AIAA Guidance, Navigation and Control Conference and Exhibit
20 - 23 August 2007, Hilton Head, South Carolina

AIAA 2007-6753

A Multi-UAV Targeting Algorithm for Ensemble Forecast Improvement

Han-Lim Choi* and Jonathan P. How†
Aerospace Controls Laboratory
Massachusetts Institute of Technology

This paper presents an algorithm for targeting a team of UAVs as sensor platforms to ensure the improvement in the forecast at a separate verification time and location within the ensemble forecast framework. The information gain at the verification time/space is used as a metric of the forecast uncertainty reduction, and a computationally efficient backward selection method forms the backbone of the targeting approach. This paper presents the optimal targeting problem, which is complex because it must account for vehicle limitations, such as maximum and minimum flying speed, and the full impact of sensing at a given location. This impact includes a modification of the covariance information used in the targeting, which thus further impacts the future targeting solutions. A cut-off heuristic based on the sparse approximation of the covariance structure is proposed for this full optimization, and it is shown to significantly save computation time by reducing the number of measurement candidates to be consided in payoff calculation. Several further approximations that decompose the computation and decision making into different topologies and choices on the planning horizon are also presented. These are shown to avoid the combinatorial explosion of the computation time, and the optimality of the decomposition schemes is shown to be highly dependent on the existence of agent-wise sharing of the up-to-date covariance information. Numerical simulations using the two-dimensional Lorenz-2003 chaos model are performed to validate the proposed algorithm and compare the suggested decomposition schemes.

I. Introduction

Targeting of mobile sensor networks concerns decision making about when, where, and which type of sensors should be deployed in order to achieve the best information such as the uncertainty reduction of a region of interest[1,2,3] and accurate information about a specified target,[4,5,6] adapted to the dynamic environment. Underlying technological questions in this problem are how to define the information reward in a comprehensive but computationally efficient way, how to determine the best choice fast enough to adapt to changes in the dynamic features of the environment but exhaustive enough to earn reasonable level of optimality, and how to incorporate the dynamic and kinematic constraints related to the motion of the sensor platforms in an effective way. The combinatorial aspects of the selection problem and the limited amount of computational/time resources available usually cause difficulty in balancing all these technological features.

There have been many studies in the sensor networks community regarding to the definition and evaluation of information reward. Zhao et al.[4] introduced the Mahalanobis distance measure

*PhD Candidate, Dept. of Aeronautics and Astronautics, hanlimc@mit.edu
†Professor, Dept. of Aeronautics and Astronautics, jhow@mit.edu

characterized by the location and modality of a sensor in a dynamic query network for target tracking, in order to avoid complicated entropy computation. Ertin et al.[5] showed that the mutual information measure is equivalent to the entropy measure in a distributed tracking sensor network within the Bayesian filtering framework. Wang et al.[6] presented an entropy-based heuristic to approximate the computationally expensive mutual information in the target tracking task, while Guestrin et al.[1] proposed an approximation algorithm based on local kernels for maximizing mutual information in Gaussian processes, and extended it to a randomized algorithm for a graphical model.[3] However, relatively little effort has been devoted to the selection scheme, since reward computations are too expensive in practice to take any better than greedy strategy into account. Reward calculation should be fast enough to adopt a sophisticated selection scheme whose time complexity could be faster than linear in time. Recently, Choi et al.[7] proposed a *backward selection* framework relying on the commutativity of the mutual information to compute the information gain in the context of targeting of sensor networks; it achieves significant savings of computation time by reducing the number of times that the computationally expensive ensemble update should be performed; however, sensor platforms were assumed to have unlimited mobility.

This paper concerns a targeting of multiple UAVs as sensor platforms that have limited capability of movement within the context of the weather prediction problem. Complex characteristics of the weather dynamics such as being chaotic, uncertain, and of multiple time- and length-scales, leads to the necessity of a large sensor network. However, the capability of the static network that takes measurement at the same location every time is constrained by geographic aspects – whether it is on land or over the ocean; the need for an adaptive sensor network incorporated with mobile sensor platforms has been increased. A team of autonomous UAVs equipped with required sensor packages can be an attractive physical realization of the adaptive sensor network, since it is much more cost-effective than conventional crewed aircrafts, and able to be facilitated with substantial amount of payload; its multi-agent feature provides robustness to changes in the environment and to unexpected failures, while its autonomous feature allows its usage in unfavorable weather situations.

Several researches have investigated adaptive targeting of sensor networks for weather forecast enhancement such as singular vector methods by Palmer et al.[8] that determines the most performance-sensitive place based on the linear propagation of the dynamics and ensemble-based[9] approximation of extended Kalman filter to quantify the changes in the future forecast covariance. However, due to the enormous size of the system (the state vector typically has $\mathcal{O}(10^6)$ elements and the size of adaptive network is $\mathcal{O}(10^3)$[9]), only very naive selection strategies, for instance, greedily picking the two flight trajectory out of 49 predefined flight trajectories of a conventional crewed aircraft, have been considered. Thus, the selected solution tend to be far from the optimal; vehicle capability was not effectively considered, since only small number of trajectories were assumed to be available; furthermore, the merit of multi-agent aspects have not been sufficiently reflected.

This paper proposes an efficient algorithm for targeting multiple UAVs that play a role of an adaptive mobile sensor network for the purpose of reducing the uncertainty in a specified verification space/time. Mutual information is introduced to quantify the uncertainty reduction in the random field, while minimum and maximum flying speed of UAV describes the limited mobility of the vehicle. The multi-UAV targeting problem is formulated as a static sensor selection problem with *reachability* constraints, and the *backward selection* algorithm[7] is adopted to form the core of the algorithm computing the mutual information efficiently. In order to reduce the computational waste of resources in computing payoff of suboptimal candidates, cut-off process is introduced accompanied by a simple cost-to-go function based on the sparse assumption of the covariance matrix, that approximates the actual mutual information. Also, a methodology to incorporate the proposed targeting algorithm into the ensemble-based forecast framework is represented. In addition, decomposition schemes that break down a large-scale targeting problem into subproblems by

different configurations of team topology and size of planning horizon, are suggested with discussion of their properties in the context of decentralized decision making. Numerical simulations will be presented to validate the proposed targeting algorithm and compare the approximate schemes with one another.

This paper is organized as follow. Section II briefly introduces the concept of mutual information and the ensemble-based filtering. Section III poses the multiple-UAV targeting problem, which will be addressed by the algorithm described in section IV. Approximation schemes are discussed in section V, and numerical simulations results are given in section VI.

II. Preliminaries

A. Entropy and Mutual Information

Entropy represents the amount of information a random variable can provide if it is revealed. The entropy of a random variable A_1 is defined as[11]

$$h(A_1) = -E\left[\log\left(p_{A_1}(a_1)\right)\right]. \tag{1}$$

The joint entropy of two random variables $h(A_1, A_2)$ is defined in a similar way. The joint entropy and the conditional entropy $h(A_2|A_1)$ are related as

$$h(A_1, A_2) = h(A_1) + h(A_2|A_1). \tag{2}$$

In this work, *mutual information*, which is also called *information gain*, is employed to measure uncertainty reduction of one random variable by another random variable. Mutual information is the amount of information one random variable can reveal about another random variable, which is defined as

$$\mathcal{I}(A_1; A_2) = h(A_1) - h(A_1|A_2). \tag{3}$$

It is noted that mutual information is commutative:[11]

$$\mathcal{I}(A_1; A_2) \equiv \mathcal{I}(A_2; A_1) = h(A_2) - h(A_2|A_1). \tag{4}$$

In the case that a set of random variables are jointly Gaussian, entropy is expressed as

$$h(A) \equiv h(A_1, \cdots, A_k) = \frac{1}{2}\log\left[(2\pi e)^k \det(\mathrm{Cov}(A))\right]. \tag{5}$$

In this work, uncertainty reduction of a random variable by a measurement of another random variable is described by the mutual information between those two, which can be computed commutatively.

B. Ensemble Square Root Filter

This work is based on an ensemble forecast system, in particular, sequential ensemble square root filter[12] (EnSRF). Ensemble forecast represents the nonlinear features of the weather system better, and mitigates computational burden of linearizing the nonlinear dynamics and keeping track of a large covariance matrix.[12,13] In EnSRF, propagation of the state and the covariance matrix amounts to nonlinear integration of ensemble members. The ensemble mean corresponds to the state estimate, and the covariance information can be obtained from the perturbation ensemble as

$$\mathbf{P} = \tilde{\mathbf{X}}\tilde{\mathbf{X}}^T/(L_E - 1), \qquad \tilde{\mathbf{X}} \in \mathbb{R}^{L_S \times L_E} \tag{6}$$

