Mutual Information-Based Communication Strategies for UAV Cooperative Odor Source Localization: 3D Simulation Experiments and Parameter Design

Ching-Po Chiu, Jen-Yi Pan, Guan-Yin Lin
Department of Communications Engineering, National Chung Cheng University, Chia-Yi, Taiwan
jypan@ccu.edu.tw

***Abstract*— Multi-UAV odor source localization (OSL) is challenging under partial observability and limited communication. We propose an event-triggered strategy based on Expected Kullback–Leibler (EKL) divergence, where messages are transmitted only when the expected information gain on teammates' beliefs exceeds a threshold. The search is formulated as a Partially Observable Markov Decision Process (POMDP) and combined with Infotaxis. 3D simulations show that EKL-gated communication substantially reduces transmissions and energy consumption while preserving localization performance, yielding higher information efficiency than conventional communication schemes.**

***Keywords*— Multi-UAV, Odor Source Localization, Infotaxis, Expected KL Divergence, POMDP.**

I. INTRODUCTION

Multi-UAV odor source localization (OSL) [1] is critical in disaster scenarios such as toxic gas leaks. However, this task is characterized by high uncertainty, partial observability, and practical communication constraints including bandwidth, energy, and environmental occlusion. To address these challenges, this study proposes an information-theoretic strategy termed Expected Kullback–Leibler Divergence (EKL) [2]-threshold event-triggered communication. Under this mechanism, message transmission is triggered only when an observation yields an expected information gain on teammates' beliefs that exceeds a threshold τ, thereby suppressing unnecessary data exchange and enabling self-adaptive communication rate. To contextualize the benefits of EKL-based triggering, we compare it against commonly used alternatives, including periodic communication, entropy-decrease-based triggering, and no-communication baselines.

The search task is modeled as a Partially Observable Markov Decision Process (POMDP) [3] coupled with the Infotaxis search strategy [4]. The simulation platform integrates gas diffusion and sensing models, Bayesian belief updates [5], and a wireless channel model incorporating path loss and shadowing effects. By systematically varying the threshold τ in three-dimensional simulations, and evaluating all methods under identical sensing, motion, and channel conditions, we analyze the trade-offs among search steps, communication frequency, energy consumption, and belief entropy convergence.

Simulation results demonstrate that, within an appropriate threshold range, EKL-threshold communication significantly reduces transmission counts and energy consumption. Crucially, the strategy maintains localization accuracy comparable to full communication baselines while improving the information efficiency of each transmission, thereby validating information theory as a foundation for adaptive UAV cooperative communication.

The contributions of this paper are as follows: (1) establishing a reproducible multi-UAV OSL simulation framework; (2) proposing an EKL-based information-value metric that drives threshold-triggered communication; (3) validating the proposed strategy under resource-constrained conditions through comparison with multiple communication baselines; and (4) providing a foundation for future extensions, including dynamic thresholds, autonomous coordination, and real-world deployment.

II. SYSTEM MODEL AND PROBLEM FORMULATION

The overall system architecture is illustrated in Fig. 1, where multiple UAVs cooperatively search for an odor source in a 3D environment under limited communication. Each UAV measures local gas concentration, updates its belief via Bayesian inference, and exchanges information through unicast links subject to path loss and shadowing.

The system model is described in the following subsections, including the environment model, sensing and communication models, and the POMDP formulation of the search problem.

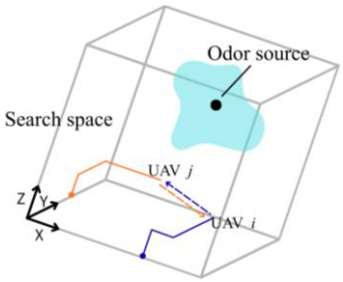

Figure 1. Multi-UAV Odor Source Search System Model. The black solid circle denotes the odor source; the blue and orange paths illustrate the trajectories of UAV-A and UAV-B, respectively; the dashed lines indicate communication links.

A. Background and Bayesian Formulation

OSL aims to enable autonomous agents—such as unmanned aerial vehicles—to infer the spatial location of an odor source

through continuous measurements of gas concentration or odor-particle hit events [6]. This task belongs to a class of dynamic search and decision-making problems characterized by high uncertainty. The inputs to the system are the sensing outcomes and the agent's own position, while the output is a probabilistic estimate of the odor source location. The research objective is to identify the true source with minimal sensing and movement steps, under limited time and energy budgets.

