgments

This work was supported in part by the University of Maryland Baltimore Institute for Clinical and Translational Research (ICTR) Accelerated Translational Incubator Pilot (ATIP) grant and the Google Research Scholar Award. This work was also supported in part by the National Science Foundation (NSF) under Grant No. 2230143, as well as by the Nanyang Technological University SUG-NAP program.

Generative AI tools were used for language polishing and figure illustration assistance. Fig. 2,4,6 contain image by Figdraw.com.

References

[1] American Institute of Ultrasound in Medicine. 2021. AIUM Practice Parameter for the Performance of Peripheral Arterial Ultrasound Examinations Using Color and Spectral Doppler Imaging. *Journal of Ultrasound in Medicine* 40, 5 (2021), E17–E24. https://doi.org/10.1002/jum.15643

[2] Ronny Andersson. 2008. Loudspeaker voice-coil temperature estimation.

[3] Bjorn A. J. Angelsen. 1980. A Theoretical Study of the Scattering of Ultrasound from Blood. *IEEE Transactions on Biomedical Engineering* BME-27, 2 (1980), 61–67. https://doi.org/10.1109/TBME.1980.326708

[4] Joachim O Arndt, Jürgen Klauske, and Frank Mersch. 1968. The diameter of the intact carotid artery in man and its change with pulse pressure. *Pflüger's Archiv für die gesamte Physiologie des Menschen und der Tiere* 301, 3 (1968), 230–240.

[5] Peter Atkinson and Peter NT Wells. 1977. Pulse-Doppler ultrasound and its clinical application. *The Yale journal of biology and medicine* 50, 4 (1977), 367.

[6] Danilo Avola, Luigi Cinque, Alessio Fagioli, Gianluca Foresti, and Alessio Mecca. 2021. Ultrasound medical imaging techniques: a survey. *ACM Computing Surveys (CSUR)* 54, 3 (2021), 1–38.

[7] Haim Azhari. 2010. *Basics of Biomedical Ultrasound for Engineers*. Wiley-IEEE Press.

[8] Mingsian R Bai and Jerwoei Liao. 2005. Acoustic analysis and design of miniature loudspeakers for mobile phones. *Journal of the Audio Engineering Society* 53, 11 (2005), 1061–1076.

[9] Ralf Bauer, Yansheng Zhang, Joseph C Jackson, William M Whitmer, W Owen Brimijoin, Michael A Akeroyd, Deepak Uttamchandani, and James FC Windmill. 2017. Influence of microphone housing on the directional response of piezoelectric MEMS microphones inspired by Ormia ochracea. *IEEE Sensors Journal* 17, 17 (2017), 5529–5536.

[10] Gottfried Behler, U Spätling, and T Arimont. 1995. Measuring the Loudspeaker's Impedance During Operation for the Evaluation of the Voice Coil Temperature. In *Audio Engineering Society Convention 98*. Audio Engineering Society.

[11] Marco Berkhout, Lutsen Dooper, and Benno Krabbenborg. 2013. A 4Ω 2.65 W class-D audio amplifier with embedded DC-DC boost converter, current sensing ADC and DSP for adaptive speaker protection. *IEEE journal of solid-state circuits* 48, 12 (2013), 2952–2961.

[12] Holger Boche and Volker Pohl. 2006. Structural properties of the wiener filter—Stability, smoothness properties, and fir approximation behavior. *IEEE transactions on information theory* 52, 9 (2006), 4272–4282.

[13] J Boersma. 1960. Computation of Fresnel integrals. *Math. Comp.* (1960).

[14] British Heart Foundation. 2023. Global Heart & Circulatory Diseases Factsheet. https://www.bhf.org.uk/-/media/files/for-professionals/research/heart-statistics/bhf-cvd-statistics-global-factsheet.pdf Accessed: 2024-11-29.

[15] British Medical Ultrasound Society. 2021. Guidance for Giant Cell Arteritis Ultrasound and Service Provision. BMUS guidance document. November 2021.

[16] William R Brody and James D Meindl. 1974. Theoretical analysis of the CW Doppler ultrasonic flowmeter. *IEEE transactions on biomedical engineering* 3 (1974), 183–192.

[17] Bykski.us. 2025. *Bykski Versatile 12V+PWM DDC POM Advanced Pump for Custom Water Cooling Loops (CP-SR-PMD3-X)*. https://www.bykski.us/products/bykski-versatile-12v-pwm-ddc-pom-advanced-pump-for-custom-water-cooling-loops-cp-sr-pmd3-x

[18] Chao Cai, Henglin Pu, Liyuan Ye, Hongbo Jiang, and Jun Luo. 2021. Active acoustic sensing for "hearing" temperature under acoustic interference. *IEEE Transactions on Mobile Computing* 22, 2 (2021), 661–673.

[19] Shirui Cao, Dong Li, Sunghoon Ivan Lee, and Jie Xiong. 2023. PowerPhone: Unleashing the Acoustic Sensing Capability of Smartphones. In *Proceedings of the 29th Annual International Conference on Mobile Computing and Networking*. 1–16.

