Distributed Pursuit–Evasion Game: Evaluation of Some Communication Schemes

Alberto Speranzon, Karl Henrik Johansson

Royal Institute of Technology
Department of Signals, Sensors and Systems
Osquldaväg 10, SE 100-44, Stockholm, Sweden
Email: {albspe,kallej}@s3.kth.se

***Abstract*— A probabilistic pursuit–evasion game from the literature is used as an example to study constrained communication in multi-robot systems. Communication protocols based on time-triggered and event-triggered synchronization schemes are considered. It is shown that by limiting the communication to events when the probabilistic map derived by the individual pursuers contain new information, as measured through their map entropy, the utilization of the communication link can be considerably improved compared to conventional time-triggered communication.**

I. Introduction

Multi-robot systems have many advantages compared to single-robot systems, including improved flexibility, sensing, and reliability. For most mobile robot systems, one need to address challenges related to sensor noise, self-localization, and partial knowledge of the environment. For a multi-robot system, the inter-robot communication adds to this list. In practice, every communication channel has a limited bandwidth, which is both due to the fundamental laws on achievable data rate and to that the channel might be shared with other users. The performance of the multi-robot system is often highly dependent on the utilization of the communication network. However, it seems like integrated design of the communication protocol so far have not been considered for multi-robot systems, cf., [1], [2].

The main contribution of this paper is to illustrate how information theory can be used in the design of a multi-robot system, in order to optimize the communication protocol with respect to some control performance. We let a pursuit–evasion game [3], [4], [5], [6] with several pursuers serve as a prototype system, since it is a good representative for several multi-robot tasks. In particular we consider a probabilistic approach for pursuit–evasion where each pursuer build a probabilistic map of the environment [5], [7]. Map entropy is used in the paper as an information measure of the probabilistic map, cf. [8], [9]. It is used to establish an event-triggered communication scheme for the pursuers, cf., [10], [11]. The considered multi-robot problem can be viewed as a benchmark for the design of integrated control and communication system [12], [13], [14].

The paper is organized as follow. In Section II we extend the pursuit–evasion model of Hespanha et al. [5] by introducing an explicit broadcasting communication protocol for the pursuers. Two particular communication schemes are discussed in Section III: time–triggered and event–triggered synchronization. The synchronization events are in the latter based on the probabilistic map entropy. In Section IV quantization is utilized to cope with bandwidth limitations. It is shown that the map entropy can be used to quantize the probabilistic map in an efficient way. Simulation results are presented in Section V.

II. Pursuit–Evasion with Communication

Consider a pursuit–evasion game with $n_p > 1$ pursuers and one randomly moving evader. Following Hespanha et al. [5], we suppose that the game is played in a finite-dimensional space, uniformly partitioned in $n_c^2 < \infty$ cells denoted $\mathcal{X} = \{1, 2, \dots, n_c^2\}$. Each cell can be occupied by the evader, the pursuers or the obstacles. Neither the evader nor the pursuers can occupy a cell with an obstacle, although the evader and a pursuer can share a cell. The latter corresponds to a capture of the evader. We assume that the time is quantized, such that to each event is associated a time instance $t \in \mathcal{T} = \{0, 1, 2, \dots\}$. The motions of the pursuers and the evader are modelled as a controlled Markov chain, see [6] for details. Pursuer i, $i = 1, \dots, n_p$, senses at each time instance $t \in \mathcal{T}$ the triple

$$\mathbf{z}_i(t) = \{\mathbf{s}_i(t), \mathbf{o}_i(t), \mathbf{e}_i(t)\}$$

where $\mathbf{s}_i(t) \in \mathcal{X}$ is the position of the pursuer, $\mathbf{o}_i(t) \subset \mathcal{X}$ is a measurement of the obstacle locations sensed by the pursuer, and $\mathbf{e}_i(t) \in \mathcal{X}$ is the corresponding measurement of the evader.[1] We assume that all sensors (detecting position, obstacles, and evader) are ideal, and thus are not affected by measurement noise etc. The measurement space is denoted $\mathcal{Z} = \mathcal{X} \times 2^{\mathcal{X}} \times \mathcal{X}$, where $2^{\mathcal{X}}$ denotes the power set of $\mathcal{X}$.