Odor transport is governed by stochastic advection, diffusion, and turbulence, leading to an intermittent and highly non-smooth concentration field. Even near the source, sensors may experience consecutive zero-hit observations, making gradient-based chemotaxis unreliable. To cope with such uncertainty, OSL is commonly formulated as a POMDP. Since the source location is unobservable, the agent maintains a belief distribution $P_t(\boldsymbol{r_s})$, representing the probability that each candidate location is the true source. Upon receiving a new measurement h_t, the belief is updated via Bayes' rule [7]:

$$P_t(\boldsymbol{r_s} \mid h_{1:t}) = \frac{P(h_t \mid \boldsymbol{r_s}) P_{t-1}(\boldsymbol{r_s} \mid h_{1:t-1})}{\sum_{\boldsymbol{r'_s}} P(h_t \mid \boldsymbol{r'_s}) P_{t-1}(\boldsymbol{r'_s} \mid h_{1:t-1})}. \quad (1)$$

In practice, the normalization term is often omitted, yielding the proportional update

$$P_{t+\Delta t}(\mathbf{r}_s) \propto P_t(\mathbf{r}_s) P(h_t \mid \mathbf{r}_s, \mathbf{r}_t), \quad (2)$$

which simplifies the computation of entropy, expected information gain, and EKL divergence.

Overall, Bayesian inference provides the sensing–update–decision loop for OSL. Combined with a Poisson hit model for the likelihood, this framework enables quantitative belief updates and information metrics, forming the basis for information-driven search strategies such as Infotaxis and the proposed communication-triggering mechanism.

B. Environment and Sensing Model

To construct a tractable yet physically grounded search environment, this study adopts a mean-plume approximation with steady-state isotropic diffusion. The odor source is modeled as a continuously emitting point source, and turbulence-induced intermittency is captured through a probabilistic hit model rather than solving the full convection–diffusion equation.

Under three-dimensional isotropic steady-state assumptions, the average concentration at location $\mathbf{r}$ produced by a source at $\mathbf{r}_s$ is expressed as

$$C(\boldsymbol{r} \mid \boldsymbol{r_s}) \propto exp(-|\boldsymbol{r} - \boldsymbol{r_s}|/\lambda)/|\boldsymbol{r} - \boldsymbol{r_s}|, \quad (3)$$

where λ is the characteristic decay length. Assuming that the sensor's hit rate is proportional to the local mean concentration, the expected hit rate is given by

$$\mu(\boldsymbol{r} \mid \boldsymbol{r_s}) = \mu_0 \frac{\lambda}{|\boldsymbol{r} - \boldsymbol{r_s}|} exp(1 - |\boldsymbol{r} - \boldsymbol{r_s}|/\lambda), \quad (4)$$

where μ_0 is the reference hit rate at distance λ.

During each time step Δt, the number of odor-particle hits is modeled as

$$h \sim Poisson(\mu(\boldsymbol{r} \mid \boldsymbol{r_s}) \triangle t). \quad (5)$$

However, the Poisson distribution has unbounded support. To maintain a finite observation space for numerical filtering and EKL evaluation, a truncated Poisson model is adopted [8], where the maximum observed hit category is N_{hits}. All values $h \geq N_{\text{hits}}$ are merged into a single tail bin. The search terminates when a UAV enters the source cell ($\mathbf{r} = \mathbf{r}_s$).

C. Hit-Count Quantization for the Observation Model

To determine a physically meaningful and adaptive truncation level, we exploit the Poisson property that the mean equals the variance. The upper bound is set based on the expected hit rate at the smallest resolvable distance—one grid spacing Δx:

$$\mu_1 \equiv \mu(|\boldsymbol{r} - \boldsymbol{r_s}| = \triangle x) \quad (6)$$

Because Δx defines the minimum spatial increment on the 3D grid, μ_1 automatically reflects the diffusion scale $\lambda/\Delta x$. The hit-count truncation level is chosen as one standard deviation above the mean, with all tail probability merged into a final category:

$$N_{\text{hits}} = \lceil \mu_1 + \sqrt{\mu_1} \rceil + 1 \quad (7)$$

Thus the discrete observation alphabet for actual computation is $h \in \{0,1,\dots,N_{hits} - 1\}$, where the last index corresponds to the aggregated tail. This design captures the dominant Poisson mass while keeping the observation dimension compact.

D. Wireless Channel Model

Communication between UAVs is modeled using a distance-dependent path-loss model with log-normal shadowing. For two UAVs separated by distance d, the average path loss (in dB) is

$$PL(d) = PL(d_0) + 10\gamma \log_{10}(d/d_0) + X_\sigma, \quad (8)$$

where $PL(d_0)$ is the reference loss at distance d_0, γ is the path-loss exponent, and $X_\sigma \sim \mathcal{N}(0, \sigma^2)$ models large-scale shadowing. A packet is lost when $PL(d)$ exceeds the communication budget. This affects only physical-layer success, whereas the decision to transmit is governed by EKL gating.

E. POMDP and Problem Definition

This work formulates OSL as a POMDP and extends it to a cooperative multi-UAV setting. The state is $s = (r_s, r_t)$, with s_Ω denoting the terminal state. The action set $A = \{\text{north, south, east, west, up, down}\}$ determines 3D grid movements. At each step, the UAV receives an observation $o \in$

$\{\Omega, 0,1, \ldots, h_{\max}\}$, where the likelihoods follow the Poisson model in (5). Conceptually, the ideal POMDP observation space would allow $h_{\max} \to \infty$. In practical computation, the implemented maximum hit category equals the truncated bound, $h_{max} = N_{hits} - 1$, as defined in Section C.