[20] Lennart Carleson. 1964. Mergelyan's theorem on uniform polynomial approximation. *Math. Scand.* 15, 2 (1964), 167–175.

[21] Tao Chen, Yongjie Yang, Xiaoran Fan, Xiuzhen Guo, Jie Xiong, and Longfei Shangguan. 2024. Exploring the feasibility of remote cardiac auscultation using earphones. In *Proceedings of the 30th Annual International Conference on Mobile Computing and Networking*. 357–372.

[22] TDK Corporation. 2020. *Using MEMS Microphones in Directional Applications*. https://invensense.tdk.com/download-pdf/an-000172-using-mems-microphones-in-directional-applications-v1-0/

[23] TDK Corporation. 2024. *SmartSound™ Analog Microphones*. https://invensense.tdk.com/smartsound/analog/

[24] Roberto D'Alessandro, Estrella Garcia Gonzalez, Paolo Falsetti, Edoardo Conticini, Miriana d'Alessandro, Elena Selvi, Francesca Bellisai, Virginia Berlengiero, Giulia Vallifuoco, Anna Paola Pata, Caterina Baldi, Luca Cantarini, Elena Bargagli, and Bruno Frediani. 2023. Peripheral Macrovascular Involvement in Systemic Sclerosis: A Cohort Study by Color and Spectral Doppler Ultrasonography. *Life* 13, 2 (2023), 487. https://doi.org/10.3390/life13020487

[25] Christian Dejaco, Sofia Ramiro, Michael Bond, et al. 2024. EULAR recommendations for the use of imaging in large vessel vasculitis in clinical practice: 2023 update. *Annals of the Rheumatic Diseases* (2024). Published as the 2023 EULAR update.

[26] DXOMARK. 2024. *Smartphones Audio Reviews*. https://www.dxomark.com/category/audio-reviews/

[27] EagleView. 2024. *EagleView: Advanced Aerial Imagery and Property Insights*. https://geteagleview.com/

[28] Isabelle Fontaine, Michel Bertrand, and Guy Cloutier. 1999. A system-based approach to modeling the ultrasound signal backscattered by red blood cells. *Biophysical journal* 77, 5 (1999), 2387–2399.

[29] Zhihui Gao, Ang Li, Dong Li, Jialin Liu, Jie Xiong, Yu Wang, Bing Li, and Yiran Chen. 2022. Mom: Microphone based 3d orientation measurement. In *2022 21st ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*. IEEE, 132–144.

[30] Oleg A Godin. 2013. Rayleigh scattering of a spherical sound wave. *The Journal of the Acoustical Society of America* 133, 2 (2013), 709–720.

[31] Mayank Goel, Elliot Saba, Maia Stiber, Eric Whitmire, Josh Fromm, Eric C Larson, Gaetano Borriello, and Shwetak N Patel. 2016. Spirocall: Measuring lung function over a phone call. In *Proceedings of the 2016 CHI conference on human factors in computing systems*. 5675–5685.

[32] Karlheinz Gröchenig. 2003. Uncertainty principles for time-frequency representations. *Advances in Gabor analysis* (2003), 11–30.

[33] Scott P Grogan and Cristin A Mount. 2023. Ultrasound physics and instrumentation. In *StatPearls [Internet]*. StatPearls Publishing.

[34] Kaiyi Guo, Tianyu Wu, Yang Gao, Qian Zhang, and Dong Wang. 2025. EchoTouch: Low-power Face-touching Behavior Recognition Using Active Acoustic Sensing on Glasses. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 9, 2 (2025), 1–33.

[35] Sidhant Gupta, Daniel Morris, Shwetak Patel, and Desney Tan. 2012. Soundwave: using the doppler effect to sense gestures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1911–1914.

[36] Monson H Hayes. 1996. *Statistical digital signal processing and modeling*. John Wiley & Sons.

[37] J Henderson, K Willson, JR Jago, and TA Whittingham. 1995. A survey of the acoustic outputs of diagnostic ultrasound equipment in current clinical use. *Ultrasound in medicine & biology* 21, 5 (1995), 699–705.

[38] GoerTek Inc. 2024. *Microphone Solutions*. https://www.goertek.com/en/content/details20_878.html

[39] GoerTek Inc. 2024. *Speakers Solutions*. https://www.goertek.com/en/content/details20_877.html

[40] International Working Group on the Diabetic Foot. 2023. IWGDF Guidelines on the Diagnosis, Prognosis and Management of Peripheral Artery Disease in Patients with Foot Ulcers in Diabetes. Guideline.

[41] R. Ismail, A. Gareeballah, A. Abdelrahim, et al. 2022. Impact of type 2 diabetes mellitus on hemodynamic and morphology of foot arteries: A duplex ultrasound evaluation. *Journal of Radiation Research and Applied Sciences* (2022).

[42] AKIRA Kamiya and TATSUO Togawa. 1980. Adaptive regulation of wall shear stress to flow change in the canine carotid artery. *American Journal of Physiology-Heart and Circulatory Physiology* 239, 1 (1980), H14–H21.