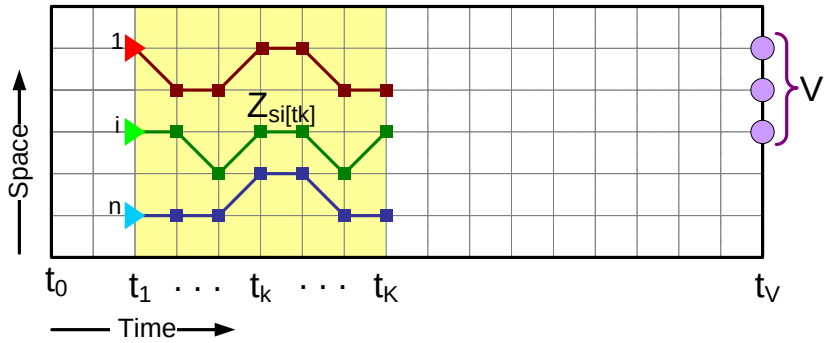

Figure 1. Multi-UAV targeting in the grid space-time

where L_S is the number of state variables and L_E is the ensemble size. $\tilde{\mathbf{X}}$ is the perturbation ensemble defined as

$$\tilde{\mathbf{X}} = \eta\left(\mathbf{X} - \bar{\mathbf{x}} \times \mathbf{1}^T\right) \tag{7}$$

where $\mathbf{X}$ is the ensemble matrix, a row concatenation of each ensemble member, and $\bar{\mathbf{x}}$ is the ensemble mean, the row average of the ensemble matrix. $\eta(\geq 1)$ is the covariance inflation factor introduced to avoid underestimation of the covariance due to the finite ensemble size. The propagation step for EnSRF is the integration

$$\mathbf{X}^f(t+\Delta t) = \int_t^{t+\Delta t} \dot{\mathbf{X}} dt, \qquad \mathbf{X}(t) = \mathbf{X}^a(t), \tag{8}$$

with $\mathbf{X}^f$ and $\mathbf{X}^a$ denoting the forecast and the analysis ensemble, respectively. The measurement update step for EnSRF is

$$\bar{\mathbf{x}}^a = \bar{\mathbf{x}}^f + \mathbf{K}(\mathbf{y} - \mathbf{H}\bar{\mathbf{x}}^f) \tag{9}$$

$$\tilde{\mathbf{X}}^a = (\mathbf{I} - \mathbf{K}\mathbf{H})\tilde{\mathbf{X}}^f \tag{10}$$

where $\mathbf{y}$ denotes the observation vector; $\mathbf{H}$ does the observation matrix under the assumption of linear observation. $\mathbf{K}$ denotes the appropriate Kalman gain, which can be obtained by solving a nonlinear matrix equation described in terms of $\mathbf{X}$.[12] The sequential update process avoids solving a nonlinear matrix equation and provides a faster method for determining $\mathbf{K}$. The ensemble update by the m-th observation is

$$\tilde{\mathbf{X}}^{m+1} = \tilde{\mathbf{X}}^m - \alpha_m \beta_m \mathbf{p}_i^m \tilde{\xi}_i^m, \tag{11}$$

$$\alpha_m = 1/\left(1 + \sqrt{\beta_m R_i}\right), \beta_m = 1/\left(\mathbf{P}_{ii}^m + R_i\right) \tag{12}$$

where measurement is taken for the i-th state variable. $\mathbf{p}_i^m$, $\tilde{\xi}_i^m$, and $\mathbf{P}_{ii}^m$ are the i-th column, the i-th row, and (i,i) element of the prior perturbation matrix $\mathbf{P}^m$, respectively. α_m is the factor for compensating mismatch of the serial update and the batch update, while $\beta_m \mathbf{p}_l^m$ corresponds to the Kalman gain.

III. Problem Statement

This paper deals with targeting of UAVs in a finite grid space-time (fig. 1). The location of a vehicle in the discrete space-time can be represented as a single positive integer regardless of the

spatial dimension; denote $s_i[t_k] \in \mathbb{Z}_+$ the location of the i-th UAV at time t_k. For convenience, the index set is defined such that

$$r = s + \tau \times d_s \tag{13}$$

when r and s represent two points that are spatially identical but temporally apart from each other by τ, positiveness of which represents posteriority; d_s denotes the spatial dimension of the grid representation. Also, an index set $\mathcal{S} \subset \mathbb{Z}_+$ is defined to be the search space-time over which UAVs may be located.

This work assumes that a single state variable (e.g. local temperature, pressure, etc.), which is measured by a sensor of a UAV, is associated with each grid point. Since the dynamics of field state variables is stochastic and sensor measurements are noisy, two random variables X_s and Z_s are introduced to represent the state estimate and the measurement at each location s, respectively. In addition, $X_\mathcal{S}$ and $Z_\mathcal{S}$ denote the random vectors representing the state estimates and measurements of the entire search space $\mathcal{S}$. Furthermore, $\mathcal{V}$ represents the random vector consisting of the state estimates over the verification region at the verification time. With slight abuse of notation, this paper does not distinguish a set of random variables from a random vector formed by the corresponding random variables. The measurement model is written as

$$Z_s = X_s + N_s;\ N_s \sim \mathcal{N}(0, R_s),\ \forall\ s \in \mathcal{S}, \tag{14}$$

$$\mathrm{Cov}(N_s, N_r) = 0, \quad \forall\ s \neq r, \tag{15}$$

$$\mathrm{Cov}(N_s, X_r) = 0, \quad \forall\ s, r \in \mathcal{S}. \tag{16}$$

Measurement is corrupted by a white Gaussian noise that is uncorrelated with any of the state variables and with any other sensing noises.

The objective of targeting is to reduce expected forecast uncertainty in $\mathcal{V}$, with considering limited capability of movent of each UAV. Adopting entropy as a metric of uncertainty, the targeting problem can be posed as decision finding the optimal control sequences $\mathbf{u}[t_k]$'s, or equivalently optimal waypoint sequences $\mathbf{s}[t_k]$'s for the following maximization:

$$\max_{\mathbf{u}[t_1],\cdots,\mathbf{u}[t_K]} \mathcal{I}(\mathcal{V}; Z_{\mathbf{s}[t_1]} \cdots Z_{\mathbf{s}[t_K]}) \triangleq h(\mathcal{V}) - h(\mathcal{V}|Z_{\mathbf{s}[t_1]} \cdots Z_{\mathbf{s}[t_{K-1}]}) \tag{17}$$

subject to

$$\mathbf{s}[t_{k+1}] = \mathbf{s}[t_k] + \mathbf{u}[t_k], \qquad \forall k \in [1, K-1] \cap \mathbb{Z}_+ \tag{18}$$

$$\mathbf{s}[t_1] = \mathbf{s}_o = \text{given} \tag{19}$$

$$s_i[t_k] \in \mathcal{S}, \qquad \forall i \in [1, n] \cap \mathbb{Z}_+, \qquad \forall k \in [1, K-1] \cap \mathbb{Z}_+ \tag{20}$$

$$\mathbf{s}[t_k] \triangleq \begin{bmatrix} s_1[t_k] & \cdots & s_n[t_k] \end{bmatrix}^T, \quad Z_{\mathbf{s}[t_k]} = \bigcup_{i=1}^{n} Z_{s_i[t_k]}, \text{ and} \tag{21}$$

$$\mathbf{u}[t_k] \triangleq \begin{bmatrix} u_1[t_k] & \cdots & u_n[t_k] \end{bmatrix}^T \quad \text{where} \tag{22}$$

$$u_i[t_k] \in \mathcal{U} \subset \mathbb{Z}_+, \quad \forall i \in [1, n] \cap \mathbb{Z}_+, \qquad \forall k \in [1, K-1] \cap \mathbb{Z}_+. \tag{23}$$

n is the number of UAVs and $u_i[t_k]$ is the control action of the i-th UAV at time t_k. The set $\mathcal{U}$ defines all the possible control options that can be represented as positive integer values in the grid space-time. Since a vehicle has limited capability of motion, $\mathcal{U}$ should be a finite set. The condition (18) describes the vehicle dynamics, when initial (at t_1) locations of vehicles are assumed to be given in (32). The condition (20) means that vehicles should remain in the grid space by their control actions - call it *admissibility* condition. With regard to the control option, this work is particularly considering limited mobility of a UAV; thus, an admissible control action leads the vehicle to its neighboring location at the next time step.

Calculation of entropy of a random vector with a general distribution is not trivial; often times requires computation of higher-order moments. However, in case a random vector is jointly Gaussian, only the covariance information is needed to evaluate entropy as given in eq. (5). This work assumes that the random vector representing the entire gridspace is Gaussian; or, more relaxed, at least entropy of any subset of the entire random vector can be computed in a Gaussian way with sufficient accuracy.

IV. Multi-UAV Targeting Algorithm

This section proposes a multi-UAV targeting algorithm for solving a targeting problem defined in the previous section. The proposed targeting algorithm is featured by four main aspects: a) the backward selection formulation to compute payoffs of solution candidates, b) the control space search to effectively incorporate vehicle mobility constraints, c) a cut-off heuristics that reduces the number of solution candidates for which payoff values are actually calculated, and d) an ensemble-based setup with taking fixed observation networks into account. Details of these aspects will be presented followed by a summary of the algorithm.

A. Backward Selection for Gaussian (BSG)

The Backward Selection for Gaussian (BSG) method that was presented in Choi et al.[7] forms the backbone of the targeting algorithm of this work. Think about a static sensor placement problem by regarding the time axis in figure 1 as another space axis and neglecting the vehicle capability constraints. Then, the selection problem can be stated as

$$\mathbf{s}_F{}^\star = \underset{\mathbf{s}\subset\mathcal{S}:|\mathbf{s}|=n_s}{\arg\max}\ \mathcal{I}(\mathcal{V};Z_\mathbf{s}) \triangleq h(\mathcal{V}) - h(\mathcal{V}|Z_\mathbf{s}) \tag{24}$$

$$\overset{G}{=} \underset{\mathbf{s}\subset\mathcal{S}:|\mathbf{s}|=n_s}{\arg\max}\ \tfrac{1}{2}\operatorname{ldet}\left(\operatorname{Cov}(\mathcal{V})\right) - \tfrac{1}{2}\operatorname{ldet}\left(\operatorname{Cov}(\mathcal{V}|Z_\mathbf{s})\right) \tag{25}$$

where n_s is the number of targeting points to select, ldet() stands for $\log\det()$ function, and $\overset{G}{=}$ denotes the equality under the Gaussian assumption. While referring to the above formulation as *forward* selection, Choi et al.[7] verified that computation time for selection decision can be dramatically reduced by reformulating it as the following *backward* selection:

$$\mathbf{s}_B{}^\star = \underset{\mathbf{s}\subset\mathcal{S}:|\mathbf{s}|=n_s}{\arg\max}\ \mathcal{I}(Z_\mathbf{s};\mathcal{V}) \triangleq h(Z_\mathbf{s}) - h(Z_\mathbf{s}|\mathcal{V}) \tag{26}$$

$$\overset{G}{=} \underset{\mathbf{s}\subset\mathcal{S}:|\mathbf{s}|=n_s}{\arg\max}\ \tfrac{1}{2}\operatorname{ldet}\left(\operatorname{Cov}(Z_\mathbf{s})\right) - \tfrac{1}{2}\operatorname{ldet}\left(\operatorname{Cov}(Z_\mathbf{s}|\mathcal{V})\right) \tag{27}$$

$$= \underset{\mathbf{s}\subset\mathcal{S}:|\mathbf{s}|=n_s}{\arg\max}\ \tfrac{1}{2}\operatorname{ldet}\left(\operatorname{Cov}(X_\mathbf{s}) + R_\mathbf{s}\right) - \tfrac{1}{2}\operatorname{ldet}\left(\operatorname{Cov}(X_\mathbf{s}|\mathcal{V}) + R_\mathbf{s}\right). \tag{28}$$

It is noted that $\mathbf{s_F}^\star \equiv \mathbf{s_B}^\star$, since mutual information is commutative.