Beliefs are updated by Bayes' rule:

$$b_{t+1}(s) = \{P(\, o_t \mid s\,) b_t(s)\} \Big/ \{\sum_{s' \in \mathcal{S}} P(\, o_t \mid s'\,) b_t(s')\}. \quad (9)$$

The uncertainty of the belief is measured by Shannon entropy [10]:

$$H_t = -\sum_{\boldsymbol{r_s} \in S} P_t(\boldsymbol{r_s})\, log\, P_t\,(\boldsymbol{r_s}), \quad (10)$$

and the expected information gain for an action a_t at position r_t is

$$\Delta H(a_t) = H_t - E_{h_{t+1}}[H_{t+1}]. \quad (11)$$

Search strategy is to select actions that maximize the expected information gain [4]:

$$a_t^* = arg\, \max_{a_t} \Delta\, H(a_t). \quad (12)$$

Two Infotaxis-driven UAVs cooperatively localize a single odor source. After a local update, UAV i incorporates successfully received neighbor measurements using distributed Bayesian fusion:

$$P_{t+\Delta t,i}(\boldsymbol{r_s}) = \frac{P_{t,i}(\boldsymbol{r_s}) \prod_{j \in \mathcal{N}_i \cup \{i\}} \left[L_j\left(\, h_j \mid \boldsymbol{r_j}; \boldsymbol{r_s}\,\right)\right]^{u_{j \to i}}}{\sum_{\boldsymbol{r_s'}} P_{t,i}(\boldsymbol{r_s'}) \prod_{j \in \mathcal{N}_i \cup \{i\}} \left[L_j\left(\, h_j \mid \boldsymbol{r_j}; \boldsymbol{r_s'}\,\right)\right]^{u_{j \to i}}}, \quad (13)$$

where $u_{j \to i} \in \{0,1\}$ indicates whether UAV i successfully receives UAV j's packet, determined by (1) EKL ≥ τ and (2) successful wireless transmission.

The search ends when any UAV reaches the source or when step/energy limits are reached.

III. Proposed Method: EKL-Gated Adaptive Communication

A. EKL Formulation & Thresholding

For UAV i, the prior belief is $P_{t,i}(r_s)$. After receiving a local observation h_i, the belief is updated via Bayes' rule:

$$P_{t+\Delta t,i}(\, \boldsymbol{r_s} \mid h_i\,) = \frac{P_{t,i}(\boldsymbol{r_s}) Pr(\, h_i \mid \boldsymbol{r_s}\,)}{\sum_{\boldsymbol{r_s'}} P_{t,i}(\boldsymbol{r_s'}) Pr(\, h_i \mid \boldsymbol{r_s'}\,)}. \quad (14)$$

The information change caused by h_i is measured using the KL divergence:

$$D_{\mathrm{KL}}(P_{t+\Delta t} | P_t) = \sum_{\boldsymbol{r_s}} P_{t+\Delta t}(\, \boldsymbol{r_s} \mid h\,)\, log \frac{P_{t+\Delta t}(\, \boldsymbol{r_s} \mid h\,)}{P_t(\boldsymbol{r_s})}. \quad (15)$$

Since the sender does not yet know the actual observation before transmission, the communication decision uses the expected KL divergence:

$$\mathrm{EKL_{i \to j}} = E_{h_i \sim P(h_i)} \left[D_{\mathrm{KL}} \left(P_{t+\triangle t,j}(\, \boldsymbol{r_s} \mid h_i\,) | P_{t,j}(\boldsymbol{r_s}) \right) \right], \quad (16)$$

Where $P(h_i) = \sum_{\boldsymbol{r_s}} P_{t,i}(\boldsymbol{r_s}) Pr(\, h_i \mid \boldsymbol{r_s}, \boldsymbol{r_i}\,)$. This quantity is equivalent to the mutual information [9] between r_s and h_i, and therefore measures the expected entropy reduction.

A binary threshold rule determines whether UAV i transmits to j:

$$u_{i \to j} = \begin{cases} 1, \mathrm{EKL_{i \to j}} \geq \tau, \\ 0, \mathrm{EKL_{i \to j}} < \tau. \end{cases} \quad (17)$$

Small τ approximates full communication, whereas larger τ suppresses low-value transmissions. Because the optimal τ depends on factors such as odor decay length, geometry, and channel conditions, this work treats τ as a tuning parameter and evaluates its impact on search efficiency and communication cost through simulation.

