Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

# Toward Secure Decentralized Coordination and Acoustic Source Localization in Autonomous Drone Swarms

**Md Shahanur Islam Shagor**

smshagor.ru@gmail.com

Voronezh State University of Forestry and Technologies https://orcid.org/0009-0003-7730-3202

**Research Article**

**Keywords:** autonomous drone swarms, acoustic source localization, mathematical modeling, time difference of arrival, secure coordination, fault-tolerant autonomy

**Posted Date:** July 31st, 2026

**DOI:** https://doi.org/10.21203/rs.3.rs-10530148/v1

**License:** This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

**Additional Declarations:** The authors declare no competing interests.

# Toward Secure Decentralized Coordination and Acoustic Source Localization in Autonomous Drone Swarms

**Md Shahanur Islam Shagor**
Department of Computer Science and Microelectronics Engineering,
Voronezh State University of Forestry and Technologies, Voronezh, Russia
*Corresponding author: smshagor.dev@gmail.com*

## Abstract

Secure autonomous drone swarms require perception, coordination, communication protection, and fault response to remain mutually consistent under uncertain sensing and latency. This article develops a mathematical and implementation-grounded analysis of a simulation-first multi-drone platform that combines event-driven leader/follower coordination, acoustic source localization, dynamic-obstacle avoidance, latency-aware fallback, AES-256-GCM communication, and an Ed25519-signed replicated flight ledger. Rather than treating architectural documentation as validated evidence, active source code is mapped to explicit equations and evaluated through reproducible code-equivalent experiments. The acoustic subsystem is formalized as a hyperbolic time-difference-of-arrival problem, including residual Jacobians, geometry conditioning, integer-lag quantization, robust nonlinear least squares, and confidence calibration. Four-sensor trials at 48 kHz yielded a median error of 0.0018 m for ideal impulses, 1.21 m at noise standard deviation 0.05, and 12.30 m at 0.10. The latency-triggered three-sensor mode produced 11 high-confidence errors above 5 m in 36 trials, demonstrating that residual fit alone is not a safe confidence measure in an exactly determined geometry. Sequential ledger replication followed the analytical model $T_{rep}(N) \approx 2.44 + 111.20(N-1)$ microseconds with $R^2 = 0.999998$, while leader-score selection remained linear in candidate count. The results support a mathematically transparent simulation platform, but not yet secure physical-swarm autonomy. A staged validation path is therefore specified.

**Keywords:** autonomous drone swarms; acoustic source localization; mathematical modeling; time difference of arrival; secure coordination; fault-tolerant autonomy

## 1 Introduction

Cooperative unmanned aerial vehicles can extend sensing coverage, reduce mission completion time, and preserve capability after individual vehicle failures. These advantages arise only when coordination, perception, networking, and low-level safety remain synchronized. A swarm that elects a leader rapidly but accepts forged commands is unsafe; a swarm that encrypts traffic but cannot respond to delayed perception is brittle; and an acoustic estimator that reports a small residual without accounting for weak geometry can confidently direct the formation toward an incorrect location. Secure autonomy must therefore be treated as a coupled mathematical and systems problem.

Multi-robot coordination has been studied through consensus, flocking, task allocation, behavior-based fault tolerance, and distributed control (Cao et al., 1997; Parker, 1998; Olfati-Saber, 2006; Ren and Beard, 2008). Aerial swarms add strict energy, aerodynamic, localization, collision, and communication constraints (Chung et al., 2018). Flying ad hoc network research similarly shows that mobility, topology change, intermittent connectivity, and rapidly varying link quality complicate dependable command exchange (Bekmezci et al., 2013; Sahingoz, 2014). These issues become more severe when GNSS, camera visibility, or centralized network infrastructure is unavailable.

Acoustic sensing is a useful complementary modality in darkness, smoke, visual occlusion, and scenarios in which an emitting target is easier to hear than to observe. Time-difference-of-arrival (TDOA) localization infers a source location from relative propagation delays measured at spatially separated sensors. Generalized cross-correlation with phase transform (GCC-PHAT) is a classical delay estimator, while hyperbolic least-squares methods convert delay differences into a position estimate (Knapp and Carter, 1976; Chan and Ho, 1994). In an

aerial swarm, microphone positions can form a large distributed array, but rotor noise, clock error, moving receivers, reverberation, and transport delay challenge the ideal model.

The analyzed repository implements an end-to-end Python and C++ research platform with a PyQt operator interface, event-driven swarm management, optional MAVSDK connections, TDOA acoustic localization, dynamic-obstacle prediction, latency monitoring, AES-GCM communication, Ed25519-signed flight-ledger records, and low-level motor-health logic. The default runtime constructs five simulated drones and populates a dynamic obstacle field. Real-drone connections are optional and environment-controlled. The appropriate claim is therefore a simulation-first implementation study, not a flight-certified autonomous system.

A central motivation for rewriting the system as mathematics is documentation drift. The repository contains multiple prose descriptions of leader scoring and security behavior, but the active code is more specific. The active leader election uses battery, signal strength, and processing capability with weights 0.4, 0.3, and 0.3; motor health is not present in the implemented score. Likewise, the in-process event bus, AES-GCM transport module, and Ed25519 ledger are distinct mechanisms. Combining them into one conceptual "signed command" path would overstate the implementation. This study therefore treats executable source at the analyzed commit as authoritative and reconstructs only equations that correspond to active operations.

The principal contributions are:

1. a mathematical reconstruction of the active coordination, acoustic, avoidance, latency, communication, and ledger mechanisms;
2. a geometry-aware analysis of TDOA identifiability, Jacobian conditioning, integer-lag resolution, and confidence failure under sensor reduction;
3. reproducible code-equivalent experiments for noise, sample rate, latency gating, AES-GCM cost, ledger replication, tamper rejection, and leader-election scaling;
4. analytical complexity models linked directly to measured coefficients rather than only asymptotic notation; and
5. a strict boundary between simulation evidence and the additional calibration, identity, synchronization, and physical-safety evidence required for real deployment.

# 2 Related work

## 2.1 Multi-robot coordination and aerial swarms

Early cooperative robotics research organized coordination problems around communication, task decomposition, architecture, and conflict resolution (Cao et al., 1997). ALLIANCE demonstrated behavior-based fault-tolerant cooperation through distributed motivational mechanisms (Parker, 1998). Consensus and flocking work supplied mathematical tools for agreement and cohesive motion in networks of mobile agents (Olfati-Saber, 2006; Ren and Beard, 2008). Surveys of swarm robotics and aerial swarms emphasize the interactions among agent count, communication, energy, localization, and collision avoidance (Brambilla et al., 2013; Bayindir, 2016; Chung et al., 2018).

The analyzed platform uses a leader/follower organization and an event bus rather than a distributed consensus protocol. A central `SwarmManager` filters eligible drones, computes a scalar suitability score, and publishes explicit leader commands through registered handlers. This provides deterministic software routing and leader replacement after a detected failure, but it does not establish agreement under network partitions, Byzantine behavior, or concurrent leaders. The manuscript therefore separates mathematical score selection from the stronger guarantees associated with distributed election protocols.

## 2.2 Acoustic delay estimation and hyperbolic localization

The generalized correlation framework of Knapp and Carter (1976) remains foundational for acoustic delay estimation. PHAT weighting normalizes cross-spectrum magnitude and emphasizes phase, often improving robustness to source coloration and reverberation. A TDOA observation defines a hyperbola because it constrains a difference of distances rather than an absolute range (Chan and Ho, 1994). Microphone-array

literature identifies sensor geometry, synchronization, reverberation, signal bandwidth, and estimator variance as central sources of error (Brandstein and Ward, 2001; Cobos et al., 2017).