[43] Esther SH Kim, Aditya M Sharma, Robert Scissons, David Dawson, Robert T Eberhardt, Marie Gerhard-Herman, Joseph P Hughes, Steve Knight, Ann Marie Kupinski, Guillaume Mahe, et al. 2020. Interpretation of peripheral arterial and venous Doppler waveforms: A consensus statement from the Society for Vascular Medicine and Society for Vascular Ultrasound. *Vascular Medicine* 25, 5 (2020), 484–506.

[44] Jaroslaw Krejza, Zenon Mariak, and Janusz Lewko. 2003. Standardization of flow velocities with respect to age and sex improves the accuracy of transcranial color Doppler sonography of middle cerebral artery spasm. *American Journal of Roentgenology* 181, 1 (2003), 245–252.

[45] Heinrich Kuttruff. 2007. *Acoustics: an introduction*. CRC Press.

[46] Eric C Larson, Mayank Goel, Gaetano Boriello, Sonya Heltshe, Margaret Rosenfeld, and Shwetak N Patel. 2012. SpiroSmart: using a microphone to measure lung function on a mobile phone. In *Proceedings of the 2012 ACM Conference on ubiquitous computing*. 280–289.

[47] Yonggu Lee, Jun-Young Park, Yeon-Woo Choi, Hyun-Kyung Park, Seok-Hyun Cho, Sung Ho Cho, and Young-Hyo Lim. 2018. A novel non-contact heart rate monitor using impulse-radio ultra-wideband (IR-UWB) radar technology. *Scientific reports* 8, 1 (2018), 13053.

[48] Dong Li. 2026. From Hearing to Feeling: Unlocking Through-Skin Acoustic Sensing on Smartphones. *GetMobile: Mobile Computing and Communications* 29, 4 (2026), 5–10.

[49] Dong Li, Shirui Cao, Sunghoon Ivan Lee, and Jie Xiong. 2022. Experience: practical problems for acoustic sensing. In *Proceedings of the 28th Annual International Conference on Mobile Computing And Networking*. 381–390.

[50] Dong Li, Jialin Liu, Sunghoon Ivan Lee, and Jie Xiong. 2020. FM-track: pushing the limits of contactless multi-target tracking using acoustic signals. In *Proceedings of the 18th Conference on Embedded Networked Sensor Systems*. 150–163.

[51] Dong Li, Jialin Liu, Sunghoon Ivan Lee, and Jie Xiong. 2022. Lasense: Pushing the limits of fine-grained activity sensing using acoustic signals. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 1 (2022),

1–27.

[52] Dong Li, Jialin Liu, Sunghoon Ivan Lee, and Jie Xiong. 2022. Room-scale hand gesture recognition using smart speakers. In *Proceedings of the 20th ACM Conference on Embedded Networked Sensor Systems*. 462–475.

[53] H Liang, Sakari Lukkarinen, and Iiro Hartimo. 1997. Heart sound segmentation algorithm based on heart sound envelogram. In *Computers in Cardiology 1997*. IEEE, 105–108.

[54] Xiaoxuan Liang, Zhaolong Wei, Shirui Cao, Quan Zhang, Longfei Shangguan, Dong Li, and Jeremy Gummeson. 2025. Acoustoscillogram: Unlocking Arterial Health Insights via Acoustic Sensing with Low-cost Wired Earphones. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 9, 4, Article 194 (Dec. 2025), 27 pages. https://doi.org/10.1145/3770690

[55] Xiaoxuan Liang, Zhaolong Wei, Dong Li, Jie Xiong, and Jeremy Gummeson. 2024. Sondar: Size and shape measurements using acoustic imaging. In *Proceedings of the Twenty-fifth International Symposium on Theory, Algorithmic Foundations, and Protocol Design for Mobile Networks and Mobile Computing*. 361–370.

[56] Stefan Liebich, Johannes Fabry, Peter Jax, and Peter Vary. 2018. Signal processing challenges for active noise cancellation headphones. In *Speech Communication; 13th ITG-Symposium*. VDE, 1–5.

[57] Chi Lin, Yingqi Chen, Jie Xiong, Yang Chi, Xin Fan, Qiang Tong, and Guowei Wu. 2025. IceSound: Acoustic-based Ice Thickness Sensing with Smartphones. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 9, 3 (2025), 1–25.

[58] Jialin Liu, Dong Li, Lei Wang, and Jie Xiong. 2021. BlinkListener: " Listen" to Your Eye Blink Using Your Smartphone. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 2 (2021), 1–27.

[59] Jialin Liu, Dong Li, Lei Wang, Fusang Zhang, and Jie Xiong. 2022. Enabling contact-free acoustic sensing under device motion. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 3 (2022), 1–27.

[60] Jinyi Liu, Youwei Zeng, Tao Gu, Leye Wang, and Daqing Zhang. 2021. WiPhone: Smartphone-based respiration monitoring using ambient reflected WiFi signals. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 1 (2021), 1–19.

