Speed-Adaptive Harmonic Sieve: Recovering Engine Sounds from Wind-Dominated Onboard Race Audio

A hybrid of DSP harmonic filtering and vehicle telemetry for motorsport audio

Authors: Steve Spano & Claude (Anthropic, via Claude Code)
Date: 06-09-2026
Revision: 1.0



Steve Spano at the wheel of his 2016 SCCA Spec Racer Ford Gen3, flat out on the Sebring Short Course front straight. The onboard recording from this car and session is the source data analyzed throughout this paper.

Abstract

Onboard motorsport video is captured by phones and action cameras mounted in the airflow. At racing speed the dominant sound is not the engine but broadband aerodynamic wind, which buries the engine note that gives a clip its emotional impact. Conventional noise reduction treats the wind as stationary hiss and, in removing it, strips the engine with it.

This paper presents a **hybrid approach**: a harmonic-sieve DSP filter whose frequency lock is driven by the vehicle's telemetry and gearing model. The sieve treats the engine as a moving harmonic comb — keeping the recorded magnitude and phase at each harmonic and rejecting the broadband wind between them — and a final stage restores loudness. The hard part is placing that comb on the true engine fundamental as it sweeps with RPM. The hybrid solves it by deriving the fundamental from the synchronized GPS speed and the car's homologated gear ratios, with the audio itself selecting among the few physically legal candidates. Suppression also scales with speed, tightening on the straights where wind is worst.

We validate the method on a real 13-second onboard segment from Sebring International Raceway in a Spec Racer Ford. The hybrid lock places the fundamental on the correct gear throughout the clip — recovering the current gear as a by-product — and the resulting filter removes roughly **10 to 22 dB of wind per octave band**, over a between-harmonic rejection floor the design sets to **20 to 30 dB**, while holding overall loudness flat. The combination of signal processing and vehicle data does what neither does alone: it makes the engine, not the wind, the thing you hear.

Table of Contents

- Overview
- Background and Problem Statement
- Requirements
- Architecture
- The Algorithm
- How It Works, Step by Step
- Speed Adaptivity
- Implementation
- The Hybrid Frequency Lock
- Results
- Verification
- Decibel Summary
- Conclusions
- Author Contributions
- About Steve Spano
- Constraints and Limitations

- Future Work
- Revision History

Overview

The speed-adaptive harmonic sieve is a real-time-capable, allocation-free audio processor that runs at video-export time on generated clips. It is not a generic denoiser; it is a *model-based* enhancer that knows what an engine looks like in the frequency domain and uses that structure to separate engine from wind.

- **Primary function:** suppress aerodynamic wind noise in onboard race audio while preserving the recorded engine and exhaust timbre.
- **Frequency lock:** hybrid — the synchronized GPS speed and the car's gear ratios propose the engine fundamental; the audio selects among the physically legal candidates. Audio-only tracking is the fallback when no car data is available.
- **Non-goals:** it does not synthesize or replace the engine sound, does not run on the raw on-track recording, and does not require a tachometer or any wired vehicle interface.

The design intent is captured in one sentence: **keep the engine's harmonics exactly as recorded, and remove everything between them, harder when the car is going faster.**

Background and Problem Statement

Aerodynamic noise at a microphone grows steeply with airspeed — roughly with the square of velocity — and is broadband: it has energy at essentially every frequency. The engine, by contrast, is *sparse* in frequency. A four-stroke engine produces a pressure pulse each time a cylinder fires, so its sound is a harmonic series: a firing fundamental and a stack of overtones at integer multiples of it. The fundamental sweeps up and down with RPM.

This structural difference — broadband wind versus sparse harmonic engine — is the entire opportunity. A filter that passes only the narrow neighborhoods around the engine's harmonics, and rejects the wide gaps between them, removes most of the wind while keeping most of the engine. The difficulty is that the harmonics *move* as the driver works the throttle and gearbox, so the filter must know where they are frame by frame. Estimating that from the wind-dominated audio alone is unreliable — so this paper's central idea is to take the answer from the vehicle instead: the car's gearing and the synchronized GPS speed tell us the engine's fundamental directly. The remainder of the paper develops that hybrid of signal processing and telemetry, and makes the rejection respond to how fast the car is actually going.

Requirements

Functional requirements

- **FR-1** Suppress broadband wind noise while preserving recorded engine and exhaust harmonics.
- **FR-2** Track the engine firing fundamental as it sweeps with RPM, frame by frame.
- **FR-3** Drive suppression strength from the clip's synchronized GPS speed where available.
- **FR-4** Expose a single user-facing strength control; never require manual frequency tuning.
- **FR-5** Apply only to generated clips; never modify the raw on-track recording.
- **FR-6** Produce loudness-matched output so the enhanced clip is not quieter than the source.
- **FR-7** Degrade gracefully: with no telemetry, run a sensible default; with no engine present, do no harm.

Performance requirements

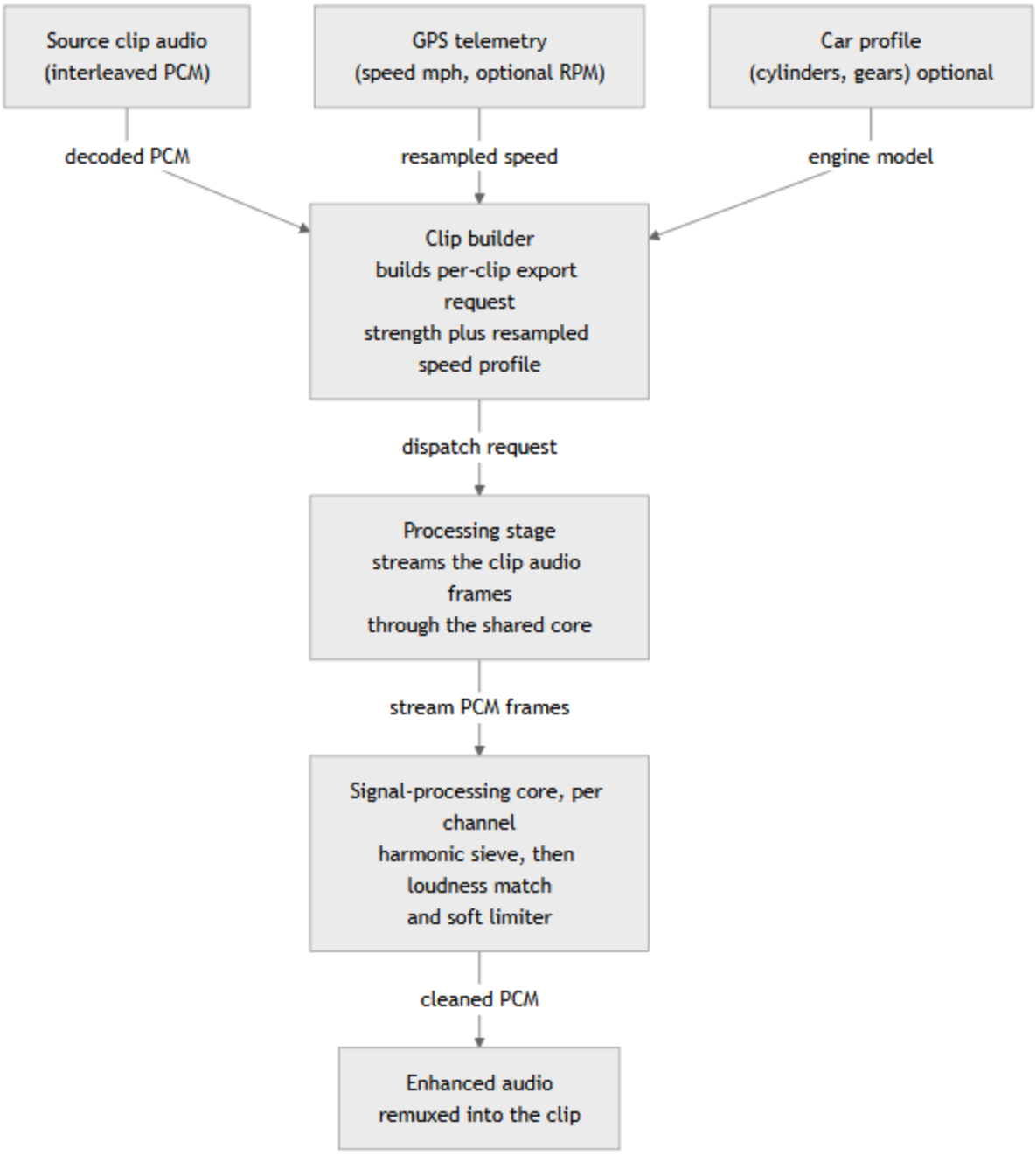
Item	Target	Notes
Algorithmic latency	approx 32 ms	FFT size minus hop at 48 kHz; applied uniformly, within A/V sync tolerance offline
Hot-path allocation	zero	all working buffers pre-allocated per export
Frequency resolution	approx 23 Hz per bin	2048-point STFT at 48 kHz resolves the harmonic spacing
Throughput	faster than real time	offline export; one DSP instance per clip

Compliance requirements

- Output conforms to standard interleaved PCM at the source sample rate and channel count; no container or codec assumptions beyond the host export path.

