BeBeC-2026-S10

Source Velocity Beamforming via Inter-Frequency Coherence and Frequency-Domain Doppler Estimation

A. Goudarzi[1]
[1]German Aerospace Center (DLR)
Bunsenstraße 10, 37073, Göttingen

Abstract

We present a proof-of-concept *source velocity beamforming* framework that estimates moving-source kinematics from microphone-array cross-spectral statistics. The core operator is *inter-frequency coherence*, which compares spectral components across frequency after compensating deterministic phase drift induced by time scaling. Under a local-stationarity approximation, this yields an efficient frequency-domain surrogate for time-stretch alignment and enables estimation of pairwise relative Doppler factors without repeated time-domain resampling. Per short-time Fourier transform (STFT) frame, we estimate (i) relative delays (TDOA) using GCC-PHAT-style phasor alignment and (ii) pairwise relative Doppler factors with respect to a reference microphone via inter-frequency coherence maximization. These complementary measurements are fused in a nonlinear state-space model with an Unscented Kalman Filter (UKF) and constant-acceleration dynamics, yielding smooth position and velocity estimates, informative acceleration trends, and uncertainty envelopes. In the synthetic proof-of-concept scenario studied here, the approach provides velocity-sensitive tracking and short-horizon forecasting without exhaustive joint position–velocity grid scanning.

1 Introduction

Tracking moving acoustic sources with microphone arrays is traditionally approached either by (i) scanning a spatial grid with a beamformer (and repeating this over time) or (ii) estimating time-difference-of-arrival (TDOA) features and then solving a geometric localization problem [2]. Both viewpoints can be limiting for moving sources: Doppler-induced time scaling violates the "same-frequency" assumptions behind standard cross-spectral estimators, while frame-wise delay estimates alone can be too noisy or ambiguous to yield stable kinematic estimates [1, 6, 9].

Notation. We use $p(t)$ for acoustic pressure. Microphone positions are denoted by $\mathbf{x}_m \in \mathbb{R}^3$. The (moving) source position is denoted by $\mathbf{y}(t) \in \mathbb{R}^3$; in discrete time we write $\mathbf{y}_k$. (We keep bold capitals such as $\mathbf{C}$ and $\mathbf{\Sigma}$ for matrices.) We use frequency f in Hz; when needed, $\omega = 2\pi f$.

This paper takes a velocity-beamforming perspective. We treat short-time cross-spectral statistics as a compact representation of two complementary cues: geometric delays constrain source position, while Doppler/warp effects constrain source velocity. Rather than relying on exhaustive joint position–velocity scanning, we estimate per-frame delay and warp measurements and fuse them over time with a Bayesian state estimator, in the spirit of regularized array processing [3, 7, 11].

The main contribution is a coherent pipeline: (i) a physics-based Doppler definition tied to retarded-time propagation, (ii) an inter-frequency coherence objective for velocity-sensitive beamforming relative to a reference microphone, and (iii) UKF-based fusion of delay and warp cues for robust kinematic tracking.

A related stabilization idea in broadband beamforming is to exploit that a physical source is (to first order) located at the same position across frequency, which allows pooling information across the band in optimization and deconvolution schemes [4, 5].

A key enabler is *inter-frequency coherence*: a coherence measure that compares spectral components at different frequencies after compensating the deterministic phase drift implied by time scaling. Inter-frequency coherence can be viewed as a frequency-domain surrogate for time-stretch alignment under a local-stationarity approximation, while avoiding the computational cost of resampling long waveforms for each candidate warp.

2 Inter-frequency coherence: theory

2.1 Signal model and time-scaling

We consider a single, wide–sense stationary (WSS), band-limited scalar source with (time-varying) position $\mathbf{y}(t) \in \mathbb{R}^3$ and (locally) constant velocity $\mathbf{v} \in \mathbb{R}^3$. An array has M microphones at fixed locations $\{\mathbf{x}_m\}_{m=1}^M$. The wave speed $c > 0$ is assumed known and constant.

Define the source–microphone geometry

$$\mathbf{d}_m \triangleq \mathbf{y} - \mathbf{x}_m, \qquad r_m \triangleq \|\mathbf{d}_m\|, \qquad \mathbf{u}_m \triangleq \frac{\mathbf{d}_m}{r_m} \tag{1}$$

where $\mathbf{u}_m$ is the line-of-sight unit vector pointing from microphone m to the source. The corresponding geometric delay is

$$\tau_m \triangleq \frac{r_m}{c}. \tag{2}$$

We neglect amplitude effects (e.g., $1/r$ spreading) because the estimators below rely on phase and relative Doppler.