In finding the optimal solution using explicit enumeration, which results in the worst-case computational complexity, the forward method computes the posterior entropy of $\mathcal{V}$ for every possible measurement candidate $\mathbf{s}$. This means evaluation and determinant calculation of the posterior covariance of $\mathcal{V}$ have to be done total $\binom{|\mathcal{S}|}{n_s}$ times. On the contrary, the covariance update can be done as a constant-time process for the backward method by updating the covariance of the entire search space $\mathcal{S}$ at once; then, a combinatorial search extracts the pair of $n \times n$ submatrices consisting of identical row/column indices from the prior and posterior covariance matrices of $\mathcal{S}$, that reveal the largest difference between their determinants. As the covariance update usually incurs high computational expense – especially for the ensemble-based targeting problems – the

backward selection saves an enormously great deal of computation time. In addition, the backward formulation becomes more preferable if $|\mathcal{V}|$ is bigger than n_s, since it computes the determinants of smaller matrices.

B. Vehicle Capability

In addressing vehicle capability, the reachable set from the current UAV locations is defined as

$$\mathcal{R}(\mathbf{s}[t_k]) = \big\{ \mathbf{r} \subset \mathcal{S} : \mathbf{r} = \mathbf{s}[t_k] + \mathbf{u}[t_k];\ \mathbf{s}[t_k] \subset \mathcal{S}, \mathbf{u}[t_k] \subset \mathcal{U} \big\}, \tag{29}$$

which can be interpreted as the set of all possible location vectors by legitimate control vectors applied at t_k. Incorporating this concept of the reachable set in the BSG framework, the targeting problem can be rewritten as

$$\begin{aligned}(\mathbf{s}[t_2], \cdots, \mathbf{s}[t_K])^\star &= \underset{\mathbf{s}[t_2],\cdots,\mathbf{s}[t_K]}{\arg\max}\ h(Z_{\mathbf{s}[t_1]} \cdots Z_{\mathbf{s}[t_K]}) - h(Z_{\mathbf{s}[t_1]} \cdots Z_{\mathbf{s}[t_K]} | \mathcal{V}) \\ &= \underset{\mathbf{s}[t_2],\cdots,\mathbf{s}[t_K]}{\arg\max}\ \tfrac{1}{2} \operatorname{ldet}\big(\operatorname{Cov}(X_{\mathbf{s}[1:K]}) + R_{\mathbf{s}[1:K]}\big) - \tfrac{1}{2} \operatorname{ldet}\big(\operatorname{Cov}(X_{\mathbf{s}[1:K]}|\mathcal{V}) + R_{\mathbf{s}[1:K]}\big) \end{aligned} \tag{30}$$

subject to

$$\mathbf{s}[t_{k+1}] \in \mathcal{R}(\mathbf{s}[t_k]), \qquad \forall k \in [1, K-1] \cap \mathbb{Z}_+ \tag{31}$$

$$\mathbf{s}[t_1] = \mathbf{s}_o = \text{given}, \tag{32}$$

where $X_{\mathbf{s}[1:K]} \triangleq \bigcup_{k=1}^{K} X_{\mathbf{s}[t_k]}$ and $R_{\mathbf{s}[1:K]} \triangleq \operatorname{diag}(R_{\mathbf{s}[t_k]})$. This formulation corresponds to selecting $n \times (K-1)$ points satisfying the reachability condition whose information gain is maximized by knowledge of $\mathcal{V}$.

In order to incorporate the reachability condition effectively, this work defines the search process based on the following principles: First, the search space should be exhaustive enough to consider all the admissible and reachable candidates but is preferred to be as small as possible. Second, the search process should be able to exclude the infeasible – unreachable or inadmissible – candidates from the decision process. Such a search space $\mathcal{S}_Z \subset \mathcal{S}$ can be obtained by

$$\mathcal{S}_Z = \bigcup_{k=1}^{K-1} \mathcal{R}^k(\mathbf{s}_o), \tag{33}$$

where

$$\mathcal{R}^{k+1}(\mathbf{s}_o) = \mathcal{R}\left(\mathcal{R}^k(\mathbf{s}_o)\right), \qquad \forall k \in [1, K-1] \cap \mathbb{Z}_+. \tag{34}$$

This $\mathcal{S}_Z$ is also minimal in that every element in $\mathcal{S}_Z$ will be referred to at least once in computing the information gains for the feasible candidates. As the right-hand side of (33) is a union of disjoint sets, the cardinality of S_Z becomes

$$|\mathcal{S}_Z| = \sum_{k=1}^{K-1} |\mathcal{R}^k(\mathbf{s}_o)| \le \sum_{k=1}^{K-1} |\mathcal{U}|^{nk} = \frac{|\mathcal{U}|^{nK} - |\mathcal{U}|^n}{|\mathcal{U}|^n - 1} \sim \mathcal{O}(|\mathcal{U}|^{n(K-1)}). \tag{35}$$

It is noted that not every subset of $\mathcal{S}_Z$ of size $n(K-1)$ is a legitimate candidate satisfying the reachability condition; thus, the search process should exclude such illegitimate candidates. For this purpose, this work searches over the control space rather than the index space; the search

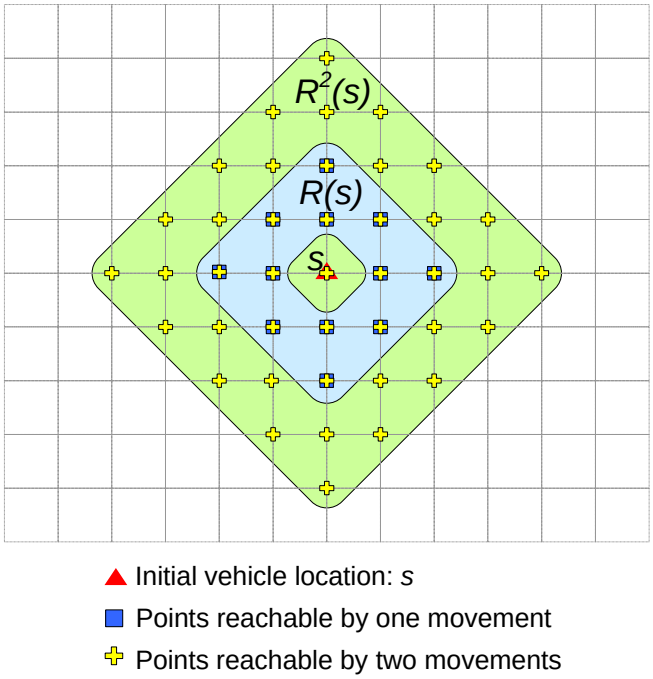

Figure 2. Reachable zone by a single agent in two-dimensional grid space

process addresses decision of

$$
\begin{aligned}
&(\mathbf{u}_Z[t_1], \cdots, \mathbf{u}_Z[t_{K-1}])^\star \\
&= \underset{\mathbf{u}_Z[t_1],\cdots,\mathbf{u}_Z[t_{K-1}]}{\arg\max} \tfrac{1}{2} \operatorname{ldet}\left(\operatorname{Cov}(X_{\mathbf{s}_Z[1:K]}) + R_{\mathbf{s}_Z[1:K]}\right) - \tfrac{1}{2} \operatorname{ldet}\left(\operatorname{Cov}(X_{\mathbf{s}_Z[1:K]} |\ \mathcal{V}) + R_{\mathbf{s}_Z[1:K]}\right) \qquad (36)
\end{aligned}
$$

subject to

$$
\mathbf{s}_Z[t_{k+1}] = \mathbf{s}_Z[t_k] + \mathbf{u}_Z[t_k], \qquad \forall k \in [1, \cdots K-1] \cap \mathbb{Z}_+ \qquad (37)
$$

$$
\mathbf{s}_Z[t_1] = \left[1, \cdots, n\right]^T \qquad (38)
$$

where $\mathbf{s}_Z$ and $\mathbf{u}_Z$ is are the newly defined location index and the control action in $\mathcal{S}_Z$. With this formulation, a close enough upper bound for the number of candidates to consider becomes simply $|\mathcal{U}|^{n(K-1)}$; this is much smaller than $\binom{|\mathcal{S}_Z|}{n(K-1)}$, which could be $\mathcal{O}(|\mathcal{U}|^{n^2(K-1)^2})$ in the worst case.