Distributed microphones mounted on aerial robots increase aperture but also introduce rotor noise, receiver motion, clock drift, and transport delay. The active implementation estimates delay relative to the lowest-index sensor, evaluates direct FFT correlation and GCC-PHAT sequences, selects a lag according to peak magnitude, and solves a robust nonlinear least-squares problem from multiple starting points. This hybrid peak-selection rule is heuristic; it is not a likelihood-ratio test or a calibrated probabilistic estimator.

## 2.3 Collision avoidance and latency-aware autonomy

Velocity-obstacle methods identify relative velocities that produce future collisions and select admissible alternatives (Fiorini and Shiller, 1998). Reciprocal methods distribute responsibility among agents and support dense multi-agent movement (van den Berg et al., 2011; Snape et al., 2011). The analyzed platform instead applies short-horizon kinematic prediction, a collision-cone probability, a learned risk memory, avoidance-vector accumulation, velocity smoothing, acceleration saturation, and waypoint replanning.

Latency-aware fallback is particularly important in mixed high-level and low-level autonomy stacks. A remote or high-level planner may improve decision quality but introduce transport and execution delay. A fallback should therefore be analyzed as a switching system: it reduces dependence on delayed data but can also reduce sensor redundancy or planning quality. The acoustic three-sensor fallback in this study provides a concrete example of that trade-off.

## 2.4 Secure communication and tamper-evident logs

Authenticated encryption protects confidentiality and integrity jointly. AES-GCM is standardized by NIST and provides an authentication tag together with counter-mode encryption (Dworkin, 2007). Ed25519 provides compact signatures with strong performance and mature implementations (Bernstein et al., 2012; Josefsson and Liusvaara, 2017). SHA3-256 supplies a standardized hash for linked evidence records (National Institute of Standards and Technology, 2015).

The active communication module derives a shared 256-bit key with PBKDF2-HMAC-SHA-256, encrypts JSON messages with AES-GCM, and embeds sender, timestamp, and sequence metadata in the authenticated plaintext. The ledger separately hashes telemetry and event payloads, links blocks with SHA3-256, signs each block hash with an Ed25519 key, and sequentially verifies replicas. These mechanisms protect different properties and should not be conflated with a Byzantine consensus protocol or a per-command public-key authorization system.

# 3 Evaluated system and research design

## 3.1 Repository snapshot and execution boundary

The analysis targets the public repository `smshagor-dev/decentralized-coordination-and-acoustic-localization-in-secure-autonomousa-drone-swarms` at commit `5505421886f8305312ba8e98e9a0599357eb36c6`. The default application creates five simulated drones, starts the event-driven manager, introduces dynamic obstacles, and launches a PyQt operator interface. Environment variables can enable MAVSDK connections, but the experiments reported here are software-based and do not represent physical flight trials.

The evaluated architecture contains four practical layers: (i) a Python coordination and sensing layer; (ii) a C++ low-level controller interface; (iii) an AES-GCM communication module; and (iv) per-drone Ed25519-signed evidence ledgers. Figure 1 shows the active boundaries without merging the in-process event path with the separate encrypted-network and signed-ledger paths.

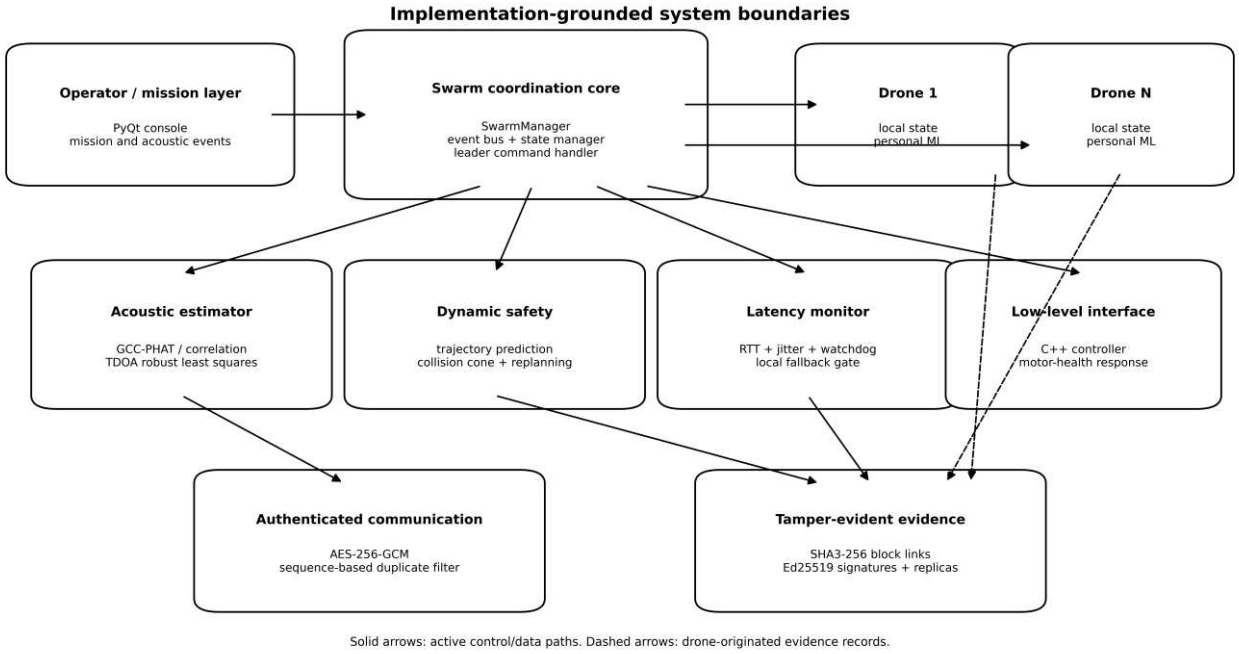



*Fig. 1 Implementation-grounded architecture and separation of control, sensing, communication, evidence, and low-level safety paths.*

**Table 1 Active implementation layers and mathematical roles**

| Layer | Active implementation | Mathematical role |
|---|---|---|
| Coordination | eligibility filtering, weighted score, event bus, state transitions | constrained argmax and finite-state transition system |
| Acoustic | correlation/GCC-PHAT, TDOA, robust least squares | nonlinear hyperbolic inverse problem |
| Avoidance | trajectory prediction, collision cone, velocity blending | switching kinematics with acceleration saturation |
| Latency | RTT, rolling mean, jitter, watchdog | thresholded mode-switching policy |
| Communication | PBKDF2-HMAC-SHA-256 and AES-GCM | authenticated-encryption cost and packet expansion |
| Evidence | SHA3-256 linked blocks and Ed25519 signatures | append/verify relation with sequential replication cost |

## 3.2 Research questions

The study addresses four questions.

**RQ1:** How do noise, sampling frequency, sensor geometry, and latency-triggered reduction from four sensors to three affect the implemented TDOA estimator?

**RQ2:** What computational and bandwidth costs are introduced by AES-GCM protection and Ed25519-signed ledger replication at prototype scale?

**RQ3:** How does the active leader-election operation scale with candidate count, and what coordination guarantees are absent from the central event-driven implementation?