For small Mach numbers $\|\mathbf{v}\|/c \ll 1$, the received acoustic pressure at microphone m is approximated by a time-scaled and delayed version of the source signal $s(t)$ [8, 10],

$$p_m(t) \approx s\big(\alpha_m\,(t - \tau_m)\big). \tag{3}$$

In the velocity-beamforming formulation below we use the line-of-sight unit vector $\mathbf{u}_m$

(pointing from microphone to source) and the normalized radial velocity

$$r_v(m) \triangleq \frac{\mathbf{u}_m^\top \mathbf{v}}{c}. \tag{4}$$

We adopt the physics-based (retarded-time) reciprocal Doppler factor

$$\alpha_m \triangleq \frac{1}{1+r_v(m)+\varepsilon}, \tag{5}$$

where $\varepsilon > 0$ is a small constant for numerical stability. With this convention, $r_v > 0$ corresponds to receding motion (redshift) and $r_v < 0$ to approaching motion (blueshift), so $\alpha_m > 1$ for approaching sources; the first-order retarded-time relation gives $\alpha_m \approx dt_e/dt_r = 1/(1+r_v(m))$. In the frequency domain this implies the dilation

$$X_m(\omega) \approx \frac{1}{\alpha_m} S\left(\frac{\omega}{\alpha_m}\right) e^{-j\omega\tau_m}. \tag{6}$$

For a fixed reference microphone r, we define the pairwise relative Doppler factor and its logarithmic measurement as

$$\beta_{mr} \triangleq \frac{\alpha_m}{\alpha_r}, \qquad \ell_{\mathrm{rel},m} \triangleq \log \beta_{mr} = \log(\alpha_m) - \log(\alpha_r). \tag{7}$$

For small Mach numbers, this yields the first-order approximation

$$\ell_{\mathrm{rel},m} \approx -\frac{(\mathbf{u}_m - \mathbf{u}_r)^\top \mathbf{v}}{c}. \tag{8}$$

2.2 Conventional coherence vs. inter-frequency coherence

Coherence measures how consistently two signals share a common component at a given frequency. Given two (complex) time–frequency representations $X(f,t)$ and $Y(f,t)$ (e.g., STFT coefficients), the magnitude-squared coherence at the *same* frequency is commonly defined as

$$\gamma^2_{XY}(f) \triangleq \frac{|P_{XY}(f)|^2}{P_{XX}(f)\,P_{YY}(f)+\varepsilon}, \qquad P_{XY}(f) \triangleq \langle X(f,t)Y^*(f,t)\rangle_t, \tag{9}$$

with $P_{XX}(f) \triangleq \langle |X(f,t)|^2\rangle_t$ and likewise for P_{YY}.

For moving sources, Doppler induces time scaling. A channel-dependent time scaling does not merely add a phase shift: it maps spectral content from f to a different frequency in the other channel. Consequently, a conventional "same-bin" coherence $\gamma^2_{mn}(f)$ between microphones may drop even when both microphones observe the same physical source, simply because each channel sees the source energy at slightly different frequencies [8].

Inter-frequency idea. Rather than forcing the comparison to the same frequency bin, we explicitly compare f in the reference channel to $\beta_{mr} f$ in microphone m. However, comparing two different frequencies introduces a deterministic phase drift over time because the Fourier

kernel depends on the time stamp. If this drift is not compensated, the cross-product averages out and coherence collapses. We treat each STFT frame as locally stationary and interpret the resulting inter-frequency coherence as a practical heuristic for small Doppler factors and short windows.

2.3 STFT/Welch formulation with phase-drift compensation

Let $X_m(f_k, t_\ell)$ denote the complex short-time Fourier transform (STFT) of the microphone signal $p_m(t)$. Because the tracker uses reference-relative measurements, we formulate the coherence for a microphone m relative to a fixed reference microphone r. For a candidate pairwise relative factor β_{mr}, we map bins and frames using nearest-neighbor rules

$$j(k,\beta_{mr}) \triangleq \arg\min_j |f_j - \beta_{mr} f_k|, \qquad \ell'(\ell,\beta_{mr}) \triangleq \arg\min_{\ell'} \left|t_{\ell'} - t_\ell/\beta_{mr}\right|. \tag{10}$$