[61] Knowles Electronics, LLC. 2024. *SiSonic™ Surface Mount MEMS Microphones*. https://www.knowles.com/subdepartment/dpt-microphones/subdpt-sisonic-surface-mount-mems

[62] Teledyne FLIR LLC. 2025. *FLIR ONE Pro Pro-Grade Thermal Camera for Smartphones*. https://www.flir.com/products/flir-one-pro/?model=435-0007-03

[63] Giulio Lovisotto, Henry Turner, Simon Eberz, and Ivan Martinovic. 2020. Seeing red: PPG biometrics using smartphone cameras. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*. 818–819.

[64] Aosong Electronic Co., Ltd. 2025. *AFD5x Series Vortex Liquid Flowmeter*. http://www.aosong.com/products-218.html

[65] Wenjie Luo, Qun Song, Zhenyu Yan, Rui Tan, and Guosheng Lin. 2022. Indoor smartphone SLAM with learned echoic location features. In *Proceedings of the 20th ACM Conference on Embedded Networked Sensor Systems*. 489–503.

[66] Dev Maulik. 2005. Physical principles of Doppler ultrasonography. In *Doppler ultrasound in Obstetrics and Gynecology*. Springer, 9–17.

[67] miniDSP Ltd. 2025. *UMIK-2 USB Reference Measurement Microphone*. https://www.minidsp.com/products/acoustic-measurement/umik-2

[68] Angelo Nagari, Emmanuel Allier, François Amiard, Vincent Binet, and Christian Fraisse. 2012. An 8\$Omega$2.5 W 1%-THD 104 dB (A)-Dynamic-Range Class-D Audio Amplifier With Ultra-Low EMI System and Current Sensing for Speaker Protection. *IEEE journal of solid-state circuits* 47, 12 (2012), 3068–3080.

[69] Rajalakshmi Nandakumar, Shyamnath Gollakota, and Jacob E Sunshine. 2019. Opioid overdose detection using smartphones. *Science translational medicine* 11, 474 (2019), eaau8914.

[70] Rajalakshmi Nandakumar, Shyamnath Gollakota, and Nathaniel Watson. 2015. Contactless sleep apnea detection on smartphones. In *Proceedings of the 13th annual international conference on mobile systems, applications, and services*. 45–57.

[71] Rajalakshmi Nandakumar, Vikram Iyer, Desney Tan, and Shyamnath Gollakota. 2016. Fingerio: Using active sonar for fine-grained finger tracking. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 1515–1525.

[72] TR Nelson and DH Pretorius. 1988. The Doppler signal: where does it come from and what does it mean? *American Journal of Roentgenology* 151, 3 (1988), 439–447.

[73] Paul G Newman and Grace S Rozycki. 1998. The history of ultrasound. *Surgical clinics of north America* 78, 2 (1998), 179–195.

[74] Xiaoran Ning, Ivan W Selesnick, and Laurent Duval. 2014. Chromatogram baseline estimation and denoising using sparsity (BEADS). *Chemometrics and Intelligent Laboratory Systems* 139 (2014), 156–167.

[75] Jaakko Nisula, Juha Holm, and Aki Mäkivirta. 2013. Calculating sound radiation from loudspeaker enclosures using the finite element analysis. In *Audio Engineering Society Conference: 51st International Conference: Loudspeakers and Headphones*. Audio Engineering Society.

[76] Ammar A Oglat, MZ Matjafri, Nursakinah Suardi, Mohammad A Oqlat, Mostafa A Abdelrahman, and Ahmad A Oqlat. 2018. A review of medical doppler ultrasonography of blood flow in general and especially in common carotid artery. *Journal of medical ultrasound* 26, 1 (2018), 3–13.

[77] Augustus R Okoyenta, Haijun Wu, Xueliang Liu, and Weikang Jiang. 2020. A short survey on Green's function for acoustic problems. *Journal of Theoretical and Computational Acoustics* 28, 02 (2020), 1950025.

[78] Yasuharu Onishi, Jun Kuroda, Yukio Murata, Motoyoshi Komoda, Kazuyuki Tsunoda, Yukio Yokoyama, Yasuhiro Oikawa, and Yoshio Yamasaki. 2010. The acoustical design of mobile phones. In *20th Int. Congr. Acoust. 2010, ICA 2010-Inc. Proc. 2010 Annu. Conf. Aust. Acoust. Soc.*, Vol. 2. 1250–1257.

[79] World Health Organization. 2021. Cardiovascular diseases (CVDs). https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds). Accessed: 2024-11-26.

[80] Swadhin Pradhan, Ghufran Baig, Wenguang Mao, Lili Qiu, Guohai Chen, and Bo Yang. 2018. Smartphone-based acoustic indoor space mapping. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 2 (2018), 1–26.