**RQ4:** Which technical gaps must be closed before simulation evidence can support claims about secure physical drone-swarm autonomy?

## 3.3 Evidence classification

Evidence is classified as implemented behavior, automated test specification, code-equivalent measurement, or deferred capability. The presence of a test assertion is not treated as proof that every platform configuration passed. No independently verifiable complete cross-platform CI result was available for the analyzed commit; accordingly, the article reports inspected test contracts and independently generated measurements without claiming universal CI success.

# 4 Mathematical formulation

## 4.1 Notation

Let $N$ denote the number of active drones, $M$ the number of acoustic sensors used in one estimate, $\mathbf{p}_i \in \mathbb{R}^2$ sensor position, $\mathbf{s} \in \mathbb{R}^2$ acoustic-source position, $c = 343$ m/s speed of sound, and $f_s$ sample rate. Let $T_{rtt}$ be round-trip latency, $T_a$ the acoustic latency threshold, and $H$ the obstacle-prediction horizon.

**Table 2 Principal notation**

| Symbol | Meaning |
|---|---|
| $B_i, L_i, C_i$ | normalized battery, link/signal, and processing indicators for drone $i$ |
| $S_i$ | implemented leader suitability score |
| $d_i(\mathrm{s})$ | source-to-sensor range |
| $\Delta t_{1j}$ | TDOA between reference sensor 1 and sensor $j$ |
| $\mathrm{r}(\mathrm{s})$ | stacked TDOA residual vector |
| $J(\mathrm{s})$ | residual Jacobian with respect to source coordinates |
| $\kappa(J^T J)$ | geometry-conditioning indicator |
| $\mathrm{p}_o, \mathrm{v}_o, \mathrm{a}_o$ | obstacle position, velocity, and acceleration |
| $T_{enc}, T_{dec}, T_{rep}$ | encryption, decryption, and replication time |

## 4.2 Leader eligibility and weighted selection

Define an eligibility indicator $e_i \in \{0,1\}$ that is zero when drone $i$ is inactive, crashed, emergency-landing, grounded, assigned an emergency role, or under an active emergency condition. The active source code computes

$$S_i = 0.4B_i + 0.3L_i + 0.3C_i,$$

and chooses

$$i^* = \arg\max_{1 \le i \le N} e_i S_i.$$

If the eligible set is empty, no leader is elected. For one pass through $N$ drones, score evaluation and maximum selection require

$T_{elect}(N) = \alpha + \beta N + o(N)$, $\mathcal{O}(N)$ time and $\mathcal{O}(N)$ temporary score storage in the present implementation.

A streaming maximum could reduce temporary storage to $\mathcal{O}(1)$, but this would not change the central authority or failure assumptions. The implemented score is also not a reliability probability. It is a weighted ranking function; weights sum to one, but the inputs are not statistically calibrated and may be stale or self-reported.

After selection, a command event is represented by

$$\mathcal{E}_k = (u_k, \theta_k, i^*, t_k),$$

where $u_k$ is command type, $\theta_k$ is payload, $i^*$ is issuing leader, and $t_k$ is issue time. A state transition for drone $j$ is

$$x_j(k+1) = \delta\big(x_j(k), \mathcal{E}_k\big),$$

where $\delta$ is the registered event handler. This is a central finite-state transition system, not a distributed consensus relation.

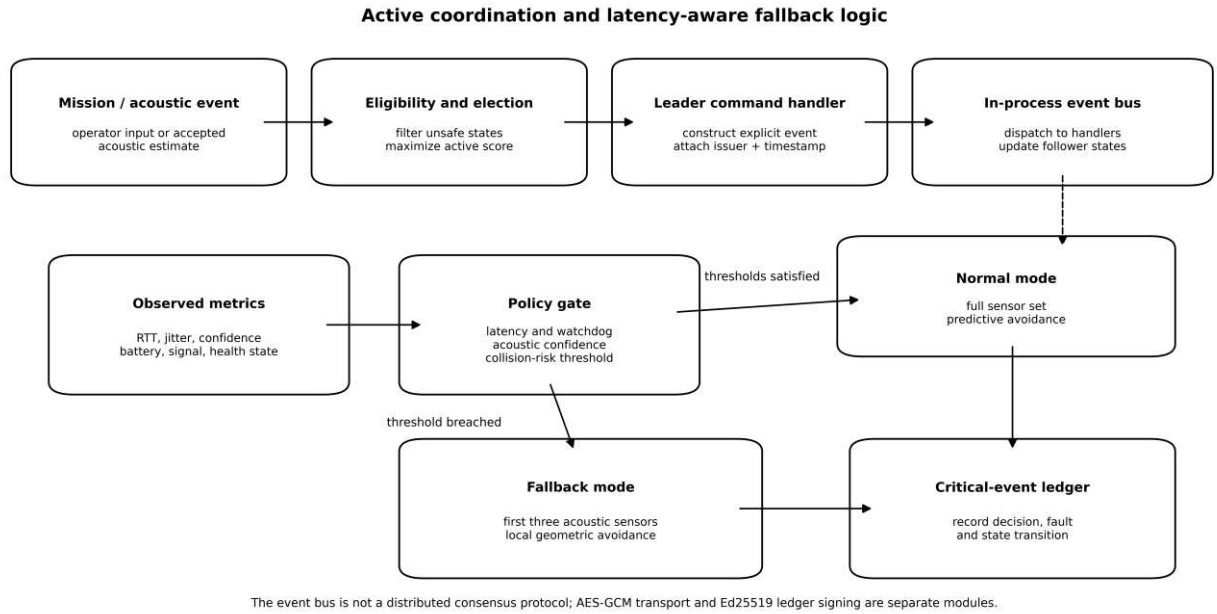



*Fig. 2 Active leader-event and latency-aware fallback paths. AES-GCM transport and Ed25519 ledger signing remain separate implementation modules.*

## 4.3 Acoustic propagation and TDOA observation model

For source $\mathbf{s} = [x_s, y_s]^T$ and sensor $i$ at $\mathbf{p}_i = [x_i, y_i]^T$, range is

$$d_i(\mathbf{s}) = \| \mathbf{s} - \mathbf{p}_i \|_2.$$

An arrival-time observation can be written

$$t_i = t_0 + \frac{d_i(\mathbf{s})}{c} + b_i + \varepsilon_i,$$

where $t_0$ is unknown emission time, $b_i$ is sensor-clock bias, and $\varepsilon_i$ is delay-estimation error. Using sensor 1 as reference removes $t_0$ but not differential clock bias:

$$\Delta t_{1j} = t_j - t_1 = \frac{d_j(\mathbf{s}) - d_1(\mathbf{s})}{c} + \big(b_j - b_1\big) + \big(\varepsilon_j - \varepsilon_1\big).$$

The implementation assumes synchronized sensors in the synthetic study, so $b_j - b_1 = 0$. Its range-difference observation is

$$y_j = c\Delta t_{1j}, \qquad j = 2, \dots, M.$$

Integer-lag estimation gives

$$\widehat{\Delta t}_{1j} = \frac{\hat{\ell}_{1j}}{f_s},$$

with quantization step $1/f_s$. A single range-difference sample has half-step magnitude

$$q_r = \frac{c}{2f_s}.$$

For 16, 48, and 96 kHz, $q_r$ is approximately 10.72, 3.57, and 1.79 mm. This is not a direct lower bound on two-dimensional position error because multiple range-difference constraints and geometry transform delay errors into coordinates.