The reciprocal frame mapping ensures that the same emission time is compared across the two receive-time scales. Nearest-neighbor remapping is computationally cheap but approximate; interpolation can reduce bias for larger Doppler factors. We rephase each STFT coefficient by its time stamp,

$$\widetilde{X}_r(f_k,t_\ell) \triangleq X_r(f_k,t_\ell)\, e^{+j2\pi f_k t_\ell}, \qquad \widetilde{X}_m(\beta_{mr}f_k,t_{\ell'}) \triangleq X_m(f_{j(k,\beta_{mr})},t_{\ell'})\, e^{+j2\pi(\beta_{mr}f_k)t_{\ell'}}. \tag{11}$$

The Doppler-compensated cross-spectrum becomes

$$P_{mr}(f_k,\beta_{mr}) \triangleq \left\langle \widetilde{X}_r(f_k,t_\ell)\, \widetilde{X}_m(\beta_{mr}f_k,t_{\ell'})^* \right\rangle_\ell, \tag{12}$$

leading to the per-frequency coherence

$$\gamma^2_{mr}(f_k,\beta_{mr}) \triangleq \frac{|P_{mr}(f_k,\beta_{mr})|^2}{P_{rr}(f_k,\beta_{mr})\, P_{mm}(\beta_{mr}f_k,\beta_{mr}) + \varepsilon}, \tag{13}$$

where $P_{rr}(f_k,\beta_{mr}) \triangleq \langle|\widetilde{X}_r(f_k,t_\ell)|^2\rangle_\ell$ and $P_{mm}(\beta_{mr}f_k,\beta_{mr}) \triangleq \langle|\widetilde{X}_m(\beta_{mr}f_k,t_{\ell'})|^2\rangle_\ell$. Let $\mathscr{B}$ denote the retained set of positive-frequency STFT bins in the analysis band. The broadband objective is then

$$\Gamma_{mr}(\beta_{mr}) \triangleq \frac{1}{|\mathscr{B}|} \sum_{k\in\mathscr{B}} \gamma^2_{mr}(f_k,\beta_{mr}). \tag{14}$$

3 Inter-frequency coherence velocity beamforming and tracking

We cast moving-source tracking as a velocity-beamforming front-end followed by Bayesian state fusion. The source state leaves a structured signature in short-time cross-spectral statistics: relative delays encode geometry, while Doppler/warp encodes radial motion.

3.1 Forward model for validation (retarded time propagation)

For synthetic validation, microphone signals are generated in the time domain via a retarded-time mapping, similar in spirit to moving-source modeling approaches used in array-based studies

[10]. Let t_r denote receive time and t_e emission time. For microphone m, the retarded time satisfies

$$t_r = t_e + \frac{R_m(t_e)}{c}, \qquad R_m(t_e) \triangleq \|\mathbf{x}_m - \mathbf{y}(t_e)\|, \tag{15}$$

where $\mathbf{x}_m$ is the microphone position and $\mathbf{y}(t)$ the source trajectory. We solve t_e by fixed-point iteration and synthesize

$$p_m(t_r) = a_m(t_e)\, s(t_e), \qquad a_m(t_e) \propto \frac{1}{R_m(t_e)}. \tag{16}$$

Linearizing the retarded-time relation yields $dt_e/dt_r \approx 1/(1+r_v(m))$, consistent with the Doppler factor in equation (5) under the above sign convention.

3.2 Measurement extraction and UKF fusion

Per STFT frame (optionally averaged over a short temporal neighborhood), we compute a short-time cross-spectral matrix (CSM). We choose a fixed reference microphone r and extract two measurement types relative to r. Here $\mathscr{A} \subset \mathbb{R}_+$ denotes the discrete search grid of candidate pairwise relative Doppler factors centered around unity, and $\mathscr{B}$ denotes the retained positive-frequency analysis band used for both GCC-PHAT and inter-frequency coherence. The self-entry $m = r$ is omitted from the measurement vector, so both $\boldsymbol{\tau}_{\mathrm{rel}}$ and $\boldsymbol{\ell}_{\mathrm{rel}}$ contain $M-1$ entries.