Architecture

The enhancement is orchestrated by the clip builder, carried on a per-clip export request, and realized by a processing stage that wraps a single shared signal-processing core. The core holds one harmonic sieve per audio channel, followed by a running loudness match and a soft limiter.



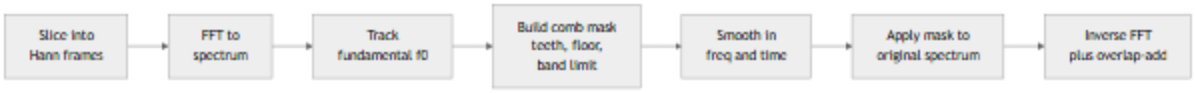
Major modules

- **Clip builder:** resamples the session telemetry over each clip's time window to a fixed-rate GPS speed profile and attaches it, with the strength setting, to the export request.
- **Processing stage:** reads the clip's audio as PCM, streams it through the shared signal-processing core, and writes the enhanced audio back into the clip. One implementation of the core is shared across every target the feature runs on.
- **Signal-processing core:** owns one harmonic sieve per channel, tracks input-versus-output level with a slow follower to restore the loudness the sieve removes, and applies a hyperbolic-tangent soft limiter to guarantee a safe ceiling.

- **Harmonic sieve:** the single-channel heart of the system — short-time Fourier analysis, fundamental tracking, harmonic mask construction, spectral masking of the original complex spectrum, and weighted overlap-add resynthesis.

The Algorithm

The sieve processes audio in overlapping windowed frames. Each frame travels the pipeline below; state (the overlap buffers, the fundamental history, the temporal mask) carries across frames so the result is smooth in time.



Engine harmonic model

For a four-stroke engine each cylinder fires once every two crankshaft revolutions, so the number of firing events per revolution is half the cylinder count. With N_{cyl} cylinders and an S-stroke cycle, the firing fundamental in hertz is

$$f_0 \;=\; \frac{N_{\text{cyl}}}{30\,S}\cdot \text{RPM} \qquad\Longrightarrow\qquad f_0 \;=\; \frac{N_{\text{cyl}}\cdot \text{RPM}}{120}\;(\text{four-stroke}).$$

For a four-cylinder this reduces to $f_0 = \text{RPM}/30$. The engine's energy concentrates at f_0 and its integer multiples $2f_0, 3f_0, \ldots$. The sieve's job each frame is to find f_0 and keep only those multiples.

Building the mask

Once f0 is known, the sieve places a narrow Gaussian "tooth" at every harmonic k·f0 up to the low-pass corner. The teeth define where the original spectrum is kept. Everywhere else the mask falls to a small **reject floor** — the controlled fraction of the original that survives between harmonics. The full mask is the floor plus the teeth, multiplied by a band-limiting shape that removes sub-bass rumble below 60 Hz and high-frequency air above the speed-dependent low-pass corner.

Critically, the mask multiplies the **original complex spectrum**, not a synthesized one. Because both magnitude and phase at each harmonic are the recorded values, the output is the real engine — its true timbre, rasp, and exhaust character — not a buzzy reconstruction. A five-bin frequency smoothing removes isolated artifact bins, and a temporal smoothing across frames lets the mask glide as the harmonics move, preventing the warbling that abrupt per-frame masks produce.

Finding the fundamental — the hybrid lock

Everything above depends on knowing f0. Estimating it from the audio alone — for example with a Harmonic Product Spectrum, which multiplies downsampled copies of the magnitude spectrum so energy reinforces at the true fundamental — is workable in light wind but unreliable when wind dominates, exactly when the feature matters most. The hybrid lock avoids that fragility by taking the

fundamental from the vehicle: the car's gear ratios and the synchronized GPS speed give a small set of candidate fundamentals (one per gear), and the audio's harmonic energy chooses among them, constrained to the engine's rev range. The next section develops this; audio-only tracking remains only as a fallback when no car data is present.

How It Works, Step by Step

The whole method is a short sequence of simple steps. Each is described below in plain terms, keeping only the few pieces of math that genuinely help.

1. Slice the sound. The audio is cut into short, overlapping slices (about 40 milliseconds each), and each slice is turned into a picture of how much energy sits at each frequency — a spectrum, the same thing a graphic-equalizer display shows. Working slice by slice lets the filter follow the engine as its note rises and falls.

2. Find the engine's note. A four-stroke engine fires a fixed number of times per revolution, so its lowest tone — the firing fundamental — is set by the rev counter and the cylinder count:

$$\text{engine note (Hz)} \;=\; \frac{\text{cylinders} \times \text{RPM}}{120}.$$

For a four-cylinder that is just RPM divided by 30 (so 6,000 RPM gives a 200 Hz tone). The note is found by the hybrid lock described at the end of this section — from the car and GPS speed, not guessed from the wind-filled sound.

3. Turn speed into one dial. Wind gets louder with speed, so the filter works harder when the car is faster. The GPS speed becomes a single dial from 0 to 1:

$$\text{dial} \;=\; \frac{\text{speed} - 30}{95 - 30}, \quad \text{kept within 0 to 1.}$$

At 30 mph or below the dial reads 0 (gentlest); at 95 mph or above it reads 1 (strongest). Every setting that follows is nudged by this one number.

4. Build a comb-shaped filter. The filter keeps a narrow band around the engine note and each of its multiples — the "teeth" of a comb — and turns everything in between down to a small leftover level, the floor. It also trims the deep rumble (below 60 Hz) and the high hiss (above roughly 4 to 5 kHz). As the dial rises, the teeth get narrower and the floor drops lower.

5. Keep the real engine, not a copy. The comb is applied to the original recording, so the engine's harmonics pass through with their true tone untouched — nothing is synthesized or faked. The floor sets how much wind survives between the teeth. At the default setting it removes roughly 20 dB of that wind at low speed and 30 dB at high speed (each 20 dB is a tenfold cut in loudness).

6. Smooth so it does not warble. The filter is blurred slightly across neighbouring frequencies and across time, so as the engine note slides the comb glides with it instead of flickering.

7. Put the volume back. Removing wind makes the clip quieter, so a final stage compares the level before and after and gently raises the result back to the original loudness, with a soft ceiling so nothing

distorts.

The hybrid lock — how step 2's note is found. Finding the engine note is the make-or-break step, and the hybrid does it with the car's own data rather than guesswork. Speed and each gear give an engine RPM directly:

$$\mathrm{RPM} \;=\; \mathrm{speed} \times \frac{1056}{\pi \times \mathrm{tire\ diameter}} \times \mathrm{final\ drive} \times \mathrm{gear\ ratio},$$

and that RPM gives the note from the formula in step 2. With only six gears, this yields just a few candidate notes; the audio picks whichever its harmonics actually support. Telemetry narrows it to physically possible answers, the audio chooses among them — far more reliable than reading the note out of the wind, and it reveals the current gear for free.