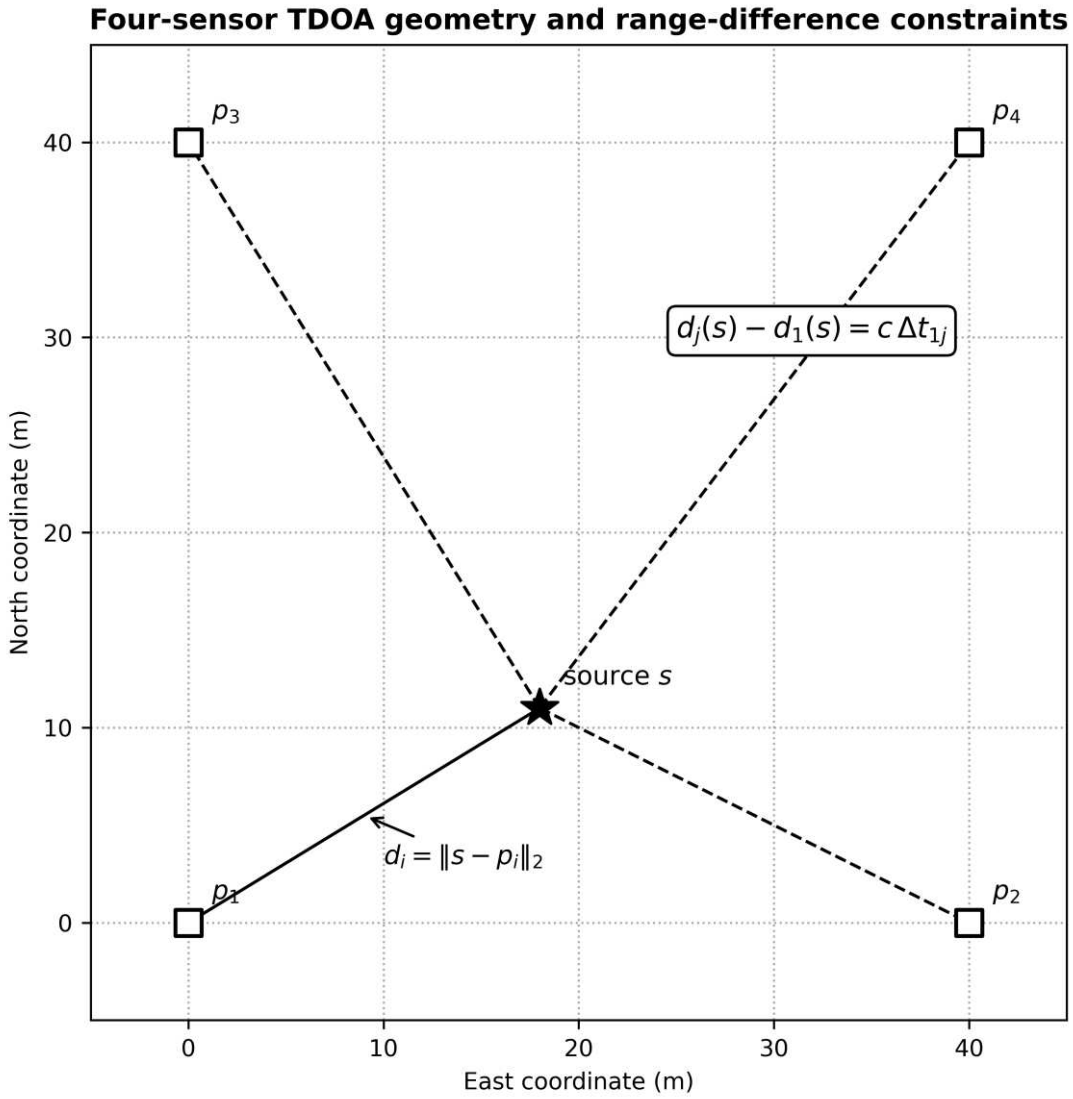



*Fig. 3 Four-sensor geometry and the hyperbolic range-difference constraints used in the synthetic experiments.*

## 4.4 Robust nonlinear least squares and local sensitivity

Define residual for each non-reference sensor:

$$r_j(\mathbf{s}) = d_j(\mathbf{s}) - d_1(\mathbf{s}) - y_j.$$

The active solver approximately computes

$$\hat{\mathbf{s}} = \underset{\mathbf{s}}{\operatorname{argmin}} \sum_{j=2}^{M} \rho\left(r_j(\mathbf{s})\right),$$

where $\rho$ is SciPy's soft-$L_1$ loss. The gradient of one residual is

$$\nabla r_j(\mathbf{s}) = \frac{\mathbf{s} - \mathbf{p}_j}{\| \mathbf{s} - \mathbf{p}_j \|_2} - \frac{\mathbf{s} - \mathbf{p}_1}{\| \mathbf{s} - \mathbf{p}_1 \|_2}.$$

Stacking these row vectors forms $J(\mathbf{s}) \in \mathbb{R}^{(M-1)\times 2}$. A local weighted Gauss-Newton update has the form

$$\Delta\mathbf{s} = -(J^T W J + \lambda I)^{-1} J^T W \mathbf{r},$$

where $W$ contains robust weights and $\lambda \geq 0$ represents damping. When the delay-error covariance is approximately $\Sigma_y$, a first-order position covariance is

$$\Sigma_s \approx \left(J^T \Sigma_y^{-1} J\right)^{-1}.$$

Thus geometry affects error through eigenvalues of $J^T J$. A useful conditioning diagnostic is

$$\kappa_g = \kappa(J^T J) = \frac{\lambda_{max}(J^T J)}{\lambda_{min}(J^T J)}.$$

Large $\kappa_g$ means a small delay perturbation can produce a large coordinate displacement. With four sensors, $J$ has three residual rows and the system is overdetermined. With three sensors, it has two equations for two unknown coordinates. If the solver reaches any branch satisfying both equations, the residual can be nearly zero even when the position is wrong. This explains why residual-only confidence can fail after sensor reduction.

The implemented confidence is

$$q_r = \frac{1}{1 + \mathrm{RMSE}(\mathbf{r})/6}.$$

A safer decision confidence should include fit, geometry, and correlation ambiguity, for example

$$q_{safe} = q_r \exp\left[-\gamma \log\left(1 + \kappa_g\right)\right] q_{peak}\, q_{subset},$$

where $q_{peak}$ measures correlation-peak separation and $q_{subset}$ measures consistency across sensor subsets. This expression is proposed for future work and is not claimed as implemented.

## 4.5 Dynamic-obstacle prediction and collision geometry

For obstacle state $(\mathbf{p}_o, \mathbf{v}_o, \mathbf{a}_o)$, the active predictor uses

$$\mathbf{p}_o(t) = \mathbf{p}_o + \mathbf{v}_o t + 1/2\, \mathbf{a}_o t^2, \qquad 0 < t \leq H.$$

Under constant relative velocity, define

$$\mathbf{r}_0 = \mathbf{p}_o - \mathbf{p}_d, \qquad \mathbf{v}_{rel} = \mathbf{v}_o - \mathbf{v}_d.$$

The closest-approach time is

$$t_{ca} = \mathrm{clip}_{[0,H]}\left(-\frac{\mathbf{r}_0^T \mathbf{v}_{rel}}{\| \mathbf{v}_{rel} \|_2^2}\right),$$

and closest distance is $d_{ca} = \| \mathbf{r}_0 + \mathbf{v}_{rel} t_{ca} \|_2$. The implementation also computes a collision-cone half-angle

$$\theta_c = \sin^{-1}\left(\frac{R_o}{\max(R_o + 1, \| \mathbf{r}_0 \|_2)}\right),$$

and returns zero cone risk when the relative direction lies outside the cone or $t_{ca} \notin [0,4]$ s. Inside the cone, risk is proportional to angular closeness and temporal imminence.