Relative delays (GCC-PHAT). Let $C_{mr}(f) \triangleq X_m(f)X_r^*(f)$ and $G_{mr}(f) \triangleq C_{mr}(f)/|C_{mr}(f)|$. We estimate

$$\widehat{\tau}_{\mathrm{rel},m} \triangleq \arg\max_{\tau} \left| \sum_{f\in\mathscr{B}} G_{mr}(f)\, e^{j2\pi f\tau} \right|. \tag{17}$$

Relative log-warps (inter-frequency coherence). For each non-reference microphone, we maximize the broadband objective in equation (14) against the reference:

$$\widehat{\beta}_{mr} \triangleq \arg\max_{\beta\in\mathscr{A}} \Gamma_{mr}(\beta), \qquad \ell_{\mathrm{rel},m} \triangleq \log \widehat{\beta}_{mr}, \tag{18}$$

The warp measurements are thus explicit pairwise relative quantities with respect to the reference microphone.

UKF measurement model. We fuse both streams in an Unscented Kalman Filter (UKF) with a constant-acceleration (CA) state $\mathbf{s}_k = [\mathbf{y}_k^\top, \mathbf{v}_k^\top, \mathbf{a}_k^\top]^\top$ and covariance $\boldsymbol{\Sigma}_k$. The measurement vector and model are

$$\mathbf{z}_k = \begin{bmatrix} \boldsymbol{\tau}_{\mathrm{rel},k} \\ \boldsymbol{\ell}_{\mathrm{rel},k} \end{bmatrix} = \mathbf{h}(\mathbf{s}_k) + \mathbf{n}_k, \qquad \mathbf{n}_k \sim \mathscr{N}(\mathbf{0}, \mathbf{R}), \tag{19}$$

with

$$\tau_{\mathrm{rel},m}(\mathbf{y}) = \frac{\|\mathbf{x}_m - \mathbf{y}\| - \|\mathbf{x}_r - \mathbf{y}\|}{c}, \qquad \ell_{\mathrm{rel},m}(\mathbf{y},\mathbf{v}) = \log\beta_{mr} = \log\left(\frac{\alpha_m}{\alpha_r}\right), \tag{20}$$

where α_m is defined in equation (5) and the small-Mach interpretation is given by equation (8). For simplicity, we use a diagonal noise model $\mathbf{R} = \mathrm{diag}(\sigma_\tau^2\mathbf{I}, \sigma_\ell^2\mathbf{I})$.

3.3 Forecasting strategies after online tracking

After the last measurement update at index k_0, we forecast future states using one of the following strategies. The baseline CA-Gaussian forecast uses

$$\mathbf{s}_{k+1} = \mathbf{F}\,\mathbf{s}_k + \mathbf{w}_k, \qquad \mathbf{w}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}), \tag{21}$$

with per-axis

$$\mathbf{F}_{\mathrm{CA}}(\Delta t) = \begin{bmatrix} 1 & \Delta t & \frac{1}{2}\Delta t^2 \\ 0 & 1 & \Delta t \\ 0 & 0 & 1 \end{bmatrix}, \qquad \mathbf{Q}_{\mathrm{CA}}(\Delta t, q_j) = q_j \begin{bmatrix} \frac{\Delta t^5}{20} & \frac{\Delta t^4}{8} & \frac{\Delta t^3}{6} \\ \frac{\Delta t^4}{8} & \frac{\Delta t^3}{3} & \frac{\Delta t^2}{2} \\ \frac{\Delta t^3}{6} & \frac{\Delta t^2}{2} & \Delta t \end{bmatrix}. \tag{22}$$

The prediction recursion is $\widehat{\mathbf{s}}_{k+1} = \mathbf{F}\widehat{\mathbf{s}}_k$ and $\mathbf{\Sigma}_{k+1} = \mathbf{F}\mathbf{\Sigma}_k\mathbf{F}^\top + \mathbf{Q}$. Adaptive-$\mathbf{Q}$ replaces fixed q_j by Normalized Innovation Squared (NIS) driven adaptation during tracking. An Interacting Multiple Model (IMM) mixes several constant-acceleration/constant-velocity (CA/CV) models with different maneuverability; Savitzky–Golay (SG) extrapolates locally smoothed derivatives; reachability produces bounded envelopes; and particle forecasting propagates samples from the final posterior [2, 6].

4 Experiments and results

We validate the proposed velocity-beamforming tracking pipeline in a fully synthetic 3D scenario (figure 1), where both the source trajectory and the ground-truth kinematics are known. Figures 2 to 4 summarize the UKF estimates for position, velocity, and acceleration, respectively. In this paper we focus on the qualitative behavior visible in the tracking plots; quantitative benchmarking and ablation studies are left for future work.