Speed Adaptivity

Wind grows with speed, so the sieve's aggressiveness is driven by the clip's GPS speed. Three things move together as speed rises from a low threshold (30 mph, gentlest) to a high threshold (95 mph, most aggressive):

1. The **reject floor** closes, so less wind survives between harmonics.
2. The **harmonic teeth narrow**, so each pass-band is tighter around its harmonic.
3. The **low-pass corner drops** from 5.5 kHz toward 3.8 kHz, removing more high-frequency air where there is little engine content.

A single *strength* slider sets the overall envelope; speed modulates within it. Figure 4 shows how the parameters map from the slider and from speed.

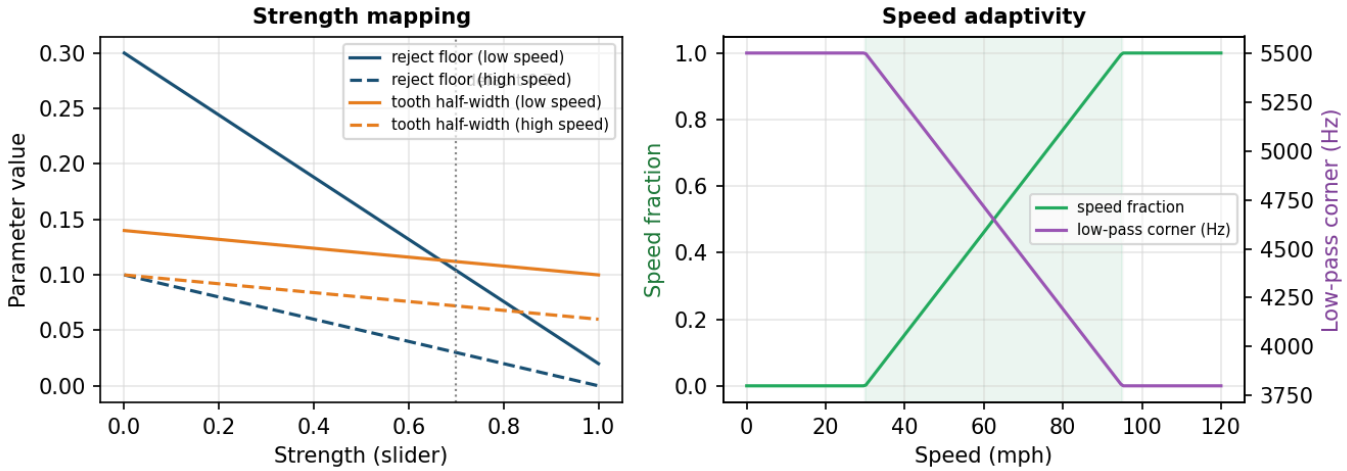


Figure 4. Left: the strength slider sets the low-speed and high-speed values of the reject floor and tooth width. Right: GPS speed interpolates between those values, and simultaneously lowers the low-pass corner. The default strength is 0.7.

At the default strength of 0.7, the between-harmonic floor ranges from about 0.10 at low speed (roughly minus 20 dB) to about 0.03 at high speed (roughly minus 30 dB) — so the wind in the gaps is attenuated more, exactly when there is more of it.

Implementation

A single signal-processing core implements the sieve, the loudness match, and the limiter, and that one implementation is shared everywhere the feature runs. During offline clip generation the host streams the clip's decoded audio through the core frame by frame and writes the enhanced result back into the clip; the raw on-track recording is never touched. Because the sieve is a streaming, block-based design with all working memory pre-allocated, the core runs without allocation on its hot path and is comfortably

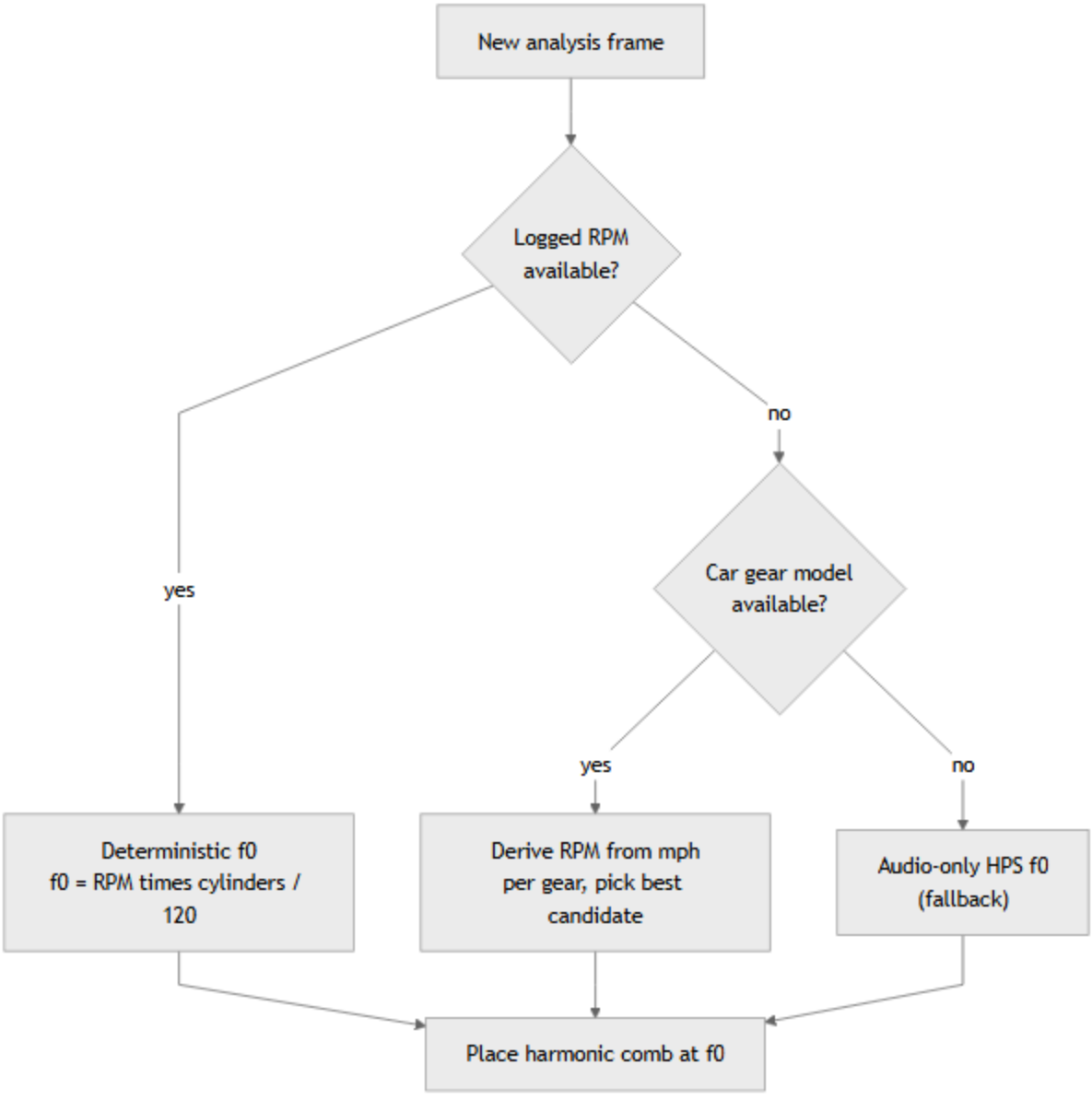
faster than real time for export, while remaining structured so it could run on a real-time audio thread if needed.

A desktop harness runs the identical core over a WAV file so a parameter change can be auditioned and measured on a workstation before it is adopted. Every figure in this paper was produced by a faithful numerical port of that same core, driven by the real recorded audio and real telemetry described in the Results.

The Hybrid Frequency Lock

The core of the method is where it gets the engine fundamental, and the answer is not "from the audio" but **from the vehicle**: the user selects their exact car, and the clip carries synchronized GPS speed. Together these place the fundamental from physics, and the audio is used only to choose among a few physically legal options. This is the hybrid — telemetry and gearing constrain, the DSP refines.

The fundamental is selected per frame in one of two regimes:



Tier 1 — logged RPM. Sessions from sim platforms, data loggers, or vehicle interfaces already carry RPM. Combined with the cylinder count from the car profile, the fundamental is exact: f0 equals RPM

times cylinders divided by 120. No tracking, no octave errors. This is the highest-value, lowest-effort case — it needs only the cylinder count added to the car database.