For predicted point $k$ at time $t_k$, the active safe distance is

$$d_{safe} = R_o + 8 + 0.25 \| \mathbf{v}_d \|_2,$$

and a pointwise risk term is

$$P_k = \min\left(1, 0.7\left(1 - \frac{d_k}{d_{safe}}\right) + 0.3m_o\right)\max\left(0.1, 1 - \frac{t_k}{H}\right),$$

when $d_k < d_{safe}$, where $m_o$ is recursive motion-risk memory. These coefficients are engineering heuristics, not probabilities estimated from labeled collision data.

## 4.6 Velocity blending, acceleration saturation, and switching

The desired velocity is

$$\mathbf{v}_{des} = \mathbf{v}_{goal} + \mathbf{v}_{avoid}.$$

With smoothing factor $\alpha \in [0,1]$,

$$\mathbf{v}_{blend} = \mathbf{v}_{cur} + \alpha(\mathbf{v}_{des} - \mathbf{v}_{cur}).$$

The implied acceleration is

$$a = \frac{\| \mathbf{v}_{blend} - \mathbf{v}_{cur} \|_2}{\Delta t}.$$

If $a > a_{max}$, the active controller scales the increment:

$$\mathbf{v}_{new} = \mathbf{v}_{cur} + \frac{a_{max}}{a}(\mathbf{v}_{blend} - \mathbf{v}_{cur}).$$

Latency and watchdog state define a switching variable

$$z_L = \mathbb{I}\left[\overline{T}_{rtt} > T_{th} \ \vee \ \text{watchdog timeout}\right],$$

while acoustic sensor reduction uses

$$z_A = \mathbb{I}[T_{rtt} > T_a], \qquad T_a = 280 \text{ ms}.$$

When $z_L = 1$, local geometric avoidance replaces the predictive path. When $z_A = 1$, only the first three acoustic sensors are used. These switches reduce dependence on delayed information but do not provide Lyapunov or reachability guarantees. A formal hybrid-system treatment would require invariant safe sets, dwell-time assumptions, and bounded sensor/actuator delay.

## 4.7 AES-GCM message-cost model

For application payload of $P$ bytes, the binary wire representation is

$$W(P) = P + 16_{IV} + 16_{tag} = P + 32.$$

The relative expansion is

$$\eta_W(P) = \frac{32}{P}.$$

Therefore, binary expansion is 25% at 128 bytes, 6.25% at 512 bytes, 1.56% at 2048 bytes, and 0.39% at 8192 bytes. Actual JSON/network expansion can be larger.

If key derivation cost is $T_K$ and one derived key protects $Q$ messages, amortized per-message cryptographic cost is

$$\overline{T}_{crypto}(P, Q) = T_{enc}(P) + T_{dec}(P) + \frac{T_K}{Q}.$$

The benchmark derives the key once, so the reported cycle time measures primitive processing rather than password-derivation cost per packet.

### 4.8 Signed linked ledger and replication complexity

For block index $k$, timestamp $t_k$, drone identifier $i$, telemetry hash $H_T$, event hash $H_E$, and previous hash $H_{k-1}$,

$$H_k = \mathrm{SHA3-256}(k \parallel t_k \parallel i \parallel H_T \parallel H_E \parallel H_{k-1}),$$

$$\sigma_k = \mathrm{Ed25519.Sign}_{sk_i}(H_k).$$

A peer accepts a block only if the index is next, the previous hash matches, the block hash recomputes, the origin public key exists, and the signature verifies. For sequential replication to $N-1$ peers,

$$T_{rep}(N) = T_{create} + \sum_{j=2}^{N} T_{verify,j} + T_{python},$$

which is $\mathcal{O}(N)$ in the current single-threaded in-memory benchmark. This relation is not a consensus proof: concurrent writers, forks, persistent storage, membership changes, and recovery are not modeled.

## 5 Threat, fault, and implementation model

The prototype considers passive observation, message alteration, duplicate messages during one process lifetime, leader loss, delayed high-level processing, dynamic obstacles, and mutation of linked ledger records. It does not address malicious authorized nodes, compromised swarm keys, Sybil admission, manager-process failure, durable replay state, concurrent ledger forks, or physical sensor spoofing.

**Table 3 Threat and fault disposition**

| Condition | Implemented control | Remaining limitation |
|---|---|---|
| passive observation | AES-GCM encryption | shared-key compromise exposes all members |
| message alteration | GCM authentication tag | authorized malicious sender remains possible |
| duplicate sequence | per-sender monotonic check | state is not restart-persistent |
| tampered ledger fields/signature | recomputation and Ed25519 verification | no global consensus or durable recovery |
| leader failure | state monitoring and reelection | central manager remains a single point of failure |
| high processing latency | watchdog and local fallback | no formal collision-free guarantee |
| acoustic ambiguity | robust least squares | no calibrated geometry-aware uncertainty |
| motor degradation | modeled return/landing response | no physical propulsion validation |

An encrypted payload can still contain a malicious command from an authorized key holder. A signed block proves possession of a signing key but not truth of the original sensor data. A low residual proves fit to selected equations but not uniqueness of the physical source. Security, integrity, estimation confidence, and safety are therefore distinct properties.

## 6 Reproducibility methodology

Code-equivalent experiments were executed on Linux 6.12 x86-64 with five AMD EPYC 9V74 virtual CPU cores, Python 3.13.5, NumPy 2.3.5, SciPy 1.17.0, and `cryptography` 46.0.4. All randomized acoustic trials used master seed 20260730. Timing values are machine-specific microbenchmarks and exclude radio transport, packet loss, storage, and real flight-control load.

Four acoustic sensors were placed at $(0,0)$, $(40,0)$, $(0,40)$, and $(40,40)$ m. Thirty-six sources were sampled uniformly inside $[5,35]^2$. Arrival signals were represented as integer-shifted impulses with additive white Gaussian noise. The noise experiment used 48 kHz and standard deviations 0, 0.01, 0.05, and 0.10. The sample-rate experiment fixed noise at 0.01 and used 16, 32, 44.1, 48, and 96 kHz. The latency experiment compared four sensors at 80 ms RTT with the first three sensors at 400 ms RTT, both at noise 0.05.

For each trial, position error was

$$e = \| \hat{\mathbf{s}} - \mathbf{s} \|_2.$$

Reported statistics include mean, median, 95th percentile, maximum, detection rate, and residual-derived confidence. Confidence calibration was additionally examined using Spearman rank correlation and the count of trials satisfying $q_r > 0.9$ while $e > 5$ m.

Ledger experiments created 120 sequential blocks for each total replica count $N \in \{3,5,10,25\}$. Mutation tests independently altered previous hash, telemetry hash, block hash, or signature in 100 fresh trials per class. AES-GCM used 600 cycles for payloads 128, 512, 2048, and 8192 bytes. Leader selection used 1500 repetitions for candidate counts 5 to 5000.