Tracking accuracy and filtering-induced lag. The position estimates remain consistently close to the ground truth over the entire trajectory (figure 2). This is expected because (i) the relative-delay (TDOA) measurements provide strong geometric constraints and (ii) position is the most directly observed component of the state. For velocity (figure 3), the estimates remain accurate but exhibit a small temporal lag. This lag is a standard filtering effect in state-space smoothing: increasing the UKF measurement averaging and/or process-noise regularization yields smoother curves, but at the cost of delayed response to rapid changes. The effect is more pronounced for acceleration (figure 4), which is (in our model) the highest-order state component and therefore inferred mainly through the dynamics prior and innovation history; as a result, it shows the largest apparent lag, yet still captures the overall trend well. Overall, the combined delay-and-warp measurement model yields stable kinematic tracking and physically plausible trajectories in this synthetic proof-of-concept setting.

Uncertainty envelopes (2σ). The shaded bands shown in figures 2 to 4 visualize a 2σ interval derived from the UKF state covariance Σ_k. In the plots, the blue shaded area denotes $\pm 2\sigma$. For each state component, we plot $\hat{x}_k \pm 2\sqrt{\Sigma_{k,ii}}$ (diagonal entries), which—under the Gaussian approximation underlying the UKF—corresponds to an approximate $\approx 95\%$ marginal credibility range. The tight envelopes in position indicate strong observability from the delay measurements, whereas the wider bands for velocity/acceleration reflect the increasing ambiguity of higher-order derivatives. These intervals are marginal and not empirically calibrated; coverage evaluation is left for future work.

Prediction after the last update. Figure 6 compares several forecasting strategies once measurements are no longer available. In our experiments, very short horizons (roughly $< 0.5\,\mathrm{s}$) are best handled by Savitzky–Golay (SG) extrapolation, which leverages local polynomial fits and therefore tracks the immediate trend well. For slightly longer horizons, the Interacting Multiple Model (IMM) and reachability-based envelopes are more robust, as they explicitly represent maneuver uncertainty (IMM) or bounded future sets (reachability), avoiding overconfident extrapolation.

Computational aspects. The dominant computational cost lies in feature extraction. Per STFT frame, the method uses short-time cross-spectral statistics, a one-dimensional search over β_{mr} for each non-reference microphone, and a UKF update on a small state vector (9 states for CA in 3D). We do not report a dedicated runtime benchmark here; the main algorithmic point is that the warp search remains low-dimensional compared with exhaustive joint position–velocity grid scanning.

5 Discussion

The experiments highlight that jointly exploiting (i) relative delays (TDOA) and (ii) relative time-scaling information (warp/Doppler proxies) yields a well-conditioned tracking formulation in this synthetic setting. From an estimation perspective, these two feature types are complementary: delay constraints dominate geometric observability and stabilize the position estimate, while inter-frequency coherence adds sensitivity to radial motion and thereby improves the identifiability of velocity components. This division of roles is consistent with the observed behavior in figures 2 to 4, where position uncertainty remains tight over the trajectory and higher-order kinematics show wider credibility bands.

A practical implication is that the proposed pipeline can be interpreted as a velocity-beamforming front-end in front of a standard Bayesian tracker. Instead of scanning a dense position–velocity grid, we compress the multichannel data into a low-dimensional measurement vector per frame. This keeps the subsequent filter update inexpensive and makes it straightforward to incorporate priors (e.g., constant-acceleration) and uncertainty quantification via Σ_k. In addition, the resulting posterior can be propagated forward for forecasting, enabling a principled comparison of extrapolation strategies (figure 6).

Limitations. The current study is intentionally proof-of-concept and uses a simplified propagation model and controlled synthetic data. The present evaluation also does not include an

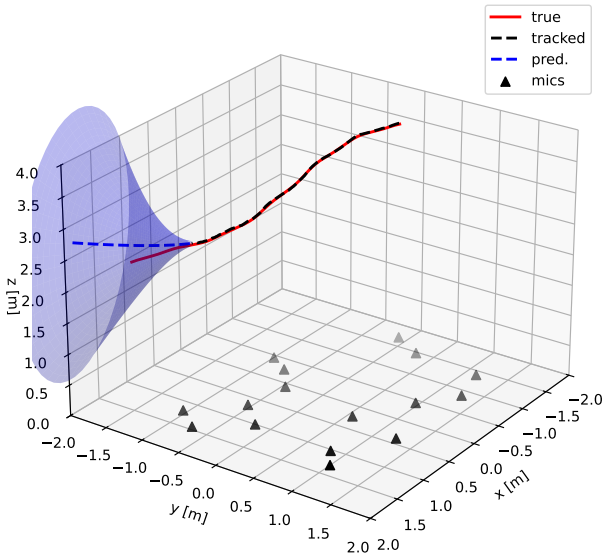

Figure 1: Synthetic 3D setup and trajectory.