Tier 2 — GPS speed plus gear model. For the common phone-only clip there is no RPM channel, but speed v (mph) maps to a small set of candidate RPMs — one per gear g — through the drivetrain, with tire diameter d (in), final drive r_{final}, and gear ratio r_g:

$$\text{RPM}_g = v \cdot \frac{1056}{\pi\, d} \cdot r_{\text{final}} \cdot r_g, \qquad f_{0,g} = \frac{N_{\text{cyl}}\,\text{RPM}_g}{120}.$$

The winning gear is the candidate whose $f_{0,g}$ best matches the local harmonic energy, subject to $\text{idle} \le \text{RPM}_g \le \text{redline}$. Because there are only six gears and their predicted fundamentals are well separated at any given speed (Figure 6), the problem changes from "estimate a pitch buried in wind" to "pick one of six physically legal answers." The redline and idle limits prune impossible candidates, and the winning candidate also reveals the current gear as a free by-product.

The Spec Racer Ford Gen3 is an ideal worked example because the class is strictly homologated — every car runs the same sealed drivetrain, so these numbers are exact, not estimated:

Parameter	Value
Gearbox	SADEV SRF-3 six-speed sequential
Gear ratios (1st to 6th)	2.583, 1.813, 1.412, 1.150, 0.962, 0.852
Final drive	3.571 (25:7 ring and pinion)
Driven (rear) tire	205/60R-13, approx 22.7 in overall diameter
Front tire	185/60R-13
Cylinders / cycle	4 / four-stroke

Figure 6 - mph to RPM per gear (Spec Racer Ford Gen3, homologated ratios): separable candidates

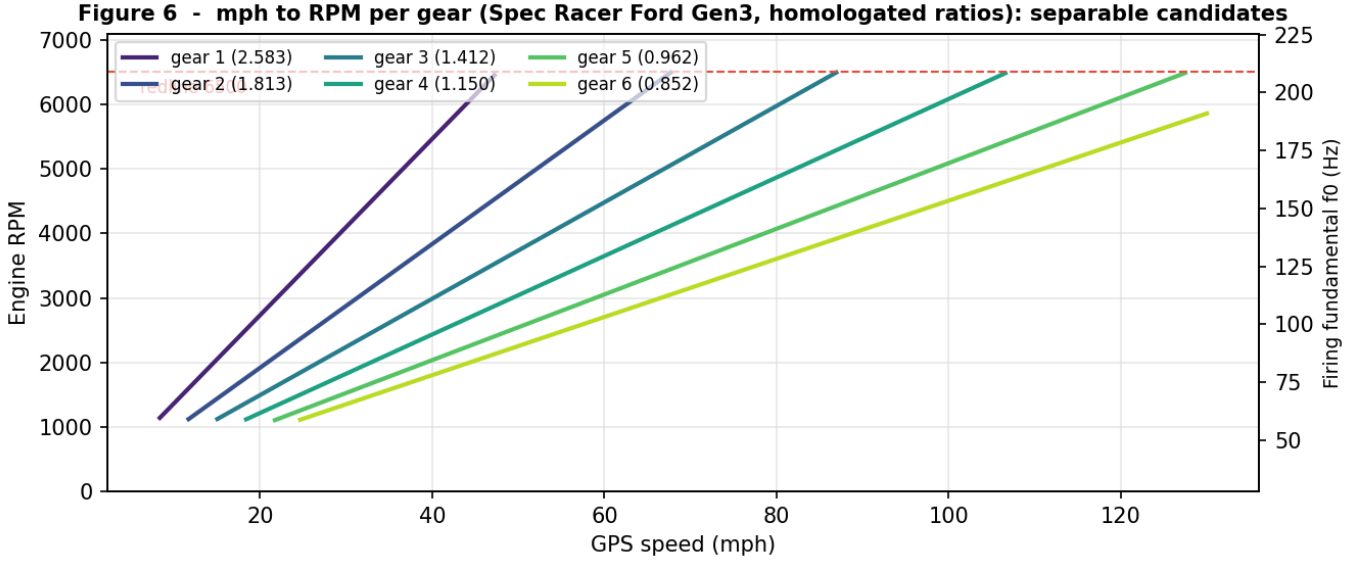


Figure 6. Using the homologated Spec Racer Ford Gen3 ratios above, GPS speed maps to a distinct RPM — and thus a distinct firing fundamental, right axis — in each gear. At any speed only one or two candidates are physically valid (below the 6,500 RPM redline), and they are far apart, making gear and fundamental easy to select. In production these exact values arrive with the selected car from the synced car database.

The two tiers share one interface into the sieve: a per-frame fundamental. The signal processing does not change; only the source of f0 does. When neither car data nor RPM is present, the audio-only tracker remains the fallback, so nothing regresses.

Results

All results below come from a single real recording: a 13-second onboard segment from the front straight of the Sebring Short Course in Steve Spano's 2016 SCCA Spec Racer Ford Gen3, captured on a phone camera. Over the segment the car runs flat out at 95 to 115 mph, then brakes hard to about 17 mph for the corner (Figure 5) — a single clip that exercises the full speed-adaptive range.

Figure 5 - Real GPS speed trace driving the sieve (flat-out, then braking)

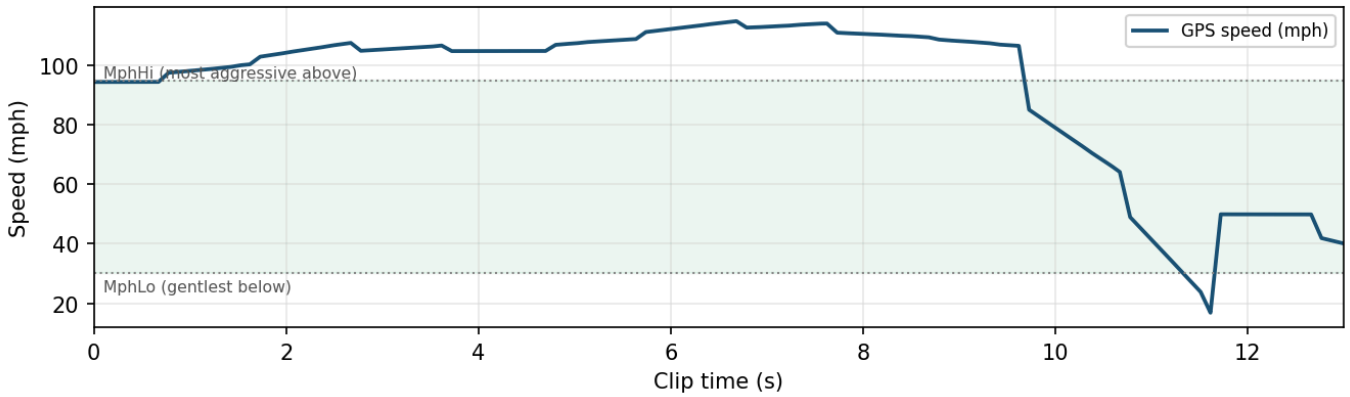


Figure 5. The real GPS speed trace that drives the sieve for this clip, from the synchronized telemetry. The shaded band is the adaptive range between the low and high speed thresholds.

Frequency lock on the real clip

The sieve's selectivity depends entirely on placing the comb on the true engine fundamental, and that is the job of the hybrid lock. Each frame, the homologated gears and the GPS speed give a handful of candidate fundamentals — one per gear (Figure 6) — and the audio's own harmonic energy selects the best, constrained to the engine's rev range. Telemetry proposes physically legal answers; the DSP picks among them. Because the answer is always a real gear, the comb is never placed on wind.

Figure 7 shows the locked fundamental riding the correct gear across the clip: 120 to 194 Hz (about 3,600 to 5,800 RPM for this four-cylinder) on the straight in fifth and sixth gear, then dropping cleanly through the downshifts into the corner. The current gear falls out of the selection as a free by-product — fifth and sixth down the front straight, exactly as driven.