**Table 4 Experimental matrix**

| Experiment | Variables | Trials / repetitions | Main outputs |
|---|---|---|---|
| Acoustic noise | noise SD 0–0.10 at 48 kHz | 144 | error distribution and confidence |
| Sampling rate | 16–96 kHz at noise 0.01 | 180 | quantization-related error |
| Latency gate | four vs three sensors | 72 | tail error and calibration failure |
| Ledger scaling | 3–25 replicas | 480 blocks | create/sign and replicate latency |
| Tamper rejection | four mutation classes | 400 | rejection proportion |
| AES-GCM | 128–8192 bytes | 2400 cycles | processing and wire expansion |
| Leader election | 5–5000 candidates | 12,000 selections | empirical linear scaling |

# 7 Results

## 7.1 Acoustic error, sampling resolution, and sensor reduction

At 48 kHz, ideal impulses and noise standard deviation 0.01 produced a median error of 0.0018 m. At noise 0.05, the median increased to 1.2075 m and the 95th percentile to 17.8044 m. At noise 0.10, the median and 95th percentile were 12.3028 and 21.0270 m. Detection remained 100% because the solver returned a solution, demonstrating that detection is not equivalent to acceptable position accuracy.

Increasing sample rate reduced idealized integer-lag error: median error fell from 4.17 mm at 16 kHz to 0.77 mm at 96 kHz. The trend is consistent with $q_r = c/(2f_s)$, although favorable multi-sensor geometry can yield coordinate error smaller than a single range-difference half-step.

The four-sensor latency-gate condition produced median error 1.63 m and 95th percentile 17.14 m. The three-sensor fallback produced a much lower median of 0.0035 m but a 95th percentile of 17.29 m. This mixture is mathematically consistent with an exactly determined system: many trials fit a correct branch almost exactly, while some fit a distant branch with similarly small residual.

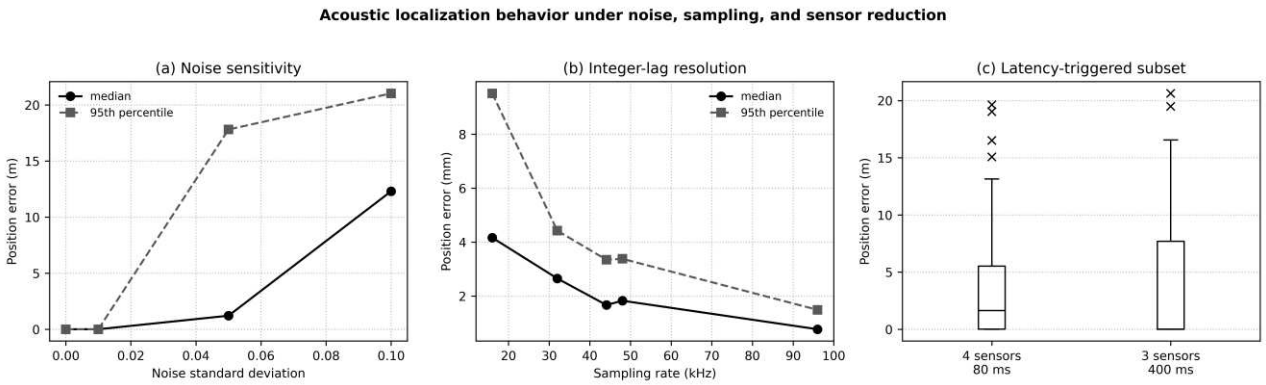



*Fig. 4 Acoustic localization under additive noise, sample-rate variation, and latency-triggered sensor reduction.*

Confidence calibration exposes the failure more directly. For four sensors in the latency experiment, Spearman correlation between error and confidence was $-0.738$, and 2 of 36 trials had confidence above 0.9 with error above 5 m. For the three-sensor fallback, the correlation was $+0.175$ and statistically uninformative, while 11 of 36 trials had confidence above 0.9 with error above 5 m. Mean and median confidence were effectively 1.0 because two residual equations can be fit by two position variables. RQ1 is therefore answered negatively with respect to safe confidence: latency-based sensor reduction is computationally simple but requires geometry-aware uncertainty or subset consistency before controlling a formation.

**Table 5 Acoustic summary**

| Condition | Median error (m) | Mean error (m) | 95th percentile (m) | Median confidence |
|---|---|---|---|---|
| Noise 0.00, 4 sensors | 0.0018 | 0.0019 | 0.0034 | 0.9998 |
| Noise 0.05, 4 sensors | 1.2075 | 4.8203 | 17.8044 | 0.8043 |
| Noise 0.10, 4 sensors | 12.3028 | 12.7614 | 21.0270 | 0.5461 |
| Latency gate off, 4 sensors | 1.6307 | 4.3178 | 17.1426 | 0.8564 |
| Latency gate on, 3 sensors | 0.0035 | 3.7981 | 17.2895 | 1.0000 |

## 7.2 Ledger, AES-GCM, and leader-election scaling

Median block creation and signing remained approximately 41.5–43.6 microseconds because origin work is independent of replica count. Sequential replication grew from 0.223 ms at three total ledgers to 2.671 ms at twenty-five. Linear regression against the number of receiving peers gives

$$\hat{T}_{rep}(N) = 2.44 + 111.20(N-1)\ \mu s,$$

with $R^2 = 0.999998$. The slope agrees with the measured median per-peer verification time of about 111 microseconds. All 100 trials in each mutation class were rejected, supporting the implemented local validation relation but not concurrent-writer consensus.

AES-GCM median encrypt/decrypt time increased from 10.746 microseconds at 128 bytes to 13.570 microseconds at 8192 bytes. One-time PBKDF2 derivation required 12.55 ms. Consequently, key derivation must be amortized across a session rather than repeated for each packet. The 32-byte IV-plus-tag expansion dominates small payloads but becomes negligible for larger messages.

Leader-score selection increased from 3.265 microseconds for five candidates to 9.079 microseconds for 5000. A linear fit gives

$$\hat{T}_{elect}(N) = 3.455 + 0.001143N\ \mu s,$$

with $R^2 = 0.9898$. This validates the expected $\mathcal{O}(N)$ computational behavior. It does not validate distributed agreement, because all candidate information is available inside one manager process.

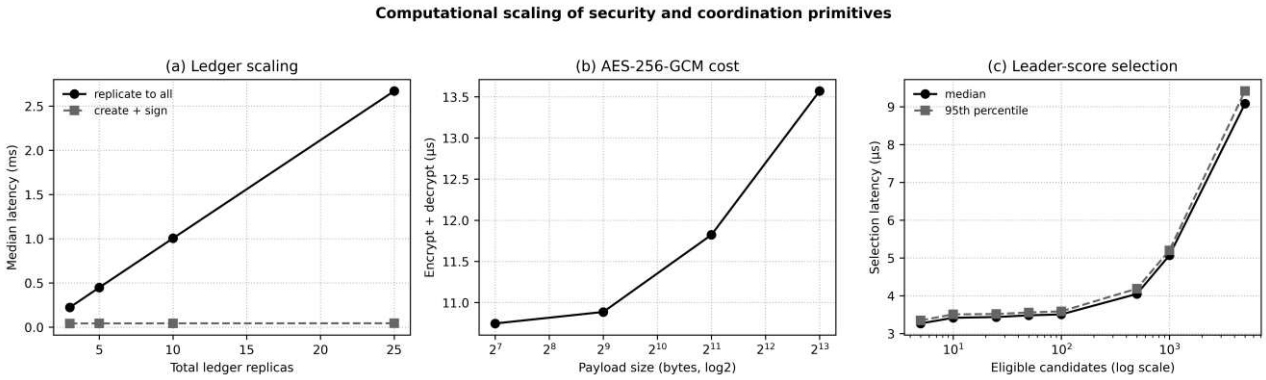



*Fig. 5 Computational scaling of sequential ledger replication, AES-GCM processing, and active leader-score selection.*