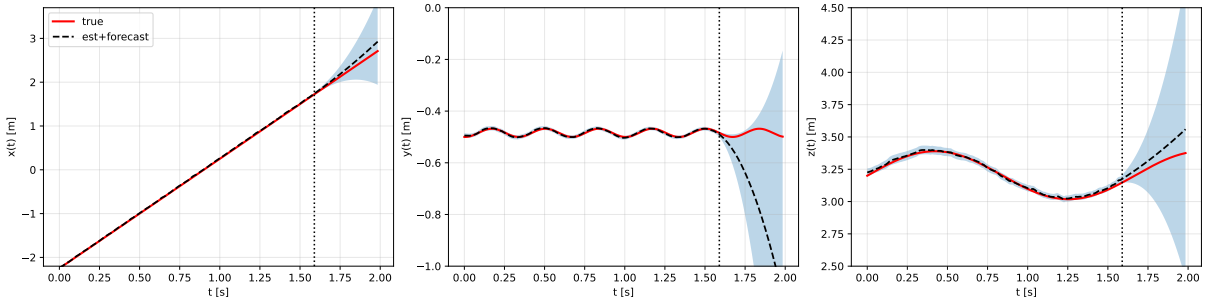

Figure 2: Position tracking results with uncertainty envelopes.

explicit ablation against a TDOA-only baseline or a same-frequency coherence baseline, so claims should be interpreted within this synthetic setting. Several factors can affect performance in real recordings: multipath and reflections may bias both delay and coherence-based warp estimates; low SNR and short analysis windows increase estimator variance; and microphone synchronization errors can appear as systematic delay offsets. Moreover, the local time-scaling approximation in equation (3) is most accurate for small Mach numbers and sufficiently short frames. Finally, while inter-frequency coherence is robust to unknown source spectra in many cases, it can degrade for highly narrowband content or rapidly time-varying source statistics. We also use nearest-neighbor time/frequency remapping for Doppler compensation; interpolation could reduce bias at larger warps. The measurement model is phase-based (PHAT/coherence) and thus largely amplitude-invariant, whereas the synthetic forward model includes $1/r$ amplitude; a

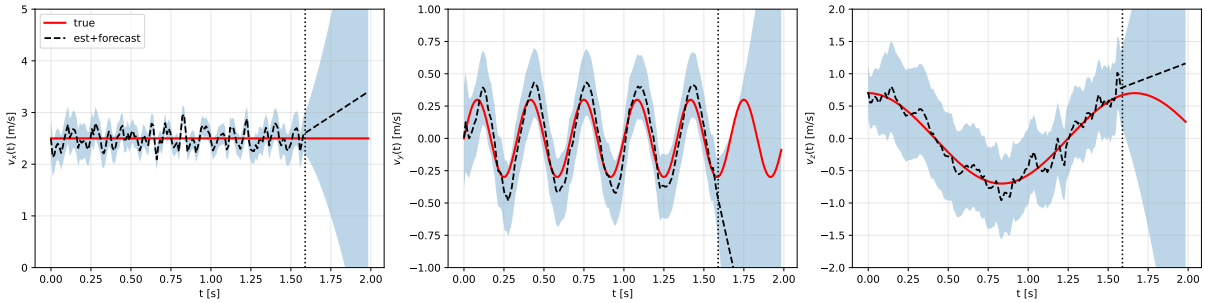

Figure 3: Velocity tracking results with uncertainty envelopes.