[81] Android Open Source Project. 2025. *AAudio and MMAP*. https://source.android.com/docs/core/audio/aaudio

[82] Android Open Source Project. 2025. *Motion sensors*. https://developer.android.com/develop/sensors-and-location/sensors/sensors_motion

[83] Kun Qian, Chenshu Wu, Fu Xiao, Yue Zheng, Yi Zhang, Zheng Yang, and Yunhao Liu. 2018. Acousticcardiogram: Monitoring heartbeats using acoustic signals on smart devices. In *IEEE INFOCOM 2018-IEEE conference on computer communications*. IEEE, 1574–1582.

[84] Ashiqur Rahaman, Steven Boor, Christopher Bradt, Sung B Lee, and Shehab Albahri. 2024. Nonlinear behavioral model of capacitive MEMS microphone for predicting ultrasound intermodulation distortion. *IEEE Sensors Journal* 25, 1 (2024), 236–243.

[85] Shalini A Rankawat, Mansi Rankawat, and Rahul Dubey. 2015. ECG artifacts detection in noncardiovascular signals using Slope Sum Function and Teager Kaiser Energy. In *2015 International Conference on BioSignal Analysis, Processing and Systems (ICBAPS)*. IEEE, 6–10.

[86] JMJG Roevros. 1974. Analog processing of CW Doppler flowmeter signals to determine average frequency shift momentaneously without the use of a wave analyser. *Cardiovascular applications of ultrasound* (1974).

[87] Mary J Roman, Tasneem Z Naqvi, Julius M Gardin, Marie Gerhard-Herman, Michael Jaff, and Emile Mohler. 2006. Clinical application of noninvasive vascular ultrasound in cardiovascular risk stratification: a report from the American Society of Echocardiography and the Society of Vascular Medicine and Biology. *Journal of the American Society of Echocardiography* 19, 8 (2006), 943–954.

[88] Ashish Saraogi. 2015. Lung ultrasound: Present and future. *Lung India* 32, 3 (2015), 250–257.

[89] Douglas Self. 2013. *Audio power amplifier design*. Routledge.

[90] Muhammad Ali Shah, Ibrar Ali Shah, Duck-Gyu Lee, and Shin Hur. 2019. Design approaches of MEMS microphones for enhanced performance. *Journal of sensors* 2019, 1 (2019), 9294528.

[91] KK Shung, BA Krisko, and JO Ballard III. 1982. Acoustic measurement of erythrocyte compressibility. *The Journal of the Acoustical Society of America* 72, 5 (1982), 1364–1367.

[92] Koping K. Shung, Rubens A. Sigelmann, and John M. Reid. 1976. Scattering of Ultrasound by Blood. *IEEE Transactions on Biomedical Engineering* BME-23, 6 (1976), 460–467. https://doi.org/10.1109/TBME.1976.324604

[93] Koping K Shung, Rubens A Sigelmann, and John M Reid. 1977. Angular dependence of scattering of ultrasound from blood. *IEEE Transactions on Biomedical Engineering* 4 (1977), 325–331.

[94] K Kirk Shung, Michael Smith, and Benjamin MW Tsui. 1992. *Principles of medical imaging*. Academic Press.

[95] Xingzhe Song, Boyuan Yang, Ge Yang, Ruirong Chen, Erick Forno, Wei Chen, and Wei Gao. 2020. SpiroSonic: monitoring human lung function via acoustic sensing on commodity smartphones. In *Proceedings of the 26th Annual International Conference on Mobile Computing and Networking*. 1–14.

[96] Statista. 2024. *Number of smartphone subscriptions worldwide from 2016 to 2023*. https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/

[97] Lee Stoner, Manning Sabatier, Kristy Edge, and Kevin McCully. 2004. Relationship between blood velocity and conduit artery diameter and the effects of smoking on vascular responsiveness. *Journal of Applied Physiology* 96, 6 (2004), 2139–2145.

[98] Péter Studinger, Zsuzsanna Lénárd, Zsuzsanna Kováts, László Kocsis, and Mark Kollai. 2003. Static and dynamic changes in carotid artery diameter in humans during and after strenuous exercise. *The journal of physiology* 550, 2 (2003), 575–583.

[99] Ke Sun and Xinyu Zhang. 2021. UltraSE: single-channel speech enhancement using ultrasound. In *Proceedings of the 27th annual international conference on mobile computing and networking*. 160–173.

[100] Sara Svedlund and Li-Ming Gan. 2011. Longitudinal wall motion of the common carotid artery can be assessed by velocity vector imaging. *Clinical physiology and functional imaging* 31, 1 (2011), 32–38.

[101] Tomotaka Tanaka, Ryosuke Doijiri, Kozue Saito, Katsufumi Kajimoto, Masafumi Ihara, Hiroshi Yamagami, Kotaro Miyashita, and Kazuyuki Nagatsuka. 2014. Validation of ultrasound parameters to assess collateral flow via ophthalmic artery in internal carotid artery occlusion. *Journal of Stroke and Cerebrovascular Diseases* 23, 5 (2014), 1177–1182. https://doi.org/10.1016/j.jstrokecerebrovasdis.2013.10.012

[102] J.-E. Trihan, G. Mahé, J.-P. Laroche, M. Dauzat, et al. 2023. Arterial Blood-Flow Acceleration Time on Doppler Ultrasound Waveforms: What Are We Talking About? *Journal of Clinical Medicine* (2023).