**Table 6 Security and coordination benchmark summary**

| Mechanism | Smallest tested scale | Largest tested scale | Main measured relation |
|---|---|---|---|
| Ledger create/sign | 3 replicas: 41.54 µs | 25 replicas: 43.64 µs | approximately constant at origin |
| Ledger replicate all | 3 replicas: 0.223 ms | 25 replicas: 2.671 ms | $2.44 + 111.20(N - 1)$ µs |
| AES-GCM cycle | 128 B: 10.746 µs | 8192 B: 13.570 µs | weak payload-size growth |
| Binary wire expansion | 25.0% at 128 B | 0.39% at 8192 B | $32/P$ |
| Leader selection | 5: 3.265 µs | 5000: 9.079 µs | $3.455 + 0.001143N$ µs |
| Tamper mutations | 4 classes | 100 trials each | 100% rejection in tested classes |

RQ2 is answered at primitive level: cryptographic processing is small relative to typical network and control times, but shared-key management, persistence, and transport were not included. RQ3 is answered computationally but not algorithmically: selection scales linearly and cheaply, while coordination remains centrally orchestrated and lacks partition tolerance, quorum agreement, authenticated membership, and split-brain prevention.

# 8 Discussion and transition to physical validation

The most important systems result is that latency reduction and estimator observability must be considered together. The fallback changes not only communication dependence but also the rank and redundancy of the localization problem. A safe switching policy should therefore depend on both delay and geometry:

$$z_{safe} = \mathbb{I}[T_{rtt} > T_a] \wedge \mathbb{I}\left[\kappa_g < \kappa_{max}\right] \wedge \mathbb{I}[q_{subset} > q_{min}].$$

If a three-sensor subset is ill-conditioned or inconsistent, the safer action may be to hold position, reconfigure sensor geometry, or move one drone to improve observability rather than immediately move the complete formation.

A second result is that low computational cost does not imply strong security or coordination guarantees. AES-GCM can be fast while key lifecycle remains weak. Ed25519 verification can be fast while the ledger lacks consensus and durable storage. Weighted leader selection can be fast while the manager remains a single point of

failure. The mathematical models make these distinctions explicit: primitive cost equations describe performance, whereas security and agreement require separate adversary and fault models.

The next acoustic implementation should estimate sub-sample delay, model clock offsets, and replace the peak-magnitude switch between direct correlation and GCC-PHAT with a statistically interpretable selection or fusion rule. Position uncertainty should be derived from $J^T \Sigma_y^{-1} J$, then validated with bootstrap or Monte Carlo calibration. Robust subset consistency can compare $\hat{\mathbf{s}}_S$ across sensor subsets and reject events when

$$\max_{S_a, S_b} \| \hat{\mathbf{s}}_{S_a} - \hat{\mathbf{s}}_{S_b} \|_2 > \tau_s.$$

For coordination, future work should introduce mission epochs, authenticated membership, persistent monotonic counters, replicated mission state, and a protocol that defines safety under delayed or partitioned communication. The leader score can include motor health only after the variable is calibrated and its reporting integrity is addressed.

**Table 7 Transition gates for physical validation**

| Stage | Mathematical/technical addition | Required exit evidence |
|---|---|---|
| Deterministic software-in-the-loop | scenario seeds, full state logs, confidence intervals | repeated trials and regression baselines |
| Hardware-in-the-loop | measured command and timestamp delays | bounded latency, watchdog, and failsafe response |
| Acoustic bench | synchronized clocks, microphone calibration, rotor-noise spectra | error covariance across SNR, geometry, drift, reverberation |
| Contained 3–5 UAV arena | ground-truth tracking and safety pilot | leader loss, obstacle, acoustic, and communication-fault trials |
| GNSS-degraded field trial | distributed identity, persistent replay state, risk assessment | mission completion, separation, interventions, false/missed events |

RQ4 is therefore answered by a concrete evidence gap: physical claims require calibrated sensing, hardware synchronization, transport measurements, persistent identity and replay protection, durable evidence storage, distributed coordination semantics, and controlled multi-UAV safety trials.

# 9 Limitations and threats to validity

All acoustic results use integer-shifted synthetic impulses and additive white Gaussian noise. They omit rotor spectra, Doppler shift, multipath, reverberation, wind, microphone directivity, receiver motion, clock drift, and packet loss. Millimeter-level ideal results quantify numerical behavior only.

Timing measurements were collected in a virtualized five-core environment and exclude radio transport, sockets in some paths, storage, packet fragmentation, retransmission, and concurrent control workload. Regression coefficients are descriptive for the analysis environment, not hard real-time guarantees.

The active coordination manager centralizes eligibility, selection, command routing, and state. The architecture has multiple agents and replicated ledgers, but it is not a formally decentralized controller. Manager failure and partitions are not recovered.

Security experiments are limited to primitive timing and four mutation classes. They do not test compromised keys, malicious authorized drones, denial of service, side channels, nonce reuse after restart, key extraction, or cryptographic misuse. The fixed PBKDF2 salt and development password are unsuitable for deployment.

The ledger benchmark has one sequential writer and in-memory replicas. It does not exercise concurrent blocks, fork reconciliation, reordering, durable persistence, or recovery. The term ledger is used rather than blockchain consensus.

Finally, no baseline comparison with alternative TDOA solvers, GCC-PHAT variants, consensus protocols, or established swarm frameworks was performed. Future evaluation should use identical scenarios, repeated seeds, uncertainty intervals, and physical ground truth.

# 10 Conclusion

This article rewrote a broad simulation-first drone-swarm implementation as an explicit mathematical system. The resulting formulation clarifies what is optimized, what is measured, and which guarantees are absent. Acoustic localization is a robust nonlinear hyperbolic inverse problem whose confidence depends on geometry as well as residual fit. Dynamic avoidance is a heuristic switching controller with kinematic prediction and acceleration saturation. AES-GCM, Ed25519 ledger verification, and leader-score selection have low primitive-level cost, but their performance does not by itself establish secure identity, consensus, or physical safety.

The central empirical finding is the three-sensor confidence failure: 11 of 36 latency-fallback trials produced confidence above 0.9 with error above 5 m. This result motivates geometry-conditioned uncertainty and subset consistency before acoustic estimates command a formation. The platform is a useful reproducible basis for mathematical and systems research, but secure physical autonomy requires synchronized acoustic hardware, persistent identity and replay protection, distributed coordination semantics, calibrated uncertainty, and staged multi-UAV validation.