As to vehicle mobility constraints, this paper is particularly interested in limited flying speed of a UAV described as

$$
V_{min} \leq \mathcal{D}_M(s_i[t_k], s_i[t_{k+1}]) \leq V_{max}. \qquad (39)
$$

$\mathcal{D}_M$ denotes the Manhattan distance defined as

$$
\mathcal{D}_M(s, r) = \sum_{j=1}^{m} |s^j - r^j| \qquad (40)
$$

where m denotes the physical dimension of the space; s^j and r^j denote the j-th physical spatial coordinate values of the gridpoints s and r in $\mathcal{S}$, respectively. Figure 2 illustrates the reachable zone by a single agent from location s in a two-dimensional grid space, in case $V_{min} = 1$ and $V_{max} = 2$. Note in this case that

$$
\mathcal{R}^k(s_o) = \left\{ r \in \mathcal{S}[t_{k+1}] \ : \mathcal{D}_M(s, r) \leq k \times V_{max} \right\}, \qquad \forall k \in [2, K] \cap \mathbb{Z}_+ \qquad (41)
$$

where $\mathcal{S}[t_{k+1}] \subset \mathcal{S}$ denotes the set of gridpoints whose time index is t_{k+1}. Thus, the size of $\mathcal{S}_Z$ for a single agent in a two-dimensional space becomes

$$|\mathcal{S}_Z|_{sg.2d} = \sum_{k=2}^{K-1} \left(1 + 4 \sum_{j=1}^{kV_{max}} j \right) + 4 \sum_{j=V_{min}}^{V_{max}} j \ . \tag{42}$$

In case each UAV has identical moving capability, the reachable zone by a team of UAV is nothing more than the union of each agent's reachable zone:

$$\mathcal{S}_Z(\mathbf{s}_o) = \bigcup_{i=1}^{n} \mathcal{S}_Z(s_i[t_1]). \tag{43}$$

Then, the size of $\mathcal{S}_Z$ for the multi-agent case will be

$$|\mathcal{S}_Z| \leq n \times |\mathcal{S}_Z|_{sg} \tag{44}$$

where equality holds when the reachable zone of each agent is disjoint each other. Note that conflicting assignment is not apparently prohibited with this $\mathcal{S}_Z$; different agents might take measurements at the same location at the same time. However, the optimal targeting decision will never select this option, since it is definitely suboptimal.

C. Cut-Off Heuristics

Although BSG does not involve a combinatorial number of covariance updates, it still requires to compute a determinant of $n(K-1) \times n(K-1)$ matrix (maximum) $|\mathcal{U}|^{n(K-1)}$ times, which is

$$|\mathcal{U}|^{n(K-1)} = \left[4 \sum_{j=V_{min}}^{V_{max}} j \right]^{n(K-1)} = \left[2(V_{max} + V_{min})(V_{max} - V_{min} + 1) \right]^{n(K-1)}, \tag{45}$$

for a two-dimensional problem. Therefore, if V_{max} and/or $n(K-1)$ becomes large, the number of submatrices to be considered in the decision rapidly grows. Moreover, as $n(K-1)$ becomes larger, the unit time for computing the determinant of one submatrix also increases - proportional to $n^3(K-1)^3$ when utilizing Cholesky factorization. Thus, in order to address a large-scale problem, it is necessary to reduce the number of candidates whose mutual information is actually calculated.

For this purpose, this paper proposes a cut-off heuristics that provides indication of which measurement choice is likely to be good or not. For notational simplicity, a static backward selection problem in (28) with simplified notations $\mathrm{Cov}(X_{\mathbf{s}}) \triangleq P_{\mathbf{s}}^{-}$ and $\mathrm{Cov}(X_{\mathbf{s}}|\mathcal{Y}) \triangleq P_{\mathbf{s}}^{+}$, is considered in this discussion; taking vehicle mobility into account in this heuristics will be an easy extension. It was noted in Burer and Lee[14] that for a legitimate covariance matrix $P \succ 0$,

$$\mathrm{ldet}(P) \leq \mathrm{ldet}(P \circ \Sigma) \tag{46}$$

where $\circ$ denotes Hadamard product, and Σ is a symmetric positive definite matrix satisfying

$$\mathrm{diag}(\Sigma) = \mathbf{1}. \tag{47}$$

With the simplest choice of $\Sigma = I$, upper bounds for the prior and posterior entropies for the candidate $\mathbf{s}$ will be

$$h(Z_{\mathbf{s}}) - \tfrac{n_s}{2}\log(2\pi e) = \tfrac{1}{2}\mathrm{ldet}\left(P_{\mathbf{s}}^{-} + R_{\mathbf{s}}\right) \leq \tfrac{1}{2}\textstyle\sum_{i=1}^{n_s} \log\left[P_{\mathbf{s}}^{-}(i,i) + R_{\mathbf{s}}(i,i)\right] \triangleq \hat{h}_{\mathbf{s}}^{-} \tag{48}$$

$$h(Z_{\mathbf{s}}|\mathcal{Y}) - \tfrac{n_s}{2}\log(2\pi e) = \tfrac{1}{2}\mathrm{ldet}\left(P_{\mathbf{s}}^{+} + R_{\mathbf{s}}\right) \leq \tfrac{1}{2}\textstyle\sum_{i=1}^{n_s} \log\left[P_{\mathbf{s}}^{+}(i,i) + R_{\mathbf{s}}(i,i)\right] \triangleq \hat{h}_{\mathbf{s}}^{+}. \tag{49}$$

Note that calculation of $\hat{h}^-_{\mathbf{s}}$ and $\hat{h}^+_{\mathbf{s}}$ is computationally much cheaper than the original entropy computation, since it only requires simple scalar operations. Based on the above upperbounds, this paper suggests to use

$$\hat{\mathcal{I}}_{\mathbf{s}} = \hat{h}^-_{\mathbf{s}} - \hat{h}^+_{\mathbf{s}} \tag{50}$$

as indication of whether the candidate $\mathbf{s}$ is worth being considered as a possibly good solution. Supposed that $\mathbf{s}$ is the m-th candidate in the list, then the cut-off decision for $\mathbf{s}$ is made by

$$c[\mathbf{s}] = \begin{cases} \texttt{pass}, & \text{if } \hat{\mathcal{I}}(\mathbf{s}) \geq \mathcal{I}_{LBD} \\ \texttt{fail}, & \text{otherwise,} \end{cases} \tag{51}$$

where $\mathcal{I}_{LBD}$ is the tightest lowerbound of the optimal information gain thus far. The actual mutual information of candidate $\mathbf{s}$ is computed, only if $c[\mathbf{s}] = \texttt{pass}$.

If $\hat{\mathcal{I}}_{\mathbf{s}} \leq \mathcal{I}(Z_{\mathbf{s}}; \mathcal{V})$ for all $\mathbf{s}$, the optimal solution is guaranteed to lie amongst the `pass`-ed candidates. Unfortunately, this is not the case in general, since mutual information is not submodular for many cases.[1] However, this work experimentally verifies that $\hat{\mathcal{I}}_{\mathbf{s}}$ is a sufficiently good approximation of the corresponding mutual information. Thus, keeping the fundament of the cut-off decision, this paper proposes the following modification of the cut-off indicator function to improve the likelihood of finding the optimal solution:

$$c_\epsilon[\mathbf{s}] = \begin{cases} \texttt{pass}, & \text{if } \hat{\mathcal{I}}(\mathbf{s}) \geq \mathcal{I}_{LBD} - \epsilon \\ \texttt{fail}, & \text{otherwise,} \end{cases} \tag{52}$$

with a positive relaxation parameter ϵ. Thus, every time a candidate $\mathbf{s}$ is on the table, its cost-to-go and cut-off indicator are evaluated; the actual payoff value is computed if it `pass`-es the cut-off criterion. An alternative way of implementating the above idea is to postpone the cut-off decision after calculation of cost-to-go values of all the candidates. By evaluating the actual information gain in the order of the largest cost-to-go heuristics, $\mathcal{I}_{LBD}$ is likely to be tighter; therefore, more candidates will be cut-off. However, this way requires more memory space than the original implementation. In addition, a better approximation of $\mathcal{I}(Z_{\mathbf{s}}; \mathcal{V})$ than $\hat{\mathcal{I}}_{\mathbf{s}}$ can be obtained by utilizing more sophisticated Σ such as block-diagonal and tri-diagonal, while it might incur more computational expense. Also, another type of cost-to-go formula can be derived in the information-theoretic context. (See Appendix A.)

D. Ensemble-Based Formulation

In weather forecast applications, targeted observations are usually not the sole source of information but some fixed - also called routine - observation network is already placed. This routine observation network periodically, usually every 6 hours in practice, takes measurement at the same location. Figure 3 illustrates the observation structure needed to be considered in the targeting problem within the time window from t_0 (decision time) to t_V (verification time). As pointed out in the figure, routine observations after the targeting time is not typically considered in weather applications, because the purpose of targeting is to enhance accuracy of the $\Delta T (= t_V - t_K)$-interval forecast broadcasted at t_K.