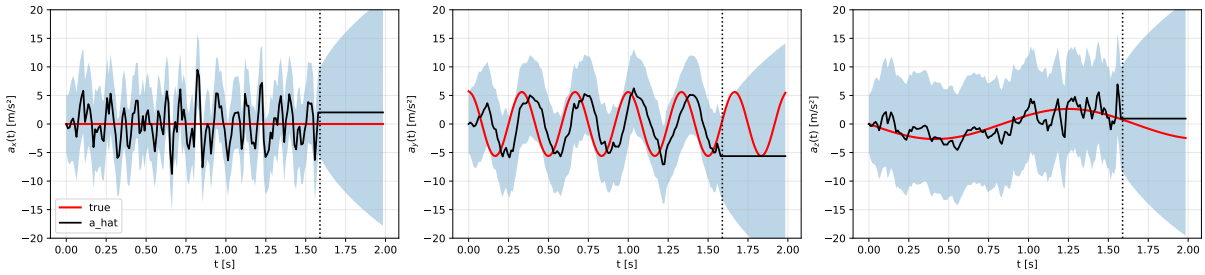

Figure 4: Acceleration tracking results with uncertainty envelopes.

fully matched amplitude-agnostic forward model or explicit amplitude modeling is left for future work.

Outlook. Future work includes validating the method on measured array data in reverberant environments; extending the measurement model to jointly account for amplitude cues and mild multipath; and integrating more expressive dynamical priors (e.g., maneuver models in an IMM, or constraints derived from vehicle dynamics). On the signal-processing side, improved bin-remapping/interpolation and outlier-robust aggregation across microphone pairs may further reduce lag and increase robustness, especially for acceleration estimation.

6 Conclusion

We introduced an inter-frequency-coherence source-velocity beamforming approach for moving acoustic sources. Its key component is a physics-consistent frequency-domain estimator of the pairwise relative Doppler factor β_{mr}, which complements conventional delay (TDOA) features. Fusing delay and warp measurements in a UKF with a constant-acceleration prior yields smooth position and velocity estimates, informative acceleration trends, and uncertainty envelopes in the synthetic proof-of-concept scenario studied here. Overall, the results indicate that velocity-sensitive beamforming based on inter-frequency coherence is a promising alternative to exhaustive joint position–velocity grid scanning for this class of tracking problem.

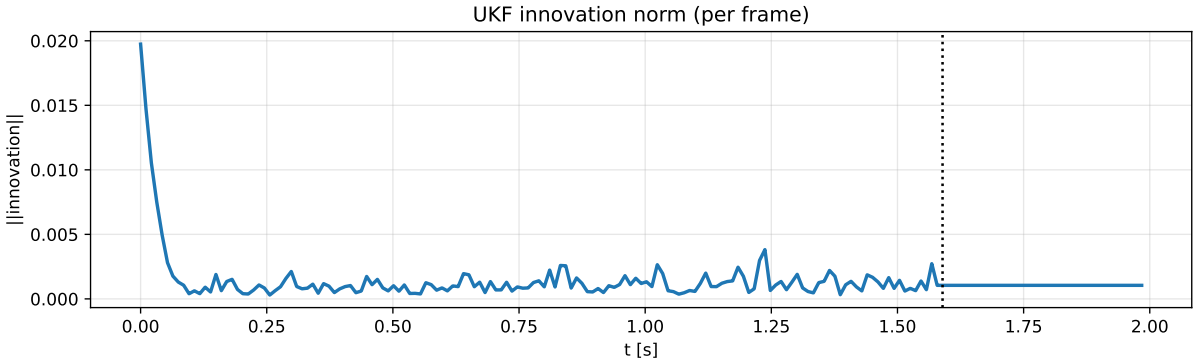

Figure 5: UKF innovation norm over time.