# References


[1] Bayindir L (2016) A review of swarm robotics tasks. *Neurocomputing* 172:292–321. https://doi.org/10.1016/j.neucom.2015.05.116

[2] Beard RW, McLain TW, Goodrich MA, Anderson EP (2002) Coordinated target assignment and intercept for unmanned air vehicles. *IEEE Trans Robot Autom* 18(6):911–922. https://doi.org/10.1109/TRA.2002.805653

[3] Bekmezci I, Sahingoz OK, Temel Ş (2013) Flying ad-hoc networks (FANETs): A survey. *Ad Hoc Netw* 11(3):1254–1270. https://doi.org/10.1016/j.adhoc.2012.12.004

[4] Bernstein DJ, Duif N, Lange T, Schwabe P, Yang BY (2012) High-speed high-security signatures. *J Cryptogr Eng* 2:77–89. https://doi.org/10.1007/s13389-012-0027-1

[5] Brandstein M, Ward D (eds) (2001) *Microphone arrays: Signal processing techniques and applications*. Springer, Berlin. https://doi.org/10.1007/978-3-662-04619-7

[6] Brambilla M, Ferrante E, Birattari M, Dorigo M (2013) Swarm robotics: A review from the swarm engineering perspective. *Swarm Intell* 7:1–41. https://doi.org/10.1007/s11721-012-0075-2

[7] Bullo F, Cortés J, Martínez S (2009) *Distributed control of robotic networks*. Princeton University Press, Princeton

[8] Cao YU, Fukunaga AS, Kahng AB (1997) Cooperative mobile robotics: Antecedents and directions. *Auton Robots* 4:7–27. https://doi.org/10.1023/A:1008855018923

[9] Chan YT, Ho KC (1994) A simple and efficient estimator for hyperbolic location. *IEEE Trans Signal Process* 42(8):1905–1915. https://doi.org/10.1109/78.301830

[10] Chung SJ, Paranjape AA, Dames P, Shen S, Kumar V (2018) A survey on aerial swarm robotics. *IEEE Trans Robot* 34(4):837–855. https://doi.org/10.1109/TRO.2018.2857475

[11] Cobos M, Antonacci F, Comanducci L, Sarti A (2017) Frequency-sliding generalized cross-correlation: A sub-band time-delay estimation approach. *IEEE/ACM Trans Audio Speech Lang Process* 25(12):2345–2357. https://doi.org/10.1109/TASLP.2017.2751353

[12] Dworkin MJ (2007) Recommendation for block cipher modes of operation: Galois/Counter Mode (GCM) and GMAC. NIST Special Publication 800-38D. National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.800-38D

[13] Fiorini P, Shiller Z (1998) Motion planning in dynamic environments using velocity obstacles. *Int J Robot Res* 17(7):760–772. https://doi.org/10.1177/027836499801700706

[14] Gupta L, Jain R, Vaszkun G (2016) Survey of important issues in UAV communication networks. *IEEE Commun Surv Tutor* 18(2):1123–1152. https://doi.org/10.1109/COMST.2015.2495297

[15] Hayat S, Yanmaz E, Muzaffar R (2016) Survey on unmanned aerial vehicle networks for civil applications: A communications viewpoint. *Ad Hoc Netw* 40:83–96. https://doi.org/10.1016/j.adhoc.2016.01.004

[16] Islam Shagor MS (2026) *Decentralized coordination and acoustic localization in secure autonomous drone swarms* [Source code]. GitHub. https://github.com/smshagor-dev/decentralized-coordination-and-acoustic-localization-in-secure-autonomousa-drone-swarms

[17] Josefsson S, Liusvaara I (2017) Edwards-curve digital signature algorithm (EdDSA). RFC 8032. Internet Engineering Task Force. https://doi.org/10.17487/RFC8032

[18] Knapp CH, Carter GC (1976) The generalized correlation method for estimation of time delay. *IEEE Trans Acoust Speech Signal Process* 24(4):320–327. https://doi.org/10.1109/TASSP.1976.1162830

[19] Meier L, Tanskanen P, Heng L, Lee GH, Fraundorfer F, Pollefeys M (2012) PIXHAWK: A micro aerial vehicle design for autonomous flight using onboard computer vision. *Auton Robots* 33:21–39. https://doi.org/10.1007/s10514-012-9281-4

[20] National Institute of Standards and Technology (2015) SHA-3 standard: Permutation-based hash and extendable-output functions. FIPS PUB 202. https://doi.org/10.6028/NIST.FIPS.202

[21] National Institute of Standards and Technology (2023) Advanced Encryption Standard (AES). FIPS PUB 197-upd1. https://doi.org/10.6028/NIST.FIPS.197-upd1

[22] Olfati-Saber R (2006) Flocking for multi-agent dynamic systems: Algorithms and theory. *IEEE Trans Autom Control* 51(3):401–420. https://doi.org/10.1109/TAC.2005.864190

[23] Parker LE (1998) ALLIANCE: An architecture for fault-tolerant multirobot cooperation. *IEEE Trans Robot Autom* 14(2):220–240. https://doi.org/10.1109/70.681242

[24] Ren W, Beard RW (2008) *Distributed consensus in multi-vehicle cooperative control*. Springer, London. https://doi.org/10.1007/978-1-84800-015-5

[25] Sahingoz OK (2014) Networking models in flying ad-hoc networks (FANETs): Concepts and challenges. *J Intell Robot Syst* 74:513–527. https://doi.org/10.1007/s10846-013-9959-7

[26] Shakhatreh H, Sawalmeh AH, Al-Fuqaha A, Dou Z, Almaita E, Khalil I, Othman NS, Khreishah A, Guizani M (2019) Unmanned aerial vehicles: A survey on civil applications and key research challenges. *IEEE Access* 7:48572–48634. https://doi.org/10.1109/ACCESS.2019.2909530

[27] Snape J, van den Berg J, Guy SJ, Manocha D (2011) The hybrid reciprocal velocity obstacle. *IEEE Trans Robot* 27(4):696–706. https://doi.org/10.1109/TRO.2011.2120810

[28] van den Berg J, Guy SJ, Lin M, Manocha D (2011) Reciprocal n-body collision avoidance. In: Pradalier C, Siegwart R, Hirzinger G (eds) *Robotics research*. Springer, Berlin, pp 3–19. https://doi.org/10.1007/978-3-642-19457-3_1

# Statements and Declarations

## Funding

The author declares that no funds, grants, or other support were received during the preparation of this manuscript.

## Competing interests

The author has no relevant financial or non-financial interests to disclose.

## Author contributions

M.S.I.S. conceptualized and designed the study; developed the system architecture, software, mathematical models, acoustic localization pipeline, coordination mechanisms, security components, validation methodology, and reproducibility experiments; conducted the investigation, benchmarking, formal analysis, and interpretation of results; prepared all figures and tables; and wrote, reviewed, and edited the manuscript. M.S.I.S. approved the final manuscript and is solely responsible for the integrity and accuracy of the work.

## Data availability

The source code analyzed in this study is publicly available at https://github.com/smshagor-dev/decentralized-coordination-and-acoustic-localization-in-secure-autonomousa-drone-swarms, with the analyzed snapshot identified by commit `5505421886f8305312ba8e98e9a0599357eb36c6`. The code-equivalent reproducibility script, raw CSV files, and figure-generation inputs used for the reported experiments are supplied as supplementary files with this manuscript and are available from the corresponding author. No proprietary, confidential, human-participant, or animal data were used.

## Code availability

The analyzed implementation is distributed through the public repository cited above. A standalone reproduction script for the manuscript's synthetic acoustic, ledger, cryptographic, and election measurements accompanies the submission package.

## Ethics approval and consent to participate

Not applicable. This study reports software analysis and synthetic simulation experiments and does not involve human participants, personal data, animals, or clinical material.

## Consent for publication

Not applicable.

## Use of generative artificial intelligence tools

A generative artificial intelligence tool was used under the sole author's direction to assist with manuscript organization, language drafting, and formatting. The author remains accountable for the study, verified the technical content against the analyzed source code and generated artifacts, revised the manuscript, and approved the final version.