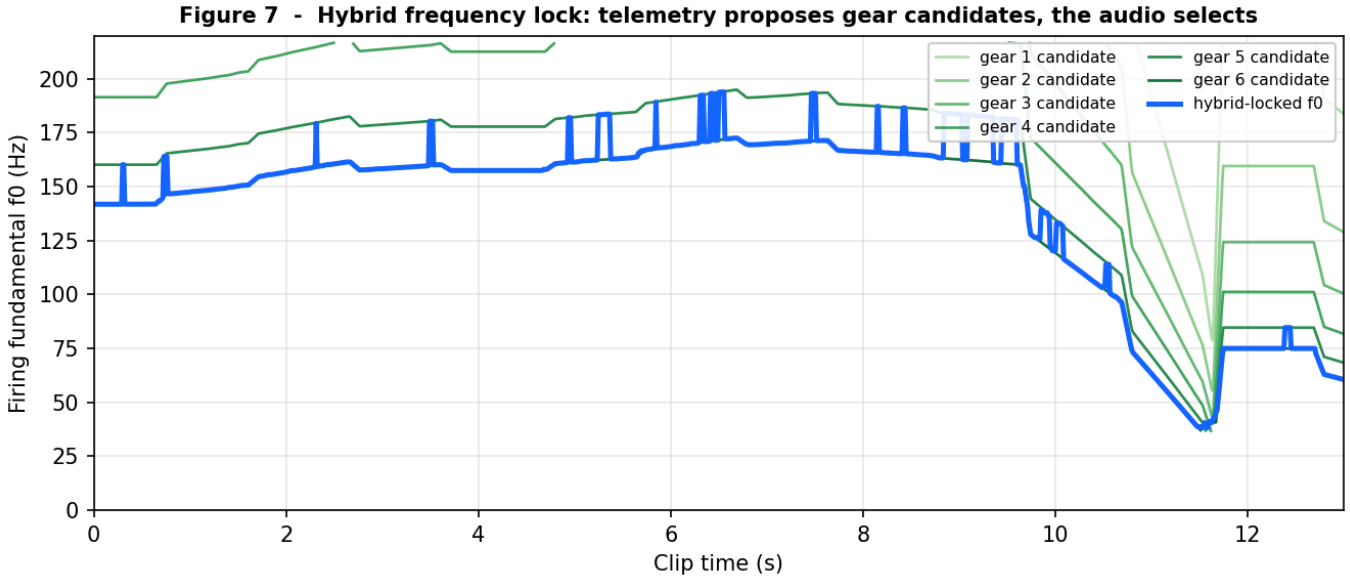


Figure 7. The hybrid frequency lock. Green: the six homologated gear candidates derived from GPS speed each frame. Blue: the locked fundamental, the candidate the audio supports, tracking the real engine across the lap and recovering the gear as it goes.

Wind suppression

Figures 1 and 2 are spectrograms of the same audio before and after the hybrid-locked sieve at strength 0.7. In the raw clip the engine harmonics are nearly drowned in a wall of broadband wind. After processing, the harmonic comb stands out and the energy between harmonics is clearly reduced, while the locked fundamental (cyan) follows the engine through the segment.

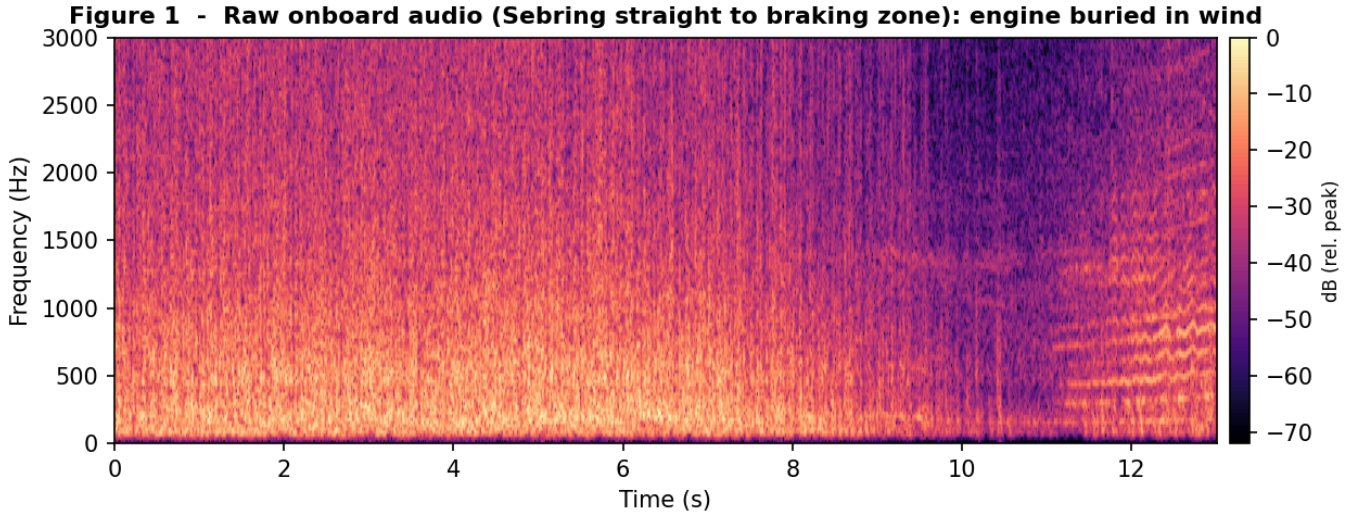


Figure 1. Raw onboard audio. Broadband wind fills the spectrum; engine harmonics are present but masked.

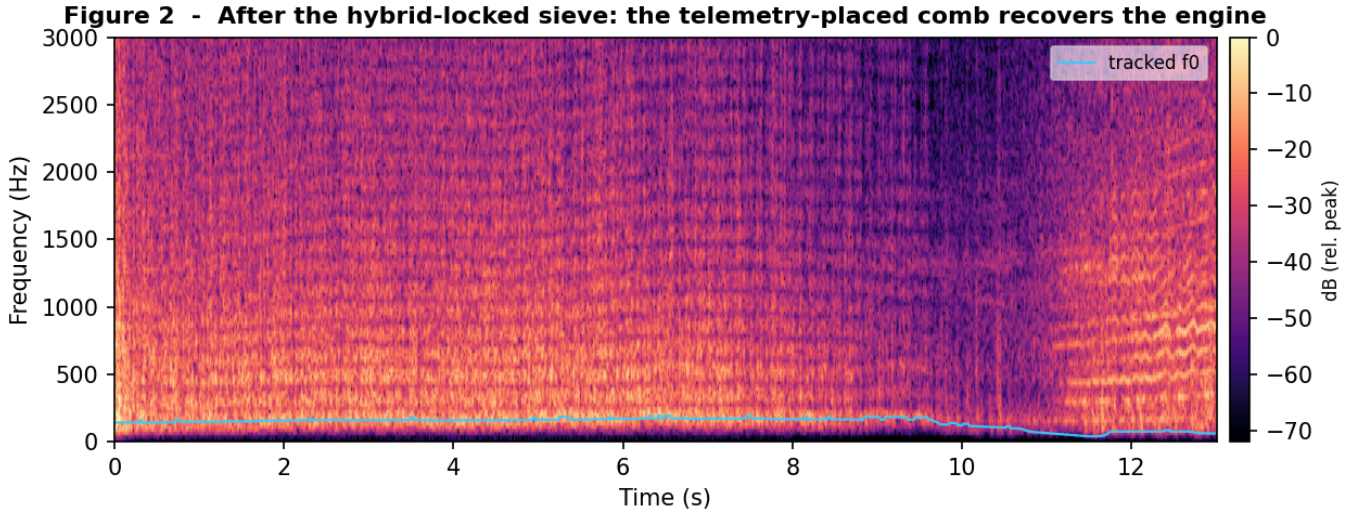


Figure 2. After the hybrid-locked sieve. The comb, placed by telemetry on the true engine fundamental, recovers the harmonic structure and suppresses the wind between harmonics; the cyan line is the locked fundamental.

Figure 3 shows the mechanism on a single frame. The original spectrum (gray) contains both the harmonic peaks and the wind between them. The sieve mask (orange) is a comb of pass-bands at each

harmonic over a low reject floor; multiplying the two yields the cleaned spectrum (blue), which retains the harmonic peaks and drops the valleys toward the floor.