B. EKL-Gated Communication Algorithm

To reduce communication load, UAV i separately evaluates its EKL contribution to each receiver j. Let $p_r = \hat{p}_{j|i}$ be the sender's cached estimate of j's prior, and let $l_h(r_s \mid r_i)$ be the Poisson likelihood map. The marginal probability of observing h and the corresponding posterior are:

$$p(h) = \langle p_r, \ell_h \rangle, \quad (18)$$

$$p^{(h)}(\mathbf{r_s}) = \frac{p_r(\mathbf{r_s}) \ell_h(\mathbf{r_s} \mid \boldsymbol{r_i}\,)}{p(h)} = normalize(p_r \odot \ell_h). \quad (19)$$

Where $\odot$ denotes element-wise multiplication and $\langle \cdot, \cdot \rangle$ is the inner product over the spatial probability grid. The expected information gain is:

$$\mathrm{EKL_{i \to j}} = \sum_h p(h) D_{\mathrm{KL}}(p_h | p_r). \quad (20)$$

If $\mathrm{EKL_{i \to j}} \geq \tau$, UAV i sends the packet$(\mathbf{r_i}, h_i)$ to j. Upon successful reception, UAV j updates its belief via:

$$p_j' = \mathrm{normalize}\left(p_j \odot L_i(\, h_i \mid \mathbf{r_i}\,) \right). \quad (21)$$

The overall procedure follows a low-complexity gate → sense → fuse single-pass structure.

TABLE I. EKL THRESHOLD-BASED COMMUNICATION ALGORITHM

ALGORITHM: EKL-GATED COMMUNICATION (SINGLE-PASS, PER STEP)

Inputs: sender i at position $\boldsymbol{r_i}$; receiver set J ($\boldsymbol{j \neq i}$) ; threshold τ
State: get_receiver_prior(i,j) // returns i's current estimate of j's prior (cached receiver-prior)
Outputs: $\boldsymbol{D_i}$; // set of receivers selected for transmission at this step
$\boldsymbol{S_i}$; // subset of receivers that successfully received the packet
$\boldsymbol{p_i'}$; // updated posterior of the sender
$\{\boldsymbol{p_j'} \mid \boldsymbol{j} \in \boldsymbol{S_i}\}$ // updated posteriors of the receivers that successfully received the packet

$\boldsymbol{D_i} \leftarrow \emptyset$
for each $\boldsymbol{j} \in$ J **do**
  $\boldsymbol{p_r} \leftarrow$ get_receiver_prior($\boldsymbol{r_i}$,$\boldsymbol{r_j}$)
  $\mathrm{EKL}_{i\to j} \leftarrow 0$
  for $h = 0,1,\ldots, N_{\mathrm{hits}} - 1$ **do**
    $\boldsymbol{\ell_h} \leftarrow$ LikelihoodField($\boldsymbol{h} \mid \boldsymbol{r_i}$)
    $\boldsymbol{p_h} \leftarrow \langle \boldsymbol{p_r}, \boldsymbol{\ell_h} \rangle$
    if $\boldsymbol{p_h} >$ EPS **then**
      $\boldsymbol{post} \leftarrow$ normalize($\boldsymbol{p_r} \odot \boldsymbol{\ell_h}$)
      $\mathrm{EKL}_{i\to j} \leftarrow \mathrm{EKL}_{i\to j} + \boldsymbol{p_h} \cdot D_{KL}(\boldsymbol{post} \mid \boldsymbol{p_r})$
    end if
  end for
  if $\mathrm{EKL}_{i\to j} \geq \tau$ then
    $\boldsymbol{D_i} \leftarrow \boldsymbol{D_i} \cup \{\boldsymbol{j}\}$
  end if
end for

Sender(No retransmission is performed; each SendPacket attempt results in a single success-or-failure outcome.)

$\boldsymbol{h_i} \leftarrow$ SenseAt($\boldsymbol{r_i}$)
$\boldsymbol{p_i'} \leftarrow$ normalize($\boldsymbol{p_i} \odot$ LikelihoodField($h_i \mid \boldsymbol{r_i}$))

Receiver

$\boldsymbol{S_i} \leftarrow \emptyset$
for each j $\in \boldsymbol{D_i}$ do
  success $\leftarrow$ SendPacket(packet($\boldsymbol{r_i}$, h_i) → j)
  if success then
    $\boldsymbol{p_j'} \leftarrow$ normalize($\boldsymbol{p_j} \odot$ LikelihoodField($\boldsymbol{h_i} \mid \boldsymbol{r_i}$))
    $\boldsymbol{S_i} \leftarrow \boldsymbol{S_i} \cup \{\boldsymbol{j}\}$
    UpdateCache($\boldsymbol{j}$, $\boldsymbol{i}$)
  end if
end for
return $\boldsymbol{D_i}$, $\boldsymbol{S_i}$, $\boldsymbol{p_i'}$, $\{\boldsymbol{p_j'} \mid \boldsymbol{j} \in \boldsymbol{S_i}\}$

Note: EPS is a very small constant introduced to ensure numerical stability (typically 10^{-12}).