[103] L.-K. Tsai, S.-J. Yeh, Y.-C. Chen, H.-M. Liu, and J.-S. Jeng. 2009. Screen for intracranial dural arteriovenous fistulae with carotid duplex sonography. *Journal of Neurology, Neurosurgery and Psychiatry* 80, 11 (2009), 1225–1229. https://doi.org/10.1136/jnnp.2008.166413

[104] Mitchell E Tublin, Ronald O Bude, and Joel F Platt. 2003. The resistive index in renal Doppler sonography: where do we stand? *American Journal of Roentgenology* 180, 4 (2003), 885–892.

[105] Guistinna Tun, Curtis Woodford, Colleen Flanagan, Peter A. Schneider, Michael S. Conte, and Clara M. Gomez-Sanchez. 2025. Validation of Pedal Acceleration Time for the Assessment of Perfusion in Patients with Chronic Limb-Threatening Ischemia. *Annals of Vascular Surgery* 121 (2025), 351–359. https://doi.org/10.1016/j.avsg.2025.07.014

[106] Haoran Wan, Lei Wang, Ting Zhao, Ke Sun, Shuyu Shi, Haipeng Dai, Guihai Chen, Haodong Liu, and Wei Wang. 2022. Vector: Velocity based temperature-field monitoring with distributed acoustic devices. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 3 (2022), 1–28.

[107] Anran Wang, Dan Nguyen, Arun R Sridhar, and Shyamnath Gollakota. 2021. Using smart speakers to contactlessly monitor heart rhythms. *Communications biology* 4, 1 (2021), 319.

[108] Wei Wang, Alex X Liu, and Ke Sun. 2016. Device-free gesture tracking using acoustic signals. In *Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking*. 82–94.

[109] Xuyu Wang, Runze Huang, and Shiwen Mao. 2017. SonarBeat: Sonar phase for breathing beat monitoring with smartphones. In *2017 26th International Conference on Computer Communication and Networks (ICCCN)*. IEEE, 1–8.

[110] Yuna Watanabe, Natasha Yamane, Aarti Sathyanarayana, Varun Mishra, and Matthew S Goodwin. 2025. Beyond Motion Artifacts: Optimizing PPG Preprocessing for Accurate Pulse Rate Variability Estimation. In *Companion of the 2025 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 1176–1182.

[111] Keith A Wear. 2001. Ultrasonic attenuation in human calcaneus from 0.2 to 1.7 MHz. *IEEE transactions on ultrasonics, ferroelectrics, and frequency control* 48, 2 (2001), 602–608.

[112] PNT Wells. 1975. Absorption and dispersion of ultrasound in biological tissue. *Ultrasound in medicine & biology* 1, 4 (1975), 369–376.

[113] Siem A. Willems, Obrecht O. van Bennekom, Abbey Schepers, Jan van Schaik, Joost R. van der Vorst, Jaap F. Hamming, and Jeroen J. W. M. Brouwers. 2025. A Diagnostic Comparison Study between Maximal Systolic Acceleration and Acceleration Time to Detect Peripheral Arterial Disease. *Annals of Vascular Surgery* 111 (2025), 203–211. https://doi.org/10.1016/j.avsg.2024.10.021

[114] Kolja Willimzik and Gryt Willimzik AB. 2024. *Studies on the Structural Vibration and Cavity Resonance Behavior of Various Loudspeaker Enclosures and Baffles*. https://www.willimzik.se/EnclosureBaffleStudies.pdf

[115] Nicholas J Willis. 2005. *Bistatic radar*. Vol. 2. SciTech Publishing.

[116] Yongjie Yang, Tao Chen, Zhenlin An, Shirui Cao, Xiaoran Fan, and Longfei Shangguan. 2025. LeakyFeeder: In-Air Gesture Control Through Leaky Acoustic Waves. In *Proceedings of the 23rd ACM Conference on Embedded Networked Sensor Systems*. 144–157.

[117] Yongjie Yang, Tao Chen, Yujing Huang, Xiuzhen Guo, and Longfei Shangguan. 2024. MAF: Exploring mobile acoustic field for hand-to-face gesture interactions. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 1–20.

[118] Yuqing Yin, Xu Yang, Jie Xiong, Sunghoon Ivan Lee, Pengpeng Chen, and Qiang Niu. 2021. Ubiquitous smartphone-based respiration sensing with Wi-Fi signal. *IEEE Internet of Things Journal* 9, 2 (2021), 1479–1490.

[119] Kuang Yuan, Dong Li, Hao Zhou, Zhehao Li, Lili Qiu, Swarun Kumar, and Jie Xiong. 2025. WindDancer: Understanding Acoustic Sensing under Ambient Airflow. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 9, 2 (2025), 1–25.

[120] YW Yuan and K Kirk Shung. 1988. Ultrasonic backscatter from flowing whole blood. I: Dependence on shear rate and hematocrit. *The Journal of the Acoustical Society of America* 84, 1 (1988), 52–58.