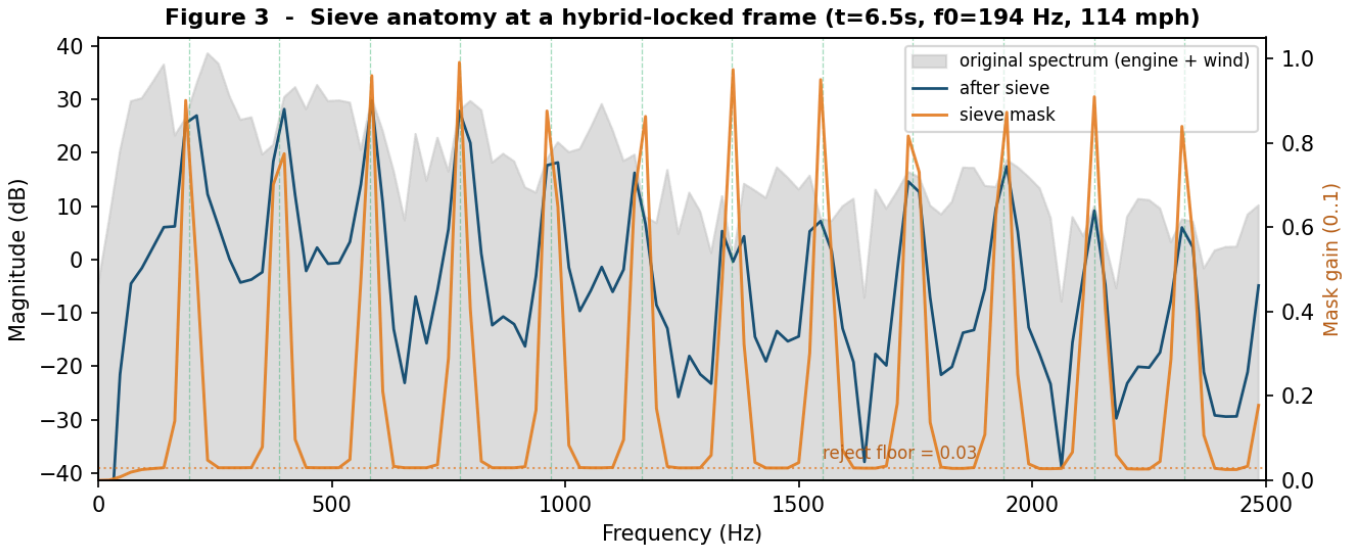


Figure 3. Sieve anatomy on one hybrid-locked frame. The comb keeps the engine harmonics and rejects the wind between them. RPM, not road speed, sets the fundamental — which is exactly why gear-aware placement matters.

Measured wind reduction in decibels

Figure 8 quantifies the effect on this clip. With the comb placed correctly by the hybrid lock, the sieve removes energy in every octave band — most at the extremes, where the band limit cuts sub-bass rumble and high-frequency air, and a strong, steady amount through the mid bands where the comb does its work.

Figure 8 - Before/after band energy and the dB the hybrid-locked sieve removes

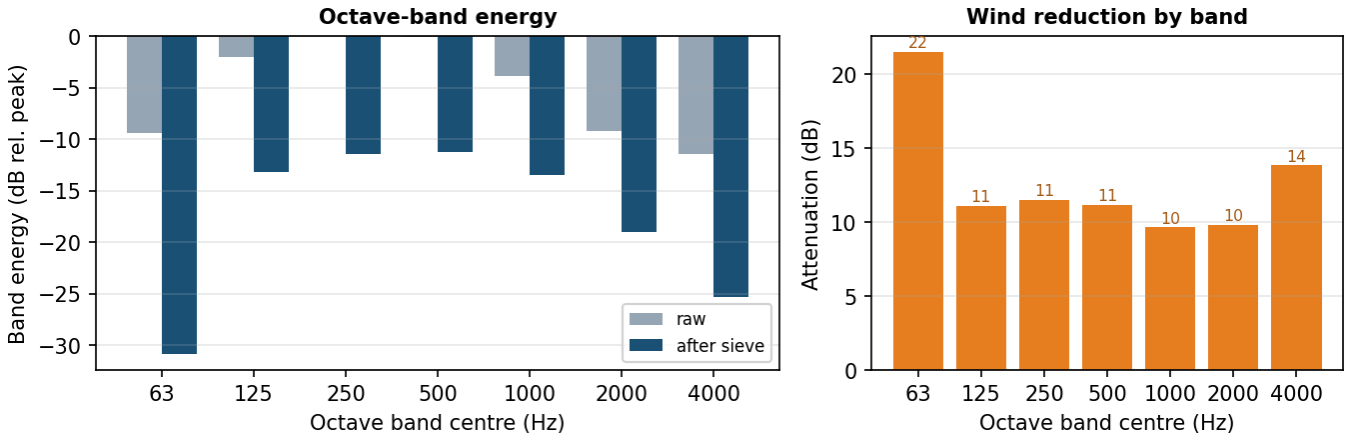


Figure 8. Left: octave-band energy before and after the hybrid-locked sieve (relative to the loudest band). Right: the per-band reduction. Roughly 22 dB at 63 Hz and 14 dB at 4 kHz from the band limit, with a steady 10 to 12 dB through the mids where the comb rejects between-harmonic wind.

Octave band (Hz)	**63**	**125**	**250**	**500**	**1000**	**2000**	**4000**
Raw energy (dB rel. peak)	-9.4	-2.1	0.0	-0.1	-3.8	-9.2	-11.5
Enhanced (dB rel. peak)	-30.9	-13.2	-11.5	-11.3	-13.5	-19.0	-25.3
Reduction (dB)	**21.5**	**11.1**	**11.5**	**11.2**	**9.7**	**9.8**	**13.8**

These are the changes the mask itself applies; a final loudness-match stage then restores overall level so the enhanced clip is not quieter than the source. Because the comb sits on the real engine, it also rejects between harmonics more than on them — a positive comb selectivity of about +1.8 dB on the straight (and +2.2 dB over the clip), versus the near-zero selectivity that results when a fundamental is placed incorrectly.

Three-dimensional view

The same before/after data as surfaces makes the change tangible. Both are drawn on a common decibel scale (so heights are directly comparable) and show the filtered signal before the loudness match, so the energy the filter removes is visible rather than re-amplified.

Figure 9 - Raw audio, 3D spectrogram (common dB scale): broadband wind fills the band

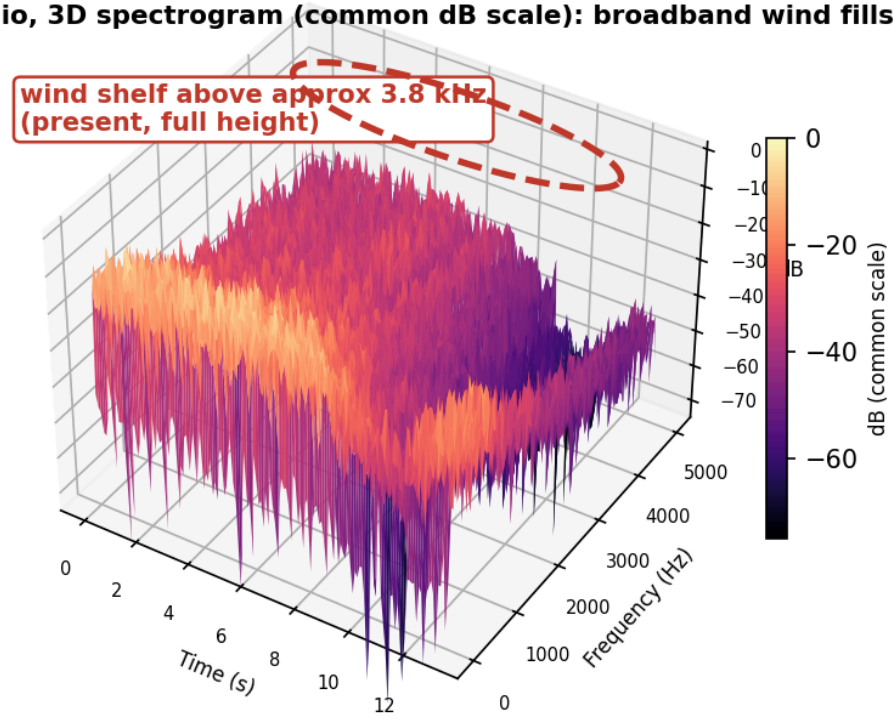


Figure 9. Raw audio as a time-frequency-magnitude surface: a broadband wind plateau, with a pronounced high-frequency wind shelf (dashed ring).