To pose a targeting problem as presented in the previous section, two features need to be notified. First, information available at t_0 is the analysis ensemble in which actual measurements up until t_0 have been taken into account. Second, given the existence of the fixed observation network, the targeting problem should be stated conditioned on the routine observations. This work addresses the first aspect by augmenting forecast ensembles for all the relevant time instances as follows.

$$\mathbf{X}^f_{[t_1:t_V]} = \left[\left(\mathbf{X}^f_{t_1}\right)^T \quad \cdots \quad \left(\mathbf{X}^f_{t_K}\right)^T \quad \left(\mathbf{X}^f_{t_V}\right)^T \right]^T, \tag{53}$$

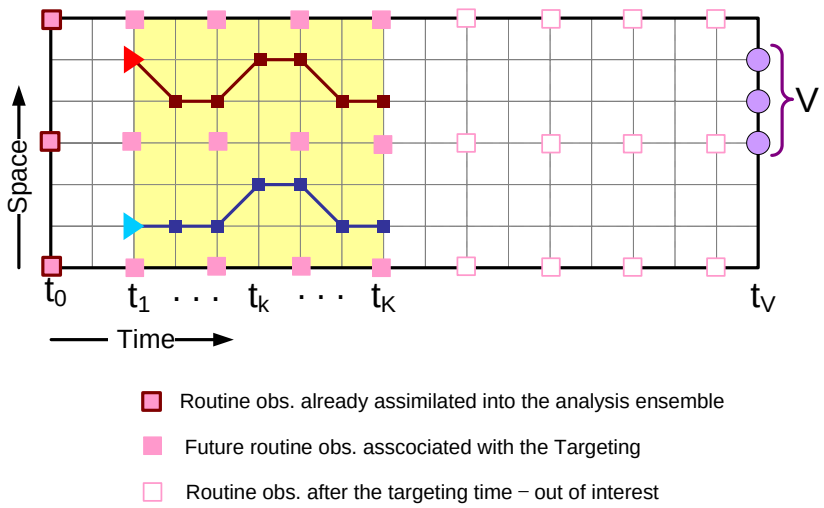

Figure 3. Observation structure over space-time

where

$$\mathbf{X}^f_{t_k} = \mathbf{X}^a_{t_0} + \int_{t_0}^{t_k} \dot{\mathbf{X}}(t)dt \tag{54}$$

and $\mathbf{X}^a_{t_0} \in \mathbb{R}^{L_S \times L_E}$ is the analysis ensemble at t_0. Then, conditioning process with the routine observations from t_1 through t_K becomes nothing more than the ensemble update for the augmented forecast ensemble $\mathbf{X}^f_{[t_1:t_V]}$. Note that an ensemble mean update cannot be (and does not need to be) done, since those fixed observations have not been actually taken and a Gaussian entropy does not require mean information. The ensemble update of the augmented ensemble by the (future) routine observations can be expressed as

$$\tilde{\mathbf{X}}^-_{[t_1:t_V]} = \left(I - \mathbf{K}^r_{[t_1:t_V]}\mathbf{H}^r_{[t_1:t_K]}\right)\tilde{\mathbf{X}}^f_{[t_1:t_V]} \tag{55}$$

where $\mathbf{K}^r_{[t_1:t_K]} \in \mathbb{R}^{L_S(K+1)\times n_r(K+1)}$ and $\mathbf{H}^r_{[t_1:t_K]} \in \mathbb{R}^{n_r(K+1)\times L_S(K+1)}$ are the corresponding Kalman gain and the observation matrix, and n_r is the size of the routine network. In the EnSRF scheme, this process can obviously be performed one observation by one observation. Then, the targeting problem is posed by treating $\mathbf{X}^-_{[t_1:t_K]}$ as the prior ensemble. For instance, the prior state covariance of the entire search space $\mathcal{S}$ can be computed as

$$\mathrm{Cov}(X_{\mathcal{S}}) = \frac{1}{L_E - 1}\tilde{\mathbf{X}}^-_{\mathcal{S}}\left(\tilde{\mathbf{X}}^-_{\mathcal{S}}\right)^T \tag{56}$$

where $\mathbf{X}_{\mathcal{S}}$ consists of the rows of $\mathbf{X}^-_{[t_1:t_K]}$ corresponding to $\mathcal{S}$.

E. Algorithm Summary

Based on the four features discussed above, algorithm 1 represents the overall algorithmic architecture for ensemble-based multi-UAV targeting. The presented one deals with the cheapest implementation that uses the cost-to-go with $\Sigma = I$ and computes the actual objective value for every `pass`-ed candidate. Other types of cost-to-go heuristics and cut-off procedure can easily be incorporated with slight modification of algorithm 1, although all the numerical results in this paper will be based on this version of algorithm.

Algorithm 1 Ensemble-Base Multi-UAV Targeting Algorithm (Memory Saving Implementation)

Initialize $\mathcal{I}_{LBD}$ as default to be zero
Determine $\mathcal{S}_Z$ from $\mathbf{s}_o$, V_{min}, and V_{max}
Extract prior perturbation ensembles corresponding to $\mathcal{S}_Z \cup \mathcal{V}$; denote it as $\mathbf{X}_{\mathcal{S}_Z \cup \mathcal{V}}$
From $\mathbf{X}_{\mathcal{S}_Z \cup \mathcal{V}}$, extract $\mathbf{X}_{\mathcal{S}_Z}$ and compute the posterior ensemble $\mathbf{X}_{\mathcal{S}_Z|\mathcal{V}}$
for all $\mathbf{u}[t_1 : t_{K-1}] \in \mathcal{U}^{n(K-1)}$ **do**
  $\mathbf{s}[t_k] \in \mathcal{S}_Z \ \forall k$? If yes, proceed; otherwise goto step 5.
  Evaluate the prior and posterior covariance matrices, $\mathrm{Cov}(Z_{\mathbf{s}[t_1:t_K]})$ and $\mathrm{Cov}(Z_{\mathbf{s}[t_1:t_K]}|\mathcal{V})$
  Compute the heuristic cost-to-go $\hat{\mathcal{I}}_{\mathbf{s}[t_1:t_K]}$
  if $\hat{\mathcal{I}}_{\mathbf{s}[t_1:t_K]} > \mathcal{I}_{LBD} - \epsilon$ **then**
    Compute $\mathcal{I}(Z_{\mathbf{s}[t_1:t_K]}; \mathcal{V})$
    if $\mathcal{I}(Z_{\mathbf{s}[t_1:t_K]}; \mathcal{V}) > \mathcal{I}_{LBD}$ **then**
      $\mathcal{I}_{LBD} = \mathcal{I}(Z_{\mathbf{s}[t_1:t_K]}; \mathcal{V})$
      $\mathbf{u}_{LBD} = \mathbf{u}[t_1 : t_K]$
    end if
  else
    Goto step 5.
  end if
end for
Return $\mathbf{u}^\star[t_1 : t_{K-1}] = \mathbf{u}_{LBD}$ and $\mathcal{I}^\star(Z_{\mathbf{s}}; \mathcal{V}) = \mathcal{I}_{LBD}$

V. Decomposition Schemes

The proposed cut-off heuristics can reduce the number of computations of matrix determinant; however, it still requires calculation of cost-to-go values combinatorially many times. Thus, a further approximation scheme that breaks down the original problem into a set of small pieces of problems is needed to address a larger-scale problem. This section presents a team-based receding horizon approximation of the original targeting problem. In addition to speeding up the computation, this decomposition provides useful implication on the decentralized decision making for the targeting approach.

Recall that the search process of targeting starts with two covariance matrices - $\mathrm{Cov}(X_{\mathcal{S}_Z})$ and $\mathrm{Cov}(X_{\mathcal{S}_Z}|\mathcal{V})$; the impact of measurements taken at the initial vehicle locations can be incorporated in advance, since initial locations are assumed to be given. Before starting a search process, the following two covariances are available.

$$\mathrm{Cov}(X_{\mathcal{S}_Z}|Z_{\mathbf{s}_o}), \text{ and } \mathrm{Cov}(X_{\mathcal{S}_Z}|\mathcal{V}Z_{\mathbf{s}_o}). \tag{57}$$

The decomposition approach in this section is applied after computing the above prior information.

Figure 4 illustrates the schematics of the problem decomposition in case the number of teams is two. Instead of finding $n(K-1)$ control actions at once, a set of subproblems in each of which T agents having an H-long lookahead window are involved, are considered. Assuming that T divides n and H divides $K-1$ for simplicity, the number of subproblems becomes $\frac{n(K-1)}{TH}$. The approximation scheme starts with solving the first subproblem that determines the optimal control actions of the first T agents until time H - decision making for team 1 for the first planning horizon in figure 4. Afterwards, the second subproblem associated with the next T agents with the same time window is taken into account. As soon as decisions for all the agents within the first time window have been made, then the "execution step," in which covariance information is updated incorporating those decisions, will follow. From the decentralized point of view, the existence of the execution step im-

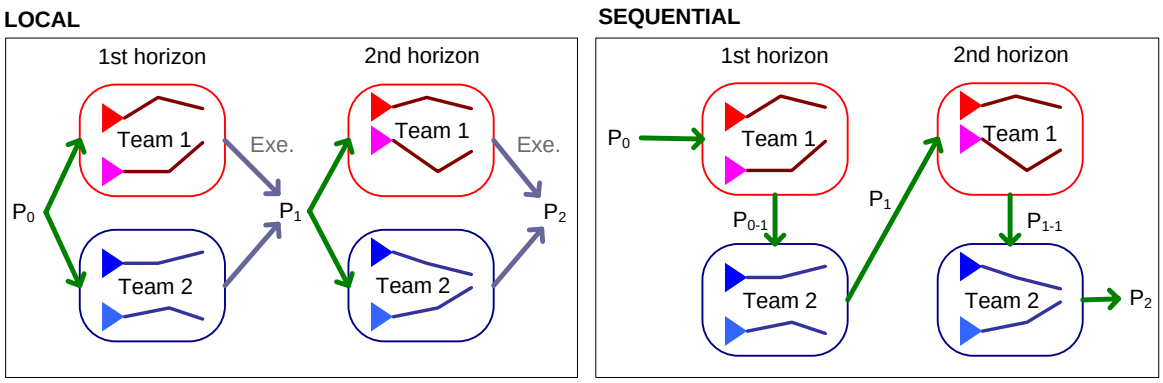

Figure 4. Schematics of team-based decomposition with two different communication topologies

plies that the actual time interval between the first and next planning horizon is sufficiently large to incorporate all the decisions made in the first horizon and to share the updated information. Thus, the $\frac{n}{T}+1$-th subproblem determines the optimal control actions for agents 1 to T for the next time window $[H+1, 2H]$ based on the updated covariances $\mathrm{Cov}(X_{\mathcal{S}_Z}|Z_{\mathbf{s}[1:H]})$, and $\mathrm{Cov}(X_{\mathcal{S}_Z}|\mathcal{V}Z_{\mathbf{s}[1:H]})$. This sequence of decision making and execution continues in the temporal direction until the final planning horizon $[t_{K-H}, t_{K-1}]$.