C. Simulation Flow and Performance Metrics

The system progresses in discrete time steps. In each iteration, the following operations are executed sequentially: action selection (Infotaxis), local measurement and Bayesian update, EKL evaluation and thresholding (unicast), receiver-side fusion and cache update, and termination checking with statistic logging. The mission succeeds when any UAV enters the detection radius of the source; otherwise, it fails upon reaching the maximum step or energy limit.

The framework supports multiple communication modes, including No-comm (no exchange), EKL-off (transmit every step), EKL-on (threshold-gated transmission), as well as periodic and entropy-triggered baselines for comparison. In periodic communication, transmissions occur deterministically every K steps $(t \bmod K = 0)$. In entropy-triggered communication, a UAV transmits when its local belief update yields an entropy reduction exceeding a threshold, $\Delta H_{\mathrm{local}} \geq \tau_{\mathrm{entropy}}$. Unlike EKL gating, this criterion depends solely on local uncertainty reduction at the sender. To quantify the performance and cost under different thresholds τ, the following metrics are used.

The total number of successful communication events (packets) is:

$$N_{\mathrm{comm}} = \sum_{t}^{T} \sum_{i} \sum_{j \neq i} u_{i\to j}(t). \tag{22}$$

Where t is the time index, T is the final step, and $u_{i\to j}(t) \in \{0,1\}$ indicates whether UAV i successfully delivered a packet to UAV j at step t. A transmission is counted when the communication condition of the selected strategy is satisfied and physical-layer delivery succeeds; for EKL-based communication, this additionally requires $\mathrm{EKL}_{i\to j} \geq \tau$. This metric represents the total number of odor-hit packets successfully received during the mission.

To compare communication load across thresholds, the normalized communication usage is defined relative to the full-communication upper bound (EKL-off), which transmits at every step:

$$E_{\mathrm{comm,norm}}(\tau) = N_{\mathrm{comm}}(\tau)/N_{\mathrm{comm}}(\text{EKL-off}). \tag{23}$$

A smaller value indicates greater communication savings. The information efficiency of each packet is:

$$\eta_{\mathrm{pkt}} = (H_0 - H_T)/N_{\mathrm{comm}}, \tag{24}$$

where H_0 and H_T denote the initial and final mean belief entropy (in bits). This measures the average information contribution per successful transmission.

Search efficiency is measured by the average number of steps N_{steps} (in steps), and reliability by the task success rate (percentage of successful runs). As τ increases, fewer packets are transmitted but search time generally increases, reflecting the trade-off between communication load and search performance.

To jointly compare these effects across thresholds, we use the composite cost:

$$J(\tau) = N_{\mathrm{steps}}(\tau) + \alpha E_{\mathrm{comm\,,\,norm}}(\tau), \tag{25}$$

where $\alpha > 0$ controls the trade-off between communication cost and search time. $J(\tau)$ serves as an analytical performance indicator rather than an optimization objective.

IV. Simulation Results and Analysis

A. Platform and Scenarios

Simulations are conducted in a three-dimensional discrete environment to evaluate the proposed EKL-based adaptive communication strategy. The search region is a cubic grid of size $N = 43$ with spacing $\Delta x = 5$ m(215 m per side). The odor source is located at $r_s = [40,40,40]$, and two UAVs start from $r_1 = [10,0,0]$ and $r_2 = [0,10,0]$, approximately 330 m $\approx 11\lambda$ away from the source, representing low-concentration initial conditions. The simulation advances with a timestep $\Delta t = 0.5$ s. The mission succeeds once any UAV reaches the source cell.

Each UAV's initializes its prior $P_{0,i}(x)$ as a uniform distribution over the grid, except that the occupied cell is assigned zero probability. To break symmetry, each UAV performs an initial measurement h_0 at its starting position, after which the likelihood update (using the Poisson model of (5)) generates an asymmetric initial belief.

Although r_1, r_2, and r_s are fixed, trial-to-trial stochasticity arises from two components: (1) the Poisson-distributed sensing process producing random hit counts; and (2) shadowing in the wireless channel (Section II.D), which induces random packet-delivery outcomes. These stochastic elements ensure variability across repeated runs even under identical geometry.

Five communication modes are evaluated:

- EKL-off (full communication): transmit all measurements at every step;
- EKL-on (adaptive communication): transmit only when $\mathrm{EKL}_{i\to j} \geq \tau$;
- Periodic communication: transmit deterministically every K steps ($t \bmod K = 0$);
- Entropy-triggered communication: transmit when the sender's local belief update yields an entropy reduction $\Delta H_{\mathrm{local}} \geq \tau_{\mathrm{entropy}}$;
- No-comm: UAVs search independently.

These models represent, respectively, an upper-bound reference, an information-adaptive strategy, two commonly used heuristic baselines, and a lower-bound baseline. Multiple thresholds τ are scanned for the EKL-on and entropy-triggered modes, and 10 independent runs are performed for each configuration to obtain statistically reliable averages.