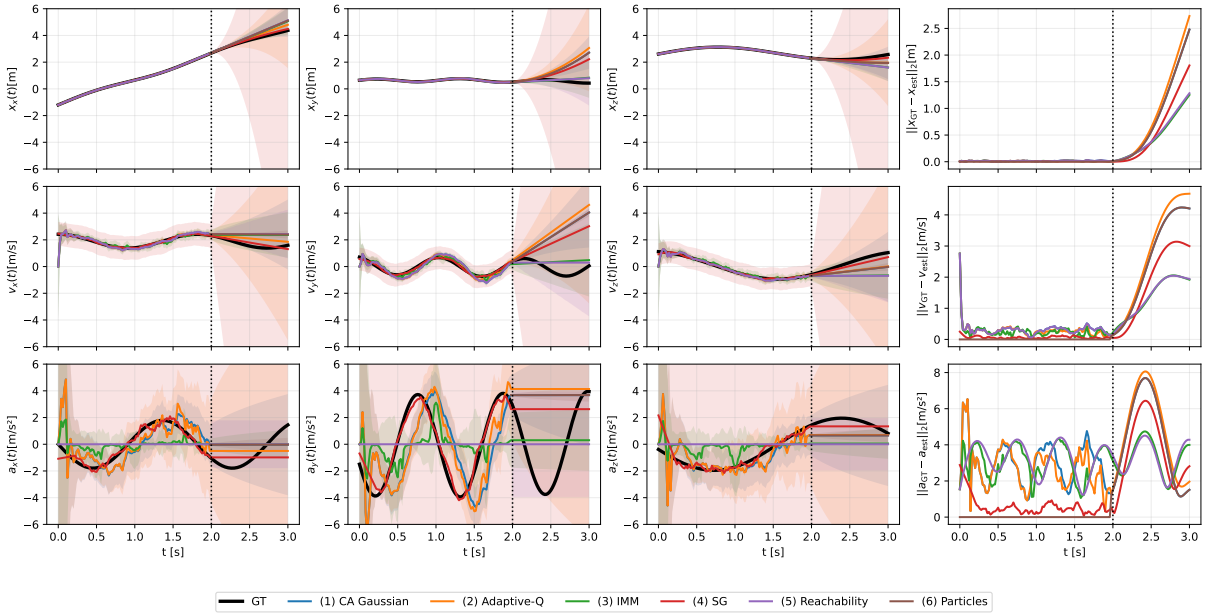

Figure 6: Comparison of forecasting strategies after tracking. The black dashed line indicates the start of the forecasting. The uncertainty is depicted as a shaded area.

References

[1] M. Basiri, F. Schill, P. U. Lima, and D. Floreano. "Robust acoustic source localization of emergency signals from micro air vehicles." In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2012. doi:10.1109/iros.2012.6385608.

[2] B. Csóka, P. Fiala, and P. Rucz. "Tracking sound sources with microphone arrays and beamforming algorithms." In *Proceedings of the 10th Convention of the European Acoustics Association Forum Acusticum 2023*, FA2023, pages 1115–1122. European Acoustics Association, 2024. doi:10.61782/fa.2023.0310.

[3] P.-A. Gauthier, C. Camier, Y. Pasco, A. Berry, E. Chambatte, R. Lapointe, and M.-A.

Delalay. "Beamforming regularization matrix and inverse problems applied to sound field measurement and extrapolation using microphone array." *Journal of Sound and Vibration*, 330(24), 5852–5877, 2011. ISSN 0022-460X. doi:10.1016/j.jsv.2011.07.022. URL http://dx.doi.org/10.1016/j.jsv.2011.07.022.

[4] A. Goudarzi. "B-CLEAN-SC: CLEAN-SC for broadband sources." *JASA Express Letters*, 3(9), 2023. ISSN 2691-1191. doi:10.1121/10.0020992.

[5] A. Goudarzi. "Global, and local optimization beamforming for broadband sources." *The Journal of the Acoustical Society of America*, 155(1), 262–273, 2024. ISSN 0001-4966. doi:10.1121/10.0024247.

[6] H. HAO and Z. DUAN. "Periodic acoustic source tracking using propagation delayed measurements." *Chinese Journal of Aeronautics*, 35(4), 390–399, 2022. ISSN 1000-9361. doi:10.1016/j.cja.2021.04.017.

[7] C. H. Kasess, W. Kreuzer, P. Soni, and H. Waubke. "Localizing uniformly moving single-frequency sources using an inverse 2.5D approach." *Journal of Sound and Vibration*, 593, 118653, 2024. ISSN 0022-460X. doi:10.1016/j.jsv.2024.118653.

[8] M. Lindfors, G. Hendeby, F. Gustafsson, and R. Karlsson. "On frequency tracking in harmonic acoustic signals." In *2017 20th International Conference on Information Fusion (Fusion)*, pages 1–8. IEEE, 2017. doi:10.23919/icif.2017.8009833.

[9] W. Ma, H. Bao, C. Zhang, and X. Liu. "Beamforming of phased microphone array for rotating sound source localization." *Journal of Sound and Vibration*, 467, 115064, 2020. ISSN 0022-460X. doi:10.1016/j.jsv.2019.115064.

[10] F. Meng. *Modeling of moving sound sources based on array measurements*, volume RWTH Aachen University. RWTH Aachen University, 2018. doi:10.18154/RWTH-2018-227109.

[11] L. Yu, Q. Guo, N. Chu, and R. Wang. "Achieving 3D beamforming by non-synchronous microphone array measurements." *Sensors*, 20(24), 7308, 2020. ISSN 1424-8220. doi:10.3390/s20247308.