Back to the transition from the first subproblem to the second subproblem, this work considers two types of transition in this direction: "local" and "sequential" scheme. In the local scheme, a decision of each team is made without any knowledge about the other teams' decisions; every team determines its best action using

$$\mathrm{Cov}(X_{\mathcal{S}_Z}|Z_{\mathbf{s}[1:kH]}), \text{ and } \mathrm{Cov}(X_{\mathcal{S}_Z}|\mathcal{V}Z_{\mathbf{s}[1:kH]}) \tag{58}$$

in the planning window $[kH+1, (k+1)H]$. On the contrary, in the sequential scheme each team knows the decisions of preceding teams and incorporates them into its decision. In other words, a team of agents $(mT+1)$ to $(m+1)T$ will compute its best control decisions with the priors of

$$\mathrm{Cov}(X_{\mathcal{S}_Z}|Z_{\mathbf{s}[1:kH]}Z_{s(1:mT)}), \text{ and } \mathrm{Cov}(X_{\mathcal{S}_Z}|\mathcal{V}Z_{\mathbf{s}[1:kH]}Z_{s(1:mT)}). \tag{59}$$

Figure 4 contrasts the information flows for these two spatial decomposition schemes. In a physical sense, the distinction between the local and the sequential scheme is related to the communication topology amongst agents. The local scheme assumes that only inner-team communication is available in a given decision horizon, while the sequential scheme assumes the existence of inter-team communication even if it may not be real-time. Also, note that conflicting assignment that is not allowed inside the team could be allowed among other teams, in particular, if the local communication scheme is adopted.

VI. Numerical Simulations

A. Lorenz-2003 Model

The Lorenz-2003 model is an extended model of the Lorenz-95 model[15] to address multi-scale feature of the weather dynamics in addition to the basic aspects of the weather motion such as energy dissipation, advection, and external forcing. In this paper, the original one-dimensional model is extended to a two-dimensional one representing the mid-latitude region $(20-70\,\mathrm{deg})$ of

the northern hemisphere. The system equations are

$$\begin{aligned}\dot{y}_{ij} =& -\xi_{i-2\alpha,j}\xi_{i-\alpha,j} + \frac{1}{2\lfloor\alpha/2\rfloor+1}\sum_{k=-\lfloor\alpha/2\rfloor}^{k=+\lfloor\alpha/2\rfloor}\xi_{i-\alpha+k,j}y_{i+k,j} \\ & -\mu\eta_{i,j-2\beta}\eta_{i,j-\beta} + \frac{\mu}{2\lfloor\beta/2\rfloor+1}\sum_{k=-\lfloor\beta/2\rfloor}^{k=+\lfloor\beta/2\rfloor}\eta_{i,j-\beta+k}y_{i,i+k} \\ & -y_{ij} + F\end{aligned} \tag{60}$$

where

$$\xi_{ij} = \frac{1}{2\lfloor\alpha/2\rfloor+1}\sum_{k=-\lfloor\alpha/2\rfloor}^{k=+\lfloor\alpha/2\rfloor} y_{i+k,j}, \qquad i = 1,\ldots,L_{on} \tag{61}$$

$$\eta_{ij} = \frac{1}{2\lfloor\beta/2\rfloor+1}\sum_{k=-\lfloor\beta/2\rfloor}^{k=+\lfloor\alpha/2\rfloor} y_{i,j+k}, \qquad j = 1,\ldots,L_{at} \tag{62}$$

The subscript i denotes the west-to-eastern grid index, while j denotes the south-to-north grid index. The dynamics of the (i,j)-th grid point depends on its longitudinal 2α-interval neighbors (and latitudinal 2β) through the advection terms, on itself by the dissipation term, and on the external forcing ($F = 8$ in this work). In case $\alpha = \beta = 1$, this model is equivalent to the two-dimension Lorenz-95 model.[7] There are $L_{on} = 36\alpha$ longitudinal and $L_{at} = 8\beta + 1$ latitudinal grid points. The dynamics in (60) are subject to cyclic boundary conditions in the longitudinal direction ($y_{i+L_{on},j} = y_{i-L_{on},j} = y_{i,j}$) and a constant advection condition($y_{i,0} = \cdots y_{i,-\lfloor\beta/2\rfloor} = 3$; $y_{i,L_{at}+1} = \cdots = y_{i,L_{at}+\lfloor\beta/2\rfloor} = 0$ in advection terms) is applied in the latitudinal direction, to model the mid-latitude area as an annulus. The length-scale of the Lorenz-2003 is proportional to the inverse of α and β in each direction: for instance, the grid size for $\alpha = \beta = 2$ amounts to 347 km $\times$ 347 km. The time-scale of the Lorenz-2003 system is such that 0.05 time units are equivalent to 6 hours in real-time.

B. Prior Setup

For numerical validation of the proposed targeting algorithm, the initial analysis ensemble for Lorenz-2003 dynamics with $\alpha = \beta = 2$ is constructed. The routine network with size 186 is assumed to be already deployed over the grid space, which is depicted with blue 'o' in figs.5). The static network is dense in two portions of the grid space called lands, while it is sparse in the other two portions of the space called oceans. The routine network takes measurement every 0.05 time unit ($\sim$ 6 hours). The leftmost part consisting of 27 grid points in the right land is selected as the verification region, over which the forecast uncertainty reduction 0.6 time units($\sim$ 3 days) after the targeting time is interested in; it is plotted with red '$\square$' in the figures. The planning window $[t_1, t_K]$ is from 0.025 time unit($\sim$ 3 hours) through 0.375 time unit($\sim$ 15 hours) after t_0 with $K = 5$; time steps are equally spaced by 0.025 time unit($\sim$ 3 hours). The analysis ensemble at t_0 with size 1224, which is same as the size of the grid space, is obtained by running an EnSRF incorporating the routine observations for 1000 cycles. In the upper plot of figure 5 shows a typical shape of the error variance field with this routine configuration. The measurement noise variance is 0.2^2 for routines and 0.02^2 for additional observations, assuming that high-resolution sensors will be equipped on the UAVs. Initial locations of the UAVs are selected based on the climatological data (see Appendix B) such that climatologically best points will be inside the reachable set of the initial locations. Flight speed of the UAV is limited as $V_{min} = 1$ grid/timestep and $V_{max} = 2$ grids/timestep.

C. Effect of Cut-Off Heuristics

As pointed out in section IV.C, the cost-to-go heuristics sometimes underestimates the actual mutual information; the relaxing parameter ϵ is required for better optimality. In order to verify the computational effectiveness of the heuristics and to figure out the effect of the relaxing parameter on the solution optimality, the following two-agent targeting problem is considered. Two UAVs are initially located at $(45, 7)$ and $(45, 9)$, which are in the middle of the eastern part of the wider ocean. A Monte-Carlo simulation with 5 different initial ensembles – five different $\mathbf{X}^a_{t_0}$ – generated from the data assimilation scheme every 1 time unit ($\sim$5 days), is performed.

Table 1 shows the average information gain, the number of candidates whose actual payoffs are computed, and the computation time, with respect to different values of ϵ. It is first noted that the cut-off heuristics enormously reduces the number of candidates to be considered in payoff calculation, while optimality degradation is very small. In fact, for four out of five cases, the heuristic search gives the optimal answer even for the case of $\epsilon = 0$. As $\epsilon = 0.2$ gives a sufficiently close answer to the optimum without substantial increase of computation time, this value of ϵ is used for later simulations. It is also seen that improvement in the actual computation time is relatively small compared with reduction in the number of candidates; this is because the heuristic search requires additional computational resource to evaluate the cost-to-go function for every candidate. This implies that more complicated cost-to-go functions might not be preferred, because it could need longer overhead computation time.

D. Comparison of Decomposition Schemes

In order to discuss the validity of the decomposition schemes, simulation studies varying the values of T and H with using either local or sequential inter-team communication, are conducted. A four-agent problem with the same setup as before, which would take thousands of years to find the solution by an exhaustive search on a usual PC, is considered. The initial locations of the UAVs are $(45, 7)$, $(45, 9)$, $(46, 7)$, and $(46, 9)$ - call them UAV-1 through UAV-4. Three values of $T = 1, 2, 4$ and three values of $H = 1, 2, 4$ are taken into account; in case $T = 1, 2$, both local and sequential communication schemes are dealt with, while $T = 4$ means full spatial cooperation among the UAVs. Thus, total 15 different combinations plus a feasible random strategy are addressed for comparison; for the cases of $(T, H) = (4, 2), (2, 4)$, the cut-off heuristics with $\epsilon = 0.2$ is implemented. Recall that the full optimal solution equivalent to $(T, H) = (4, 4)$ cannot be obtained within reasonable amount of time. Monte-Carlo simulations are performed using 10 different initial ensembles.