[121] Sangki Yun, Yi-Chao Chen, Huihuang Zheng, Lili Qiu, and Wenguang Mao. 2017. Strata: Fine-grained acoustic-based device-free tracking. In *Proceedings of the 15th annual international conference on mobile systems, applications, and services*. 15–28.

[122] Fusang Zhang, Zhaoxin Chang, Jie Xiong, Junqi Ma, Jiazhi Ni, Wenbo Zhang, Beihong Jin, and Daqing Zhang. 2023. Embracing consumer-level UWB-equipped devices for fine-grained wireless sensing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 4 (2023), 1–27.

[123] Guoming Zhang, Chen Yan, Xiaoyu Ji, Tianchen Zhang, Taimin Zhang, and Wenyuan Xu. 2017. Dolphinattack: Inaudible voice commands. In *Proceedings of the 2017 ACM SIGSAC conference on computer and communications security*. 103–117.

[124] Qian Zhang, Kaiyi Guo, Yifei Yang, and Dong Wang. 2025. WearSE: Enabling Streaming Speech Enhancement on Eyewear Using Acoustic Sensing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 9, 1 (2025), 1–30.

[125] Qian Zhang, Dong Wang, Run Zhao, Yinggang Yu, and Junjie Shen. 2021. Sensing to hear: Speech enhancement for mobile devices using acoustic signals. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 3 (2021), 1–30.

A Proportional approximation of frequency perturbation

Consider

$$s_{\mathrm{RX}}(t) \;=\; A_S e^{j(\omega t+\theta_S)} + A_D e^{j(\omega t+\theta_0+\omega'_D t)}. \tag{30}$$

Factor out the strong tone:

$$s_{\mathrm{RX}}(t) \;=\; e^{j(\omega t+\theta_S)} A_S \Big[\, 1 + \varepsilon\, e^{j(\Delta\theta_0+\omega'_D t)} \Big], \tag{31}$$

where

$$\varepsilon \;=\; \frac{A_D}{A_S}, \qquad \Delta\theta_0 = \theta_0 - \theta_S, \qquad \varepsilon \ll 1. \tag{32}$$

We now write the bracket in polar form:

$$1 + \varepsilon e^{j\phi(t)} \;=\; A(t)\, e^{j\varphi(t)}, \qquad \phi(t) = \Delta\theta_0 + \omega'_D t. \tag{33}$$

Magnitude:

$$\begin{aligned} A(t) &= \left|1 + \varepsilon e^{j\phi}\right| \\ &= \sqrt{(1+\varepsilon\cos\phi)^2 + (\varepsilon\sin\phi)^2} \\ &= \sqrt{1 + 2\varepsilon\cos\phi + \varepsilon^2}. \end{aligned} \tag{34}$$

For small ε, expand the square root to its first-order Taylor expansion:

$$A(t) \;\approx\; 1 + \varepsilon\cos\phi(t). \tag{35}$$

Phase:

$$\varphi(t) = \arg\!\left(1 + \varepsilon e^{j\phi}\right) = \arctan\!\left(\frac{\varepsilon\sin\phi}{1+\varepsilon\cos\phi}\right). \tag{36}$$

Using $\arctan(x) \approx x$ for small x, and neglecting higher–order terms in ε,

$$\varphi(t) \;\approx\; \varepsilon\sin\phi(t). \tag{37}$$

Resulting signal:

$$s_{\mathrm{RX}}(t) \;\approx\; A_S\big[1+\varepsilon\cos(\Delta\theta_0+\omega'_D t)\big]\, e^{\,j\left(\omega t+\theta_S+\varepsilon\sin(\Delta\theta_0+\omega'_D t)\right)}. \tag{38}$$

Instantaneous phase and frequency: The instantaneous phase is

$$\psi(t) \;=\; \omega t + \theta_S + \varepsilon\sin(\Delta\theta_0 + \omega'_D t), \tag{39}$$

so the instantaneous frequency is

$$\omega_{\mathrm{inst}}(t) = \frac{d\psi}{dt} = \omega + \varepsilon\,\omega'_D\cos(\Delta\theta_0 + \omega'_D t). \tag{40}$$

Hence, the frequency deviation relative to the transmitted frequency is

$$\Delta\omega(t) \;=\; \omega_{\mathrm{inst}}(t) - \omega \;\approx\; \frac{A_D}{A_S}\,\omega'_D\cos\!\big(\Delta\theta_0 + \omega'_D t\big). \tag{41}$$

Conclusion: For $A_S \gg A_D$, the resulting signal contains Doppler information $\Delta\omega$ that is approximately proportional to ω'_D, i.e.