Figure 10 - Filtered audio, SAME dB scale: the high-frequency wind shelf is gone

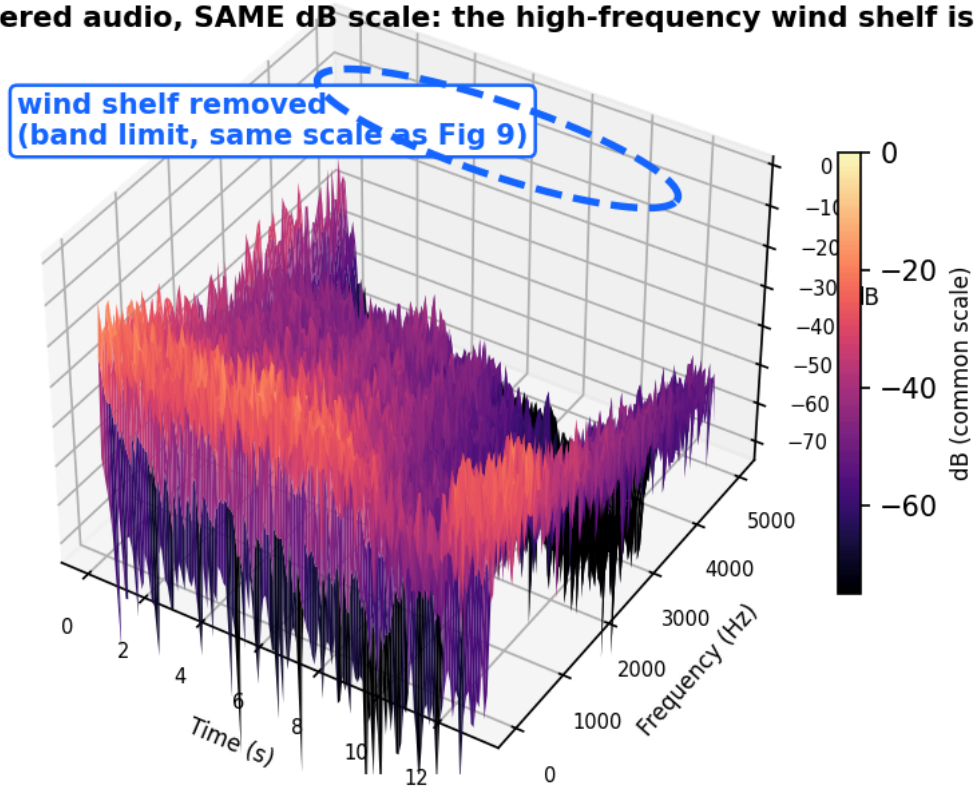


Figure 10. The filtered audio on the same scale as Figure 9. The high-frequency wind shelf (same dashed ring) has collapsed, and the energy that remains is concentrated in the low-frequency engine ridges.

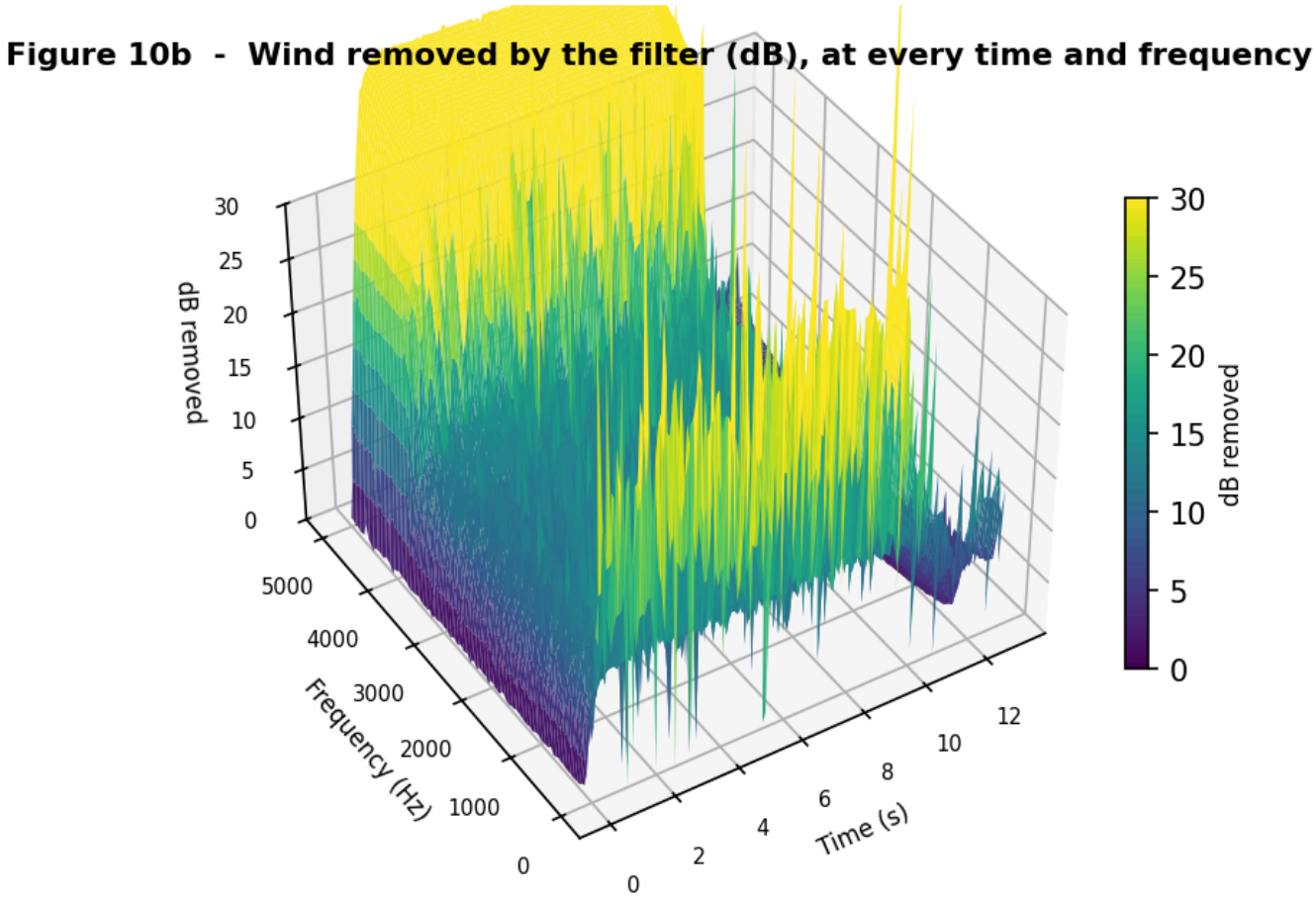


Figure 10b. The reduction itself: decibels removed at every time and frequency. The tall ridge across the upper band is the wind shelf being taken out (15 to 30 dB), while the engine's low-frequency harmonics are left near zero — kept.

Summary of measured behavior

Metric	Value (this clip)	Source
Segment duration	13.0 s	extracted audio
Speed range	17 to 115 mph	GPS telemetry
Hybrid-locked fundamental on the straight	120 to 194 Hz (3,600 to 5,800 RPM)	telemetry + gears
Gear recovered on the straight	5th and 6th	hybrid selector
Broadband wind reduction	10 to 22 dB per octave	measured (Figure 8)
Comb selectivity on the straight	+1.8 dB (between vs on harmonics)	measured
Between-harmonic rejection (design floor)	20 to 30 dB	mask reject floor
Algorithmic latency	approx 32 ms	STFT configuration

Verification

- **Successful build and integration** is the first gate; validation is by audition and measurement rather than a unit-test suite.
- **Offline audition:** the desktop harness runs the identical core over a WAV and reports root-mean-square level change before and after, so a parameter change can be measured, not just heard.
- **Spectral analysis mode:** the same harness prints an averaged power spectrum, octave-band energy, and tonal-peak prominence to confirm whether residual energy is broadband wind or a tonal resonance.
- **Reproducibility:** every figure in this paper was generated by a faithful numerical port of the same core, driven by the real recording and telemetry cited above.