Table 2 shows the average optimality of each strategy where superscript 'C' represents the use of the cut-off heuristics. It is first noticeable that the local communication causes significant optimality degradation. All of the three cases of $T = 1(L)$ topology reveal inferior optimality even to the reference random strategy. Furthermore, increase of the planning horizon degrades the optimality for these cases; this can be interpreted as selecting many points based on outdated information

Table 1. 5-simulation average performance of the cut-off heuristics for different ϵ

ϵ	$\mathcal{I}(Z_{\mathbf{s}^\star}; \mathcal{V})$	No. Cand ($\times 10^3$)	Comp. time (s)
0	6.185	454	545
0.1	6.193	782	547
0.2	6.198	1319	551
1.0	6.199	31847	754
∞	6.199	401688	3187

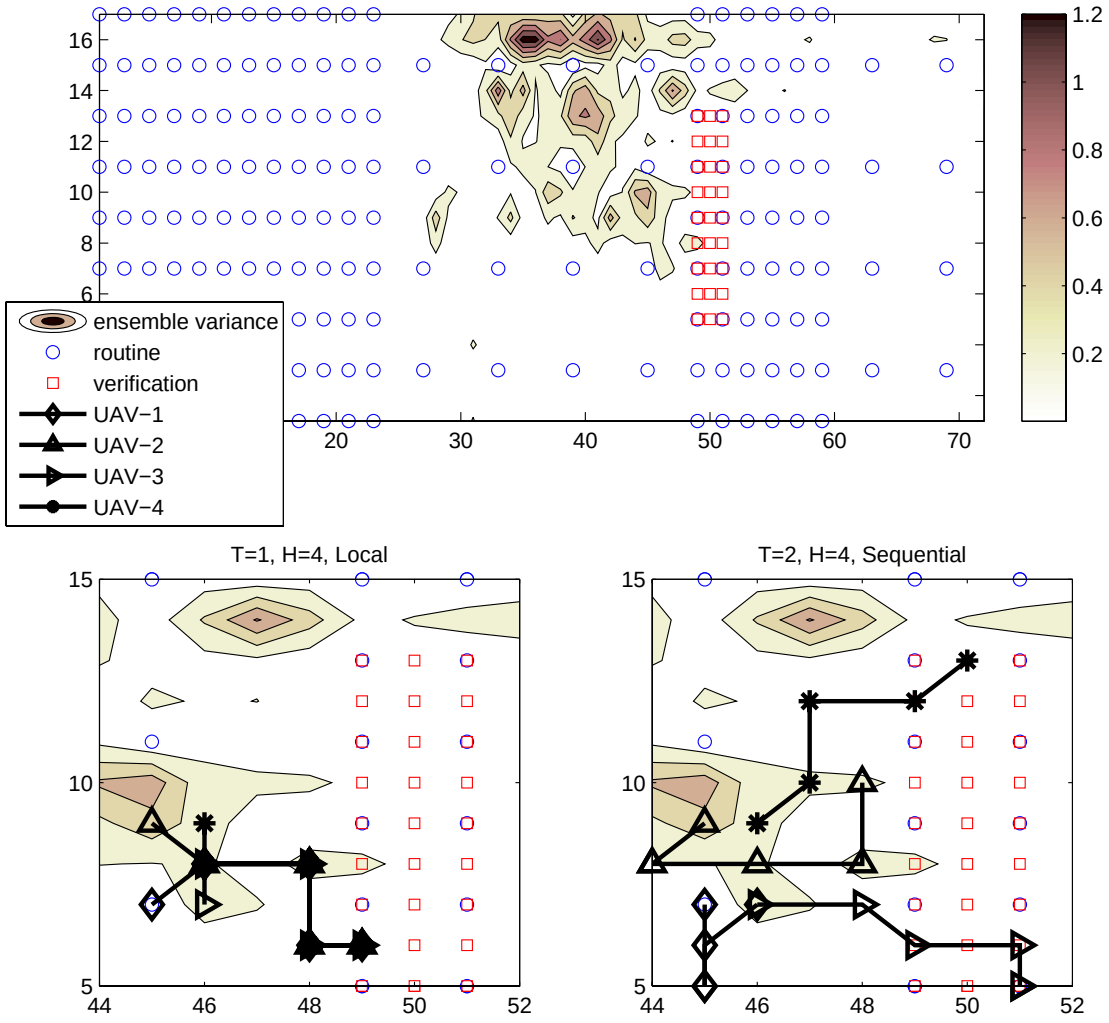

Figure 5. Targeting Solutions with s_o by climatology

is exceedingly suboptimal. Another noticeable comparison is between $T = 2(L)$ and $T = 1(S)$: although each team consists of more UAVs for $T = 2(L)$, the lack of inter-team communication results in poorer optimality. In contrast to the case of local communication, increase of the team size and/or the horizon size enhances optimality for the cases of sequential communication.

Figure 5 shows one exemplary targeting solution. The upper figure depicts the initial analysis ensemble variance field, routine networks, and verification sites. The lower figures contrast two sets of UAV paths corresponding to strategies $T = 1(L), H = 4$ (left) and $T = 2(S), H = 4$ (right). These two setups provide the worst and the best optimality, respectively. In the left picture, due to the lack of inter-agent communication all four UAVs collapse to one trajectory that must be the single-agent optimal (but obviously multi-agent suboptimal) solution, while UAVs are searching over relatively large portion of the space in the right picture.

The average computation time for each strategy is tabulated in Table 3. For every case except the three cases with $TH = 8$, the targeting decision takes just about $3 - 6$ seconds. Computation time for the remaining three cases are all about a thousand seconds. Compared with extremely long expected computation time (about thousands of years) for the full optimization problem, it can be verified that decomposition schemes considerably enhance computational effectiveness.

Table 2. Average value of $\mathcal{I}(Z_{\mathbf{s}^\star};\mathcal{V})$ for 4-agent case ($\mathcal{I}(Z_{\mathbf{s}^{rnd}};\mathcal{V}) = 6.5098$)

	$H=1$	$H=2$	$H=K-1$
$T=1$(L)	6.4419	6.1147	6.0655
$T=1$(S)	9.9142	9.9749	10.2611
$T=2$(L)	7.5253	7.7506	8.2679^C
$T=2$(S)	9.9947	10.1854	10.5615^C
$T=n$	10.0064	10.3363^C	N/A

Table 3. Average computation time (seconds) for 4-agent case

	$H=1$	$H=2$	$H=K-1$
$T=1$(L)	3.8	3.0	3.1
$T=1$(S)	5.9	4.4	4.1
$T=2$(L)	3.9	3.7	1088.0^C
$T=2$(S)	5.3	4.8	1075.0^C
$T=n$	4.7	1001.9^C	N/A

E. Additional Remarks

In the above simulation, the initial sites are selected rather close each other to ensure for climatologically best four points (in space/time) to lie in the reachable set of the UAVs' initial locations. Due to this proximity, UAV paths easily tend to collapse. Thus, additional simulation with more dispersed initial locations is conducted to see the effect of initial distance between agents on the targeting result. Other conditions being unchanged, initially the vehicles are located at $(48, 3)$, $(47, 7)$, $(47, 11)$, and $(47, 15)$. Table 4 and figure 6 show the results. The optimality gap between the local strategies and the sequential strategies becomes smaller in this case, because the chance of conflicting assignment is reduced because of initial dispersion of the agents. However, the qualitative message regarding the communication topology and the size of planning horizon is consistent with that of the previous experiments. Optimality of $T = 1(L)$ cases are decreasing as the size of planning horizon increases, and $T = 1(S)$ still outperforms $T = 2(L)$. Tendency of collapsed paths for the local communication cases still exist as seen in figure 6.

In addition, this work also wants to discuss the influence of the change in the static network configuration. By eliminating six routine observation sites in the oceanic region, the average and maximum squared analysis error are increased approximately eight times. Table 5 represents the optimality of each strategy. It is first noted that information gains are generally smaller than those for the denser fixed network cases; this implies that targeting is of less help when the prior information is not sufficiently good. Also, the optimality gaps amongst strategies decrease as the overall performance levels decrease; but, the qualitative aspects as to the communication topology are still coherent.

VII. Conclusions

This paper presented a targeting algorithm for multiple UAVs with limited mobility, which was incorporated with ensemble-based forecasting and was featured by the backward selection algorithm and the cut-off heuristics for enhancement of computational efficiency. Several decomposition schemes that breaks down a large-scale problem into small pieces of subproblems were proposed with interpretation as the notion of decentralization. Numerical simulations using an

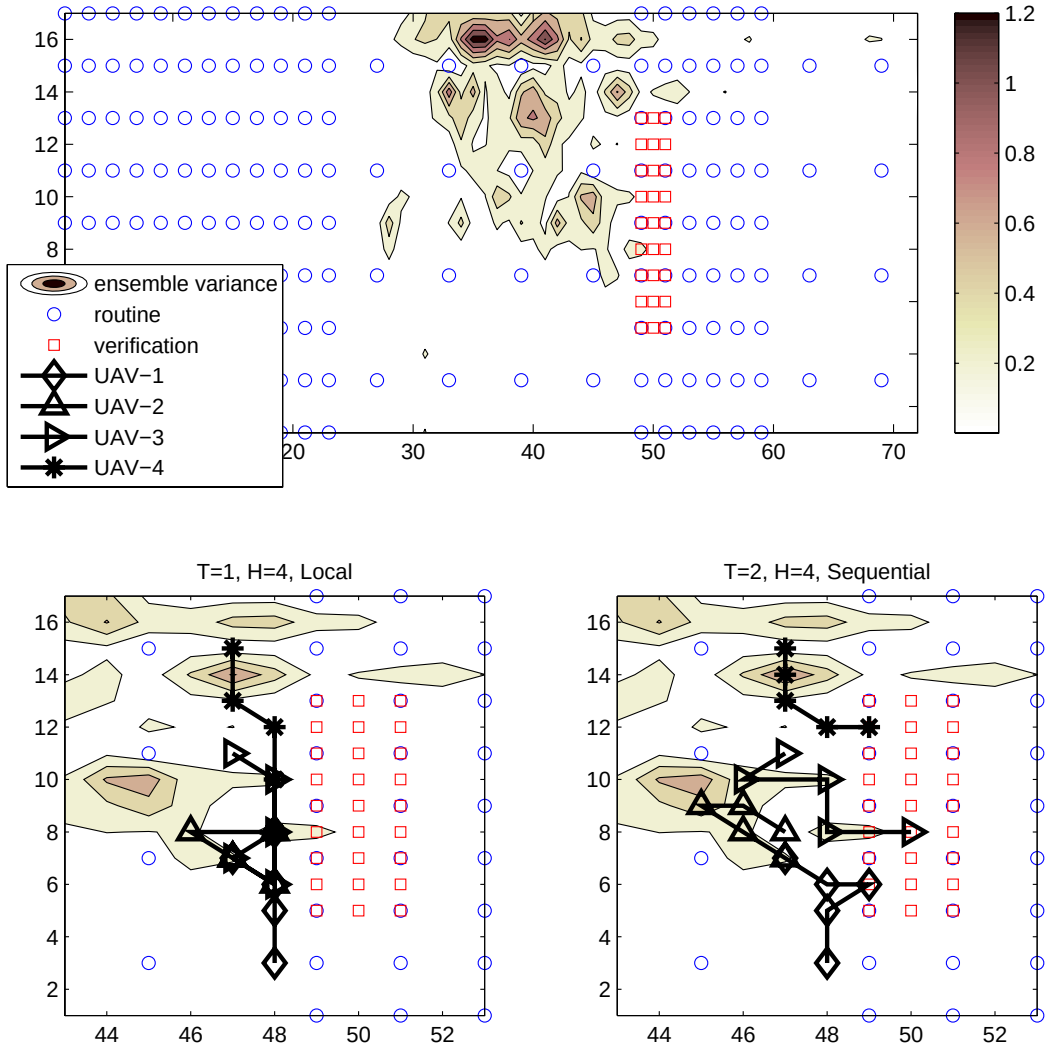

Figure 6. Targeting Solutions with dispersed s_o

simplified weather-like model revealed the validity of the proposed targeting algorithm and quantitatively supported the importance of the information sharing amongst agents. Future research will enhance the computational effectiveness of the proposed algorithm by optimizing the selection of cost-to-go heuristics and cut-off decision process, and will address a realistic weather model such as Coupled Ocean/Atmosphere Mesoscale Prediction System.[16]

Acknowledgment

This work is funded by NSF CNS-0540331 as part of the DDDAS program with Dr. Frederica Darema as the overall program manager.