$$\Delta\omega \;\propto\; \omega'_D. \tag{42}$$

B Single-bin DFT

Let the DFT input signal be the signal received in Equation (15), with a length-N discrete form:

$$x[n] = A e^{j\left(2\pi\frac{f}{F_s}n+\theta_S\right)}, \quad f = f_0 + \Delta f, \quad 0 \le n < N, \tag{43}$$

where f is the input frequency and F_s is the sampling rate. Let i be the index of the nearest-frequency bin of f:

$$i = \mathrm{round}\left(\frac{fN}{F_s}\right). \tag{44}$$

Then, the i-th bin of the N-point DFT is:

$$X[i] = \sum_{n=0}^{N-1} x[n] e^{-j2\pi\frac{i}{N}n} = A e^{j\theta_S}\sum_{n=0}^{N-1} e^{j2\pi\left(\frac{f}{F_s}-\frac{i}{N}\right)n}. \tag{45}$$

Define the fractional bin offset d as follows:

$$d = \frac{f - i f_{\mathrm{bin}}}{f_{\mathrm{bin}}} = \frac{fN}{F_s} - i, \quad \text{where } f_{\mathrm{bin}} = \frac{F_s}{N}. \tag{46}$$

Substitute d into Equation (45), we have:

$$\frac{X[i]}{A e^{j\theta_S}} = \sum_{n=0}^{N-1} e^{j2\pi\frac{d}{N}n}. \tag{47}$$

Note that this is a Dirichlet kernel. Summing it as a geometric series and applying Euler's formula give us:

$$X[i] = \frac{\sin(\pi d)}{\sin\left(\pi\frac{d}{N}\right)} e^{j\left(\pi d\frac{N-1}{N}+\theta_S\right)}. \tag{48}$$

Therefore, the phase of the i-th bin DFT when the input frequency is slightly off from the bin-centered frequency is:

$$\phi = \mathrm{Arg}(X[i]) = \pi d\frac{N-1}{N} + \theta_S. \tag{49}$$

C Artifact

C.1 Abstract

This artifact contains MATLAB demo scripts accompanying the ACM MobiSys 2026 paper *[Emerging Ideas] Phonotonos: Through-Skin Ultrasonic Blood Flow Sensing Using Smartphones.* The artifact provides an end-to-end signal processing pipeline for ultrasonic sensing of blood flow using audio recordings.

The artifact enables reproduction of key processing stages described in the paper, including non-linearity cancellation, baseline estimation and removal, and extraction of normalized blood flow velocity waveforms and Doppler indices. A sample dataset is provided to demonstrate the pipeline. Due to IRB restrictions, only limited data is available. The artifact focuses on functional reproducibility of the pipeline and qualitative correspondence to figures reported in the paper.

C.2 Description

C.2.1 How to access. The artifact is publicly available on GitHub:

- https://github.com/Phonotonos/Phonotonos

C.2.2 Hardware dependencies. No special hardware is required. The artifact can be executed on a standard desktop or laptop.

C.2.3 Software dependencies.

- MATLAB R2025a or later
- MATLAB DSP System Toolbox
- MATLAB Signal Processing Toolbox

C.2.4 Data sets. Due to IRB restrictions, data collected from human subjects cannot be publicly shared. In accordance with the MobiSys'26 artifact evaluation guidelines, we provide a demonstration video illustrating the data collection procedure:

- https://phonotonos.github.io/

To enable functional evaluation of the pipeline, we include a sample dataset:

- `DataExample/sample.wav` (recorded by one of the authors)

C.3 Installation

(1) Install MATLAB R2025b or later.
(2) Ensure that the DSP Toolbox is installed.
(3) Ensure that the DSP System Toolbox and Signal Processing Toolbox are installed.
(4) Clone the repository:

```
git clone https://github.com/Phonotonos/Phonotonos.git
```

C.4 Experiment workflow

To run the end-to-end demo:

(1) Open MATLAB.
(2) Navigate to the repository directory.
(3) Run the main script:

```
EndToEndDemo/main.m
```

(4) Navigate to the repository `Phonotonos/EndToEndDemo`
(5) Run the main script: `main.m`

The script performs the full processing pipeline, including:

- Non-linearity cancellation
- Baseline estimation and removal
- Extraction of normalized blood flow velocity waveforms
- Computation of Doppler indices

C.5 Evaluation and expected results

Running the demo produces intermediate and final outputs corresponding to key figures in the paper:

- Non-linearity cancellation results (corresponding to Fig. 7)
- Baseline estimation and removal (corresponding to Fig. 10)
- Final normalized velocity waveforms and Doppler indices (corresponding to Fig. 1)

The outputs are generated as plots and processed signals in MATLAB. The results should qualitatively match the trends and behaviors described in the paper.

C.6 Experiment customization

Users can modify key parameters at the beginning of `main.m`, including:

- `fs`: sampling rate (F_S)
- `F_T`: transmitted frequency (f_0)
- `FFT_SIZE`: single-bin DFT size (N)

By default, the pipeline assumes 48 kHz input audio and upsamples it to 192 kHz for processing. If 192 kHz input is used directly, the upsampling stage should be disabled.

C.7 Notes

- The provided dataset is limited due to IRB constraints.
- The artifact focuses on demonstrating the signal processing pipeline rather than reproducing full-scale evaluation.
- For updates and issue tracking, refer to the official GitHub repository.