The simulation ends when the source is reached, when energy is exhausted, or when 500 steps are completed. Performance is assessed using four metrics: (1) Communication cost, quantified by the total number of successfully delivered packets and the normalized communication-energy ratio defined in (23); (2) Search efficiency, evaluated using the average completion steps N_{steps} and the information-efficiency measure in (24); (3) Composite mission cost, represented by the weighted cost function in (25), where α specifies the relative importance between search time and communication energy. All simulation parameters reflect physically realistic odor dispersion and short-range wireless communication characteristics.

TABLE II. Simulation Parameter Settings

Parameter	Value / Unit
Grid spacing, Scenario size	5 m, 215 m
Time step $\boldsymbol{\Delta t}$	0.5 s
Odor decay length $\boldsymbol{\lambda}$, Detection radius $\boldsymbol{a}$	30 m, 2.5 m
Reference hit rate $\boldsymbol{\mu_0}$	0.77 hits/step
Path-loss exponent $\boldsymbol{\gamma}$	2
Path-loss budget $\boldsymbol{PL_{\mathrm{budget}}}$	90 dB
Shadowing standard deviation $\boldsymbol{\sigma}$	6 dB
Reference distance $\boldsymbol{d_0}$	1 m
Hovering power, Motion energy	100 W, 5 J/m
Sensing energy, Tx/Rx startup energy	0.05 J/step, 1 mJ / 0.5 mJ
Energy per bit, Amplifier constant	5e-8 J/bit, 1e-11 J/bit/m^η
Total UAV energy	5000 J

B. Communication Energy Model and Packet Bit Length

To enable consistent comparison across communication strategies, a simplified but repeatable communication-energy model is used. The model converts the number of packet bits into transmission energy, capturing relative performance trends rather than replicating precise hardware characteristics. For a measurement packet with bit length b transmitted over link distance d, the transmitter and receiver energy costs are

$$E_{\mathrm{tx}} = E_{\mathrm{start,tx}} + E_{\mathrm{elec}} b + \varepsilon_{\mathrm{amp}} b d^{\eta}, \tag{26}$$

$$E_{\mathrm{rx}} = E_{\mathrm{start,rx}} + E_{\mathrm{elec}} b. \tag{27}$$

A failed transmission still incurs full transmit energy (26) but no receive energy (27). The bit length of each measurement packet consists of the header, a 3-D position index, and the discretized hit category derived from the truncated Poisson model (Section II.C). The total number of bits is

$$b_{\mathrm{meas}} = b_{\mathrm{hdr}} + N_{\mathrm{dim}} \lceil \log_2 N \rceil + \lceil \log_2 N_{\mathrm{hits}} \rceil, \tag{28}$$

where $b_{\mathrm{hdr}} = 64$ bits, $N_{\mathrm{dim}} = 3$, and N is the grid size per dimension. The value of N_{hits} is determined by the adaptive truncation rule of (7), which aggregates all Poisson outcomes above the threshold into a single tail category. Because the amplifier term grows approximately linearly with mission duration under the simulated distance scale and path-loss exponent, both packet counts and normalized communication energy are reported.

C. Search Efficiency and Communication Cost

Five communication strategies are evaluated under identical conditions: EKL-on, EKL-off, Periodic, Entropy-triggered, and No-comm. In the Periodic baseline, transmissions occur every $K = 5$ steps. For the Entropy-triggered baseline, a transmission is triggered when the local entropy reduction satisfies $\Delta H_{\mathrm{local}} \geq \tau_{\mathrm{entropy}} = 0.05$. For EKL-on, the threshold is fixed at $\tau = 5 \times 10^{-4}$, corresponding to a moderate communication intensity identified in preliminary scans. All statistics are computed over 10 independent runs.

Experimental results show that the task success rates are 100% for EKL-on, EKL-off, Periodic, and Entropy-triggered strategies, while the No-comm baseline achieves approximately

90%. Table III reports the mean, standard deviation, and 95% confidence intervals of search steps, packet transmissions, and communication energy. Compared with other communication strategies, EKL-on achieves fewer average search steps while requiring fewer packets and lower energy consumption, without degrading search efficiency. As expected, the No-comm strategy exhibits significantly increased search steps due to the absence of information exchange.

TABLE III. PERFORMANCE COMPARISON OF FIVE COMMUNICATION STRATEGIES (MEAN ± STD AND 95% CI)

Method		Steps	Packets	Energy (J)
EKL_ON	mean ± std	280.4 ± 66.0	495.2 ± 73.6	0.73 ± 0.11
	95% CI	[233.2, 327.6]	[442.6, 547.8]	[0.657, 0.808]
EKL_OFF	mean ± std	303.4 ± 54.3	606.8 ± 108.6	0.89 ± 0.16
	95% CI	[264.6, 342.2]	[529.1, 684.5]	[0.776, 1.003]
No-comm	mean ± std	446.3 ± 114.2	-	-
	95% CI	[364.6, 528.0]	-	-
Periodic	mean ± std	336.6 ± 118.1	673.2 ± 236.1	1.03 ± 0.36
	95% CI	[252.2, 421.1]	[504.3, 842.1]	[0.767, 1.284]
Entropy-triggered	mean ± std	304.2 ± 32.6	608.4 ± 65.2	0.93 ± 0.10
	95% CI	[280.9, 327.5]	[561.8, 655.0]	[0.855, 0.997]

D. Dynamics of Entropy and Traffic

Fig. 2 illustrates the evolution of mean belief entropy as a function of normalized search progress. All strategies start with high entropy, indicating substantial initial uncertainty about the source location. As the search proceeds, entropy decreases steadily and converges toward zero.