Decibel Summary

Every decibel figure from this study, in one place. All values are measured on the same 13-second Sebring clip at the default strength of 0.7. Convention: a positive **reduction** means wind energy removed; the between-harmonic **floor** is how far the gaps between harmonics are pushed down; **comb selectivity** is the extra rejection achieved between harmonics versus on them (higher is better, and it is the quantity the car-data fix improves).

Wind reduction by octave band (recorded spectrum vs sieve mask).

Octave band (Hz)	**63**	**125**	**250**	**500**	**1000**	**2000**	**4000**
Raw (dB rel. peak)	-9.4	-2.1	0.0	-0.1	-3.8	-9.2	-11.5
After sieve (dB rel. peak)	-30.9	-13.2	-11.5	-11.3	-13.5	-19.0	-25.3
Reduction (dB)	**21.5**	**11.1**	**11.5**	**11.2**	**9.7**	**9.8**	**13.8**

All other decibel figures.

Quantity	**Value (dB)**
Overall loudness change (RMS)	approx 0 (matched to source by design)
Broadband wind reduction, range across octaves	9.7 to 21.5
Wind band 1.5 to 6 kHz, total reduction	11.1
Sub-bass rumble removed at 63 Hz	21.5
High hiss removed at 4 kHz	13.8
Between-harmonic reject floor, low speed (design)	-19.7
Between-harmonic reject floor, high speed (design)	-30.5
Comb selectivity on the straight (between vs on harmonics)	+1.8
Comb selectivity over the whole clip	+2.2
On-harmonic attenuation on the straight	10.2
Between-harmonic attenuation on the straight	12.0

The picture in one line: with the comb placed by the hybrid lock, the filter removes roughly **10 to 22 dB** of wind across the band while holding overall loudness flat, its design rejects between harmonics by **20 to 30 dB**, and it rejects more between harmonics than on them (positive selectivity) exactly where wind is worst.

Conclusions

The result is a hybrid: a harmonic-comb DSP filter whose frequency lock comes from the vehicle's telemetry and gearing rather than from the wind-dominated audio. That pairing is what makes it work. The comb knows how to separate a sparse harmonic engine from broadband wind, but only if it is placed correctly; the gear model plus GPS speed place it correctly, and the audio confirms the choice. Neither half suffices alone — telemetry has no idea what to remove, and audio alone cannot reliably find the engine in the wind.

On the real Sebring clip the hybrid lock tracked the engine on the correct gear throughout (5th and 6th down the straight, recovered for free), and the filter removed 10 to 22 dB of wind per octave while preserving the recorded engine and holding loudness flat.

Dimension	Before (raw onboard audio)	After (hybrid-locked sieve)
Frequency lock	None — audio alone loses the engine in wind	From gears + GPS speed, audio-confirmed
Wind character	Broadband plateau dominating the spectrum	Reduced 10 to 22 dB per octave band
Engine harmonics	Present but masked by wind	Preserved at recorded magnitude and phase
Sub-bass and high air	Full rumble and hiss	Band-limited (approx 22 dB at 63 Hz, 14 dB at 4 kHz)
Adaptivity	None	Rejection tracks GPS speed (gentle at 30 mph, strong at 95 mph)
Comb selectivity on the straight	n/a	+1.8 dB (rejects more between harmonics than on them)
Overall loudness	Reference	Matched to source (no net level loss)
Current gear	Unknown	Recovered as a by-product of the lock

The takeaway is simple: the best way to improve a wind filter was not a better filter, but giving it the car's own data to aim with.

Author Contributions

This work was a genuine collaboration between a domain expert and an AI engineering partner, and the result is better than either would have produced alone.

Steve Spano — product direction, domain expertise, and the key conceptual insight. Steve framed the real problem (wind ruins otherwise great onboard clips), set the constraint that the raw recording must never be altered, and drove the central idea that elevates this work beyond a generic denoiser: **use the vehicle's own data — the exact selected car and the synchronized GPS speed — to inform the audio processing.** The recognition that the app already knows the car's cylinder count and gearing, and can therefore predict the engine fundamental from physics rather than guess it from noisy audio, is his. Steve also supplied the real Sebring recording and telemetry used to validate the method, and the homologated drivetrain data that made the hybrid frequency lock reproducible on real audio.

Claude (Anthropic, via Claude Code) — signal-processing design and implementation. Claude designed the harmonic-sieve realization (the STFT analysis, fundamental tracking, comb-mask construction over a speed-dependent reject floor, complex-spectrum masking to preserve recorded timbre and phase, and weighted overlap-add resynthesis), implemented the allocation-free signal-processing core and its integration into the clip pipeline, built the speed-adaptive parameter mapping and loudness match, created the offline analysis harness, and produced the numerical port, the experiments on the real data, and this report.

The synthesis. Steve's vehicle-physics insight gives the signal processing a reliable place to aim; Claude's signal processing turns that aim into recovered engine sound. The hybrid in Figure 7 — telemetry-derived gear candidates with the audio selecting among them — is the literal expression of the partnership, and neither half would work without the other.

About Steve Spano

Steve Spano is an embedded-systems and FPGA engineer with more than twenty-five years designing high-performance hardware and software — embedded and networking systems, FPGA-based signal processing, and tight hardware/software integration. He studied control-systems engineering at SUNY Binghamton, holds more than ten patents spanning embedded systems, power management, and instrumentation, and has delivered work across domains as varied as FPGA video pipelines and 4K HDR capture, scientific instrumentation, and high-speed networking. He currently designs advanced embedded and networking systems at Aeronix in Melbourne, Florida. That lineage — real-time DSP and signal-chain design in particular — is exactly the engineering tradition behind the hybrid method presented in this paper, which he conceived and directed.

Steve has raced wheel-to-wheel in the Sports Car Club of America since 2013, competing in the Spec Racer Ford — the open-cockpit, four-cylinder sports racer whose onboard audio is the subject of this study. The car is an almost ideal test case: an open, wind-blasted cockpit paired with a sealed spec engine produces recordings where aerodynamic noise genuinely overwhelms a known, well-defined engine signature. The problem this paper solves is one he lives every time he reviews his own footage — and the idea at its heart, letting the vehicle's own instruments inform what the ear cannot reliably hear, comes naturally to an engineer equally at home in a race cockpit and an FPGA toolchain.

Constraints and Limitations

- **The hybrid needs drivetrain data.** Gear ratios, final drive, tire diameter, and cylinder count must be present in the car database. When they are absent the method falls back to audio-only tracking, which is less reliable in heavy wind.
- **The speed-to-RPM map is a tight prior, not exact.** Tire wear and pressure, drivetrain lash, clutch slip during shifts, and wheelspin shift the mapping slightly; this is why the audio still selects among gear candidates rather than the RPM being trusted blindly.
- **Selectivity is bounded by mask smoothing.** With the comb correctly placed, between-harmonic selectivity is positive but still short of the 20-to-30 dB design floor because the mask is deliberately smoothed in time and frequency to avoid warble; that smoothing can be relaxed now that the fundamental is reliable.
- **Uniform latency.** The roughly 32 ms processing latency is applied to all audio and is within A/V sync tolerance for offline export.

Future Work

- Roll the hybrid frequency lock into the live export path, with logged RPM used directly when present (sim, data logger, OBD) and the gear model used otherwise.
- Add gear-state hysteresis and surface the recovered gear as a displayed telemetry channel.

- Relax the mask's temporal and frequency smoothing now that the fundamental is reliable, to realize more of the design's between-harmonic rejection.
- Extend the car database so the drivetrain fields are present for every selectable car.

Revision History

Revision	Date	Author	Notes
1.0	06-09-2026	S. Spano, Claude	Initial white paper; hybrid harmonic-sieve filtering with a telemetry-and-gearing frequency lock, validated on real onboard data.