References

[1]A. Krause and C. Guestrin, "Near-Optimal Nonmyopic Value of Information in Graphical Methods," *21st Conference on Uncertainty in Artificial Intelligence*, Edinburgh, Jul. 2005.

Table 4. Average value of $\mathcal{I}(Z_{\mathbf{s}^\star};\mathcal{V})$ for 4-agent case with dispersed $\mathbf{s}_o$ ($\mathcal{I}(Z_{\mathbf{s}^{rnd}};\mathcal{V}) = 5.7066$)

	$H=1$	$H=2$	$H=K-1$
$T=1$(L)	8.7646	8.5795	7.4195
$T=1$(S)	9.5129	9.7168	9.6681
$T=2$(L)	9.1316	9.2557	9.2071^C
$T=2$(S)	9.5757	9.8560	10.2609^C
$T=n$	9.8511	9.9705^C	N/A

Table 5. Average value of $\mathcal{I}(Z_{\mathbf{s}^\star};\mathcal{V})$ for 4-agent case with $n_r = 180$ ($\mathcal{I}(Z_{\mathbf{s}^{rnd}};\mathcal{V}) = 1.8037$)

	$H=1$	$H=2$	$H=K-1$
$T=1$(L)	3.4065	3.2438	3.1887
$T=1$(S)	3.8024	4.0597	4.2118
$T=2$(L)	3.6649	3.6939	3.9021^C
$T=2$(S)	3.8177	4.0276	4.2499^C
$T=n$	3.9028	4.1575^C	N/A

[2]C. Guestrin, A. Krause, and A. Singh, "Near-Optimal Sensor Placements in Gaussican Processes," *22nd International Conference on Machine Learning*, Bonn, Aug. 2005.

[3]A. Krause, C. Guestrin, A. Gupta, and J. Kleinberg, "Near-Optimal Sensor Placements: Maximizing Information while Minimizing Communication Cost," *5th International Conference on Information Processing in Sensor Networks*, Apr. 2006.

[4]F. Zhao, J. Shin, and J. Reich, "Information-Driven Dynamic Sensor Collaboration," *IEEE Signal Processing Magazine*, Vol.19, No.2, 2002.

[5]E. Ertin, J.W. Fisher, and L.C. Potter, "Maximum Mutual Information Principle for Dynamic Sensor Query Problems," *3rd International Conference on Information Processing in Sensor Networks*, 2003.

[6]H. Wang, K. Yao, G. Pottie, and D. Estrin, "Entropy-Based Sensor Selection Heuristic for Target Localization," *3rd International Conference on Information Processing in Sensor Networks*, Apr. 2004.

[7]H.-L. Choi, J.P. How, and J.A. Hansen, "Ensemble-Based Adaptive Targeting of Mobile Sensor Networks," *American Control Conference*, July 2007.

[8]T.N. Palmer, R. Gelaro, J. Barkmeijer, and R. Buizza, "Singular Vectors, Metrics, and Adaptive Observations," *Journal of the Atmospheric Sciences*, Vol 55, No.4, pp.633-653, 1998.

[9]S.J. Majumdar, C.H. Bishop, B.J. Etherton, and Z. Toth, "Adaptive Sampling with the Ensemble Transform Kalman Filter: Part II Field Programming Implementation," *Monthly Weather Review*, Vol.130, No.3, pp.1356-1369, 2002.

[10]M. Leutbecher, "Adaptive Observations, the Hessian Metric and Singular Vectors," *ESMWF Seminar 2003*, 2003.

[11]T.M. Cover and J.A. Thomas, *Elements of Information Theory*, Wiley Interscience, 1991.

[12]J.S. Whitaker and H.M. Hamill, "Ensemble Data Assimilation without Perturbed Observations," *Monthly Weather Review*, Vol.130, No.7, pp.1913-1924, 2002.

[13]G. Evensen and P.J. van Leeuwen, "Assimilation of Altimeter Data for the Agulhas Current Using the Ensemble Kalman Filter with a Quasigeostrophic Model," *Monthly Weather Review*, Vol.123, No.1, pp.85-96, 1996.

[14]S. Burer and J. Lee, "Solving Maximum-Entropy Samping Problems Using Factored Masks," *Mathematical Programming*, Vol.109, No.2-3, pp.263-281, 2007.

[15]E.N. Lorenz and K.A. Emanuel, "Optimal Sites for Supplementary Weather Observations: Simulation with a Small Model," *Journal of the Atmospheric Sciences*, Vol.55, No.3, pp.399-414, 1998.

[16]R.M. Hodur, "The Naval Research Laboratory's Coupled Ocean/Atmosphere Mesoscale Prediction System (COAMPS)," *Monthly Weather Review*, Vol.125, No.7, pp.1414-1430, July 1997.

[17]Y. Zhang and Q. Ji, "Sensor Selection for Active Information Fusion," *National Conference for Artificial Intelligence Conference (AAAI-05)*, July 2005.

Appendix

A. Information-Theoretic Cost-to-Go Heuristics

A different type of heuristics can be derived in the information-theoretic context. The principle of "information never hurts[11]" says that

$$h(A_1, A_2) = h(A_1) + h(A_2|A_1) \leq h(A_1) + h(A_2) \tag{63}$$

for any random vectors A_1 and A_2. The implication of this inequalities is an upper bound of the entropy of a random vector can be obtained by neglecting some connections between consisting random variables. Note that expression of $\hat{\mathcal{I}}_{\mathbf{s}}$ can also be derived from this principle. If a bit more computation cost is allowed, the following heuristics based on the chain approximation of the graphical model of the information structure, can be used:

$$\hat{\mathcal{I}}^a_{\mathbf{s}} = \mathcal{I}(Z_{s_1}; \mathcal{V}) + \sum_{j=1}^{n_s-1} \mathcal{I}(Z_{s_{j+1}}; \mathcal{V}|Z_{s_j}) = \mathcal{I}(Z_{s_1}; \mathcal{V}) + \sum_{j=1}^{n_s-1} \mathcal{I}(Z_{s_{j+1}} Z_{s_j}; \mathcal{V}) - \mathcal{I}(Z_{s_j}; \mathcal{V}), \tag{64}$$

which was presented in Zhang and Ji[17] (but with incorrect interpretation). In terms of the covariance matrix, this approximation is equivalent to the tri-diagonal assumption of the information matrix, inverse of the covariance matrix. To implement this cost-to-go heuristic, mutual information values for every single-point and double-point selection should be computed and stored in advance, which are not particularly computationally expensive for a reasonable size of the search space. Thus, application of $\hat{\mathcal{I}}^a_{\mathbf{s}}$ makes sense only when n_s is greater than two.

B. Climatology-Based Fixed Observation Selection

The targeting problem proposed in this work determines the best places to take measurement (plus considering vehicle capability) exploiting the prior information based on the up-to-date analysis ensemble. Because targeting solution is adapted to initial ensemble information, different initial ensemble results in different targeting solutions.

Climatology-based observation selection takes a different objective: it is interested in determines good places to take measurement in the long-term sense independent of current ensemble information. Thus, this concept can be useful in determining locations of additional fixed observation, at which observation will be made on a regular basis. Algorithmically, this also can be done using the targeting algorithm presented in this work starting with the climatological ensemble obtained as follow.

$$\tilde{\mathcal{X}}^f_C = \begin{bmatrix} \mathcal{E}_1 & \cdots & \mathcal{E}_m & \cdots \mathcal{E}_{L_C} \end{bmatrix} \tag{65}$$

where $\mathcal{E}_m$ is the forecast error vector m-th data assimilation cycle defined as

$$\mathcal{E}_m = \begin{bmatrix} \bar{\mathbf{X}}^f(t_1) - \mathbf{x}^t(t_1) \\ \bar{\mathbf{X}}^f(t_2) - \mathbf{x}^t(t_2) \\ \vdots \\ \bar{\mathbf{X}}^f(t_K) - \mathbf{x}^t(t_K) \\ \bar{\mathbf{X}}^f(t_V) - \mathbf{x}^t(t_V) \end{bmatrix}_m \tag{66}$$

$$\bar{\mathbf{X}}^f(t_k) = \int_{t_0}^{t_k} \dot{\mathbf{X}} dt, \qquad \mathbf{X}(t_0) = \mathbf{X}^a(t_0). \tag{67}$$

The climatology-based solution can be obtained by replacing $\mathbf{X}^f_{[t_1:t_V]}$ by $\tilde{X}^f_C$ in the targeting algorithm.