EKL-on, EKL-off, Periodic, and Entropy-triggered strategies exhibit similar entropy convergence rates without noticeable delay. In contrast, the No-comm strategy shows slower entropy reduction and a higher final entropy level due to insufficient information exchange. These results indicate that EKL-on preserves the convergence quality of full communication while reducing transmission overhead. Since EKL-off represents the theoretical information upper bound, the results imply that EKL-on incurs no significant performance loss rather than outperforming full communication.

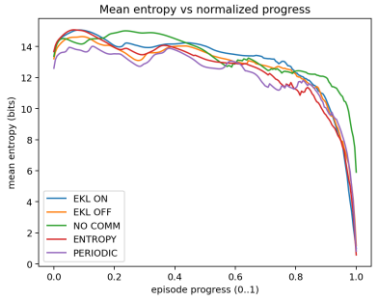

Figure 2. Variation of Average Entropy with Normalized Search Time Progress

Fig. 3 compares the evolution of mean belief entropy as a function of cumulative packet count. All communication strategies exhibit the expected trend that increased communication leads to lower entropy, reflecting the positive role of information exchange in belief concentration. Compared under the same packet budget, EKL-on achieves a faster entropy reduction in the mid-range of cumulative transmissions (approximately 350–520 packets), indicating a higher information contribution per transmitted packet.

EKL-off, Periodic, and Entropy-triggered strategies require a larger number of packets to reach comparable entropy levels, while No-comm does not appear in this plot due to zero transmissions. The oscillatory behavior observed near the end of some curves is not caused by entropy increases, but by reduced averaging samples after early task termination in successful runs.

Overall, the results indicate that EKL-based thresholding improves information efficiency per packet and reduces the communication required to achieve belief convergence, without affecting the final estimation quality.

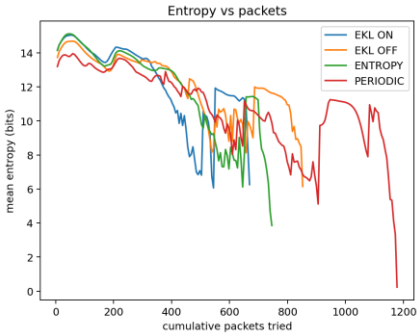

Figure 3. Average Entropy versus Cumulative Packet Count

E. τ-Sensitivity & Design Guidelines

Because varying τ across several orders of magnitude induces multiplicative changes in search behavior, all analyses use $\log_{10}(\tau)$ as the horizontal axis. The threshold sweep is divided into two ranges:

- $[10^{-4}, 10^{-3}]$: from "near-full communication" to the onset of throttling;
- $[10^{-3}, 10^{-2}]$: where throttling becomes significant and search efficiency begins to degrade.

Each range is sampled at ten points (with denser sampling around transition regions). The results in Fig. 4 show a clear nonlinear relationship between τ and both the average search steps and communication usage, forming three behavioral regimes:

1) Low-threshold regime ($\tau \leq 3 \times 10^{-4}$)

Here, both search steps and communication usage differ by less than 2%. This regime effectively corresponds to near-full communication, serving as the lower-bound reference.

2) Transition regime ($3 \times 10^{-4} \leq \tau \leq 8 \times 10^{-3}$)

This is the key trade-off region. As τ increases, low-value transmissions are filtered out, resulting in a substantial communication reduction while search performance remains

largely unaffected. Across this interval: Communication cost drops by 30–40% compared to EKL-off. Search steps increase only slightly (<5%). In our simulations, the best balance—i.e., the "knee point''—occurs near $\tau \approx 6 \times 10^{-3}$, where the weighted cost $J(\tau)$ reaches its minimum.

3) High-threshold regime ($\tau \geq 8 \times 10^{-3}$)

Excessively high thresholds over-filter communication, reducing usage by more than 50% but causing a sharp rise in search steps (10–20%+). This degradation reflects insufficient belief fusion and increasingly inconsistent posteriors across UAVs.

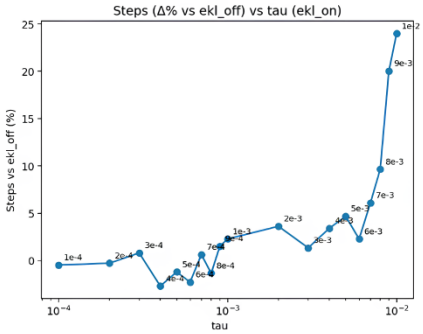

Figure 4. Variation Rate of Average Search Steps under Threshold τ(relative to EKL-off)

To jointly capture search efficiency and communication savings, the weighted cost in (25) is applied to characterize the trade-off between the two performance objectives under different threshold settings.

Fig. 5 shows that all three weight settings ($\alpha = 0.1,0.3,0.5$) exhibit a consistent Pareto-like trade-off: as τ increases, communication usage decreases monotonically while search steps increase, creating a distinct knee point. This knee shifts rightward as more emphasis is placed on energy/bandwidth savings (larger α): (1) $\alpha = 0.1: \tau^{\star} \approx 4 \times 10^{-4}$, (2) $\alpha = 0.3: \tau^{\star} \approx 6 \times 10^{-4}$, and (3) $\alpha = 0.5: \tau^{\star} \approx 6 \times 10^{-3}$. These results offer practical guidance: For time-critical missions (small α): choose $\tau \in [4 \times 10^{-4}, 6 \times 10^{-4}]$ to maintain stable convergence with minimal delay. For energy- or bandwidth-limited missions (large α): $\tau \in [6 \times 10^{-4}, 6 \times 10^{-3}]$ provides substantial communication reduction with acceptable performance loss.

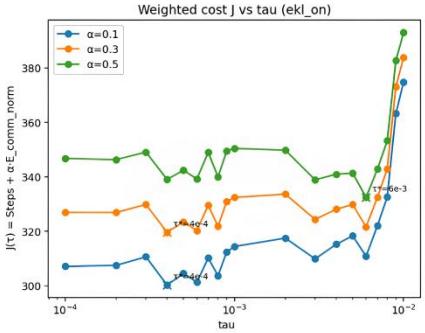

Figure 5. Relationship between Weighted Cost Function and Threshold τ

V. Conclusion

This work investigates an EKL-based information-driven communication strategy for multi-UAV odor source localization under communication constraints. By using expected KL divergence as a transmission trigger, the proposed approach enables selective information sharing while preserving effective cooperative search behavior.

Simulation results confirm that EKL-gated communication achieves a favorable balance between search performance and communication cost, with the threshold τ providing a simple and interpretable control knob. Compared with periodic, entropy-based, and full-communication baselines, the EKL mechanism improves information efficiency without introducing additional coordination complexity.

While this study focuses on a two-UAV setting to isolate fundamental entropy and communication effects, the proposed framework is inherently scalable. Future work will extend the analysis to larger UAV teams to study how communication load, channel conditions, and the effective τ operating region evolve with team size, as well as to incorporate more realistic plume dynamics and networking constraints.

Acknowledgment

This work was supported in part by the National Science and Technology Council (NSTC) of Taiwan under Grants NSTC 114-2221-E-194-034 and 114-2218-E-194-002.

References

[1] J. Monroy, V. Hernandez-Bennetts, H. Fan, A. Lilienthal, and J. Gonzalez-Jimenez, "GADEN: A 3D gas dispersion simulator for mobile robot olfaction in realistic environments," Sensors, vol. 17, no. 7, p. 1479, 2017, doi: 10.3390/s17071479.

[2] S. Kullback and R. A. Leibler, "On information and sufficiency," The annals of mathematical statistics, vol. 22, no. 1, pp. 79-86, 1951.

[3] A. R. Cassandra, "A survey of POMDP applications," in Working notes of AAAI fall symp. on planning with partially observable Markov decision processes, 1998, vol. 1724.

[4] M. Vergassola, E. Villermaux, and B. I. Shraiman, "Infotaxis as a strategy for searching without gradients," Nature, vol. 445, no. 7126, pp. 406-409, 2007.

[5] P. J. Gmytrasiewicz, S. Noh, and T. Kellogg, "Bayesian update of recursive agent models," User Modeling and User-Adapted Interaction, vol. 8, no. 1, pp. 49-69, 1998.

[6] A. T. Hayes, A. Martinoli, and R. M. Goodman, "Distributed odor source localization," IEEE Sensors Journal, vol. 2, no. 3, pp. 260-271, 2002.

[7] P. Morere, R. Marchant, and F. Ramos, "Sequential Bayesian optimization as a POMDP for environment monitoring with UAVs," in Proc. IEEE Int. Conf. Robot. Autom. (ICRA), 2017, pp. 6381-6388.

[8] R. Plackett, "The truncated Poisson distribution," *Biometrics,* vol. 9, no. 4, pp. 485-488, 1953.

[9] R. B. Ash, Information Theory. New York, NY, USA: Courier Corporation, 2012.

[10] J. Olkin, V. Parimi, and B. Williams, "Multi-agent Vulcan: an information-driven multi-agent path finding approach," in Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS), Oct. 2024, pp. 10253-10259, doi: 10.1109/IROS58592.2024.10801571.