A. Synchronization

We extend the pursuit–evasion model of Hespanha et al. by introducing limited pursuer communication. The pursuers gather individual sensor information and make local decisions, but they can communicate only at synchronization time instances $\tau \in \mathcal{T}$. We denote the data sent by pursuer i at time instance τ to the other pursuers with $\mathbf{y}^i(\tau)$.

[1] Boldface indicates a random variable and the normal typeface its realization.

Definition II.1 *A synchronization is a complete broadcasting communication in which all pursuers exchange information with each other. The data received by pursuer i is*

$$\mathbf{Y}^i(\tau) = \left\{\mathbf{y}^j(\tau)\right\}_{j\neq i}$$

for some synchronization time $\tau \in \mathcal{T}$.

The synchronization operation is depicted in Figure 1. The pursuers share network channel, which is used simultaneously when a synchronization is performed at a instance $\tau \in \mathbf{T}_s$, where $\mathbf{T}_s$ denotes the set of synchronization times. In the paper we consider two different types of synchronization: time–triggered and event–triggered synchronization.

Definition II.2 *The time-triggered synchronization set is equal to*

$$\mathbf{T}_s = \{\tau \in \mathcal{T} | \tau = k\Delta, k = 1, 2, \dots\}$$

where $\Delta \in \mathcal{T}$ is a fixed synchronization period.

Definition II.3 *The event-triggered synchronization set is equal to*

$$\mathbf{T}_s = \{\tau \in \mathcal{T} |\, \exists i : \Phi(W^i(\tau), \tau) < \lambda(\tau)\}$$

where the event map Φ depends on the data $W^i(\tau)$ available for pursuer i at time τ and the time instance $\tau \in \mathcal{T}$. The synchronization threshold is denoted $\lambda(\tau) \in \mathbb{R}$.

In the paper, the synchronized data $\mathbf{y}^i(\tau)$ are the probabilistic maps introduced next.

B. Probabilistic Map

Definition II.4 *The probabilistic map is the probability density function for the position of the evader $\mathbf{x}_e$ conditioned on the available data. For each pursuer i with the local probabilistic map is given by*

$$\tilde{p}^i_{t+1|t}(x_e, \overline{x}_e, Z_t) \triangleq P(\mathbf{x}_e(t') = x_e | \mathbf{x}_e(t) = \overline{x}_e, \mathbf{Z}^i_t = Z_t)$$

where $\mathbf{Z}^i_t \in \mathcal{Z}^*$ represents the sequence of measurements taken by pursuer i up to time t and $\mathbf{x}_e$ is the position of the evader.

The local probabilistic map is updated through a two-step algorithm: a measurement step in which $\tilde{p}^i_{t|t}(x_e, Z_t)$ is computed using the current measurements, and a prediction step in which $\tilde{p}^i_{t+1|t}(x_e, \overline{x}_e, Z_t)$ is computed using an evader model, see [5] for details. The game starts with an a-priori probabilistic map $\tilde{p}^i_{0|-1}(x_e, \emptyset, \emptyset)$ that we assume to be the uniform distribution. Let $M^i(t)$ denote the conditional probability $\tilde{p}^i_{t+1|t}(x_e, \overline{x}_e, Z_t)$ for a particular realization. Note that $M^i(t)$ is an $n_c \times n_c$ matrix with elements that sum to one. The data received by pursuer i at synchronization time $\tau \in T_s$ is then given by a collection of matrices $\mathbf{Y}^i(\tau) = \{M^j(\tau)\}_{j\neq i}$.

C. Control Policy

The control action at $t \in \mathcal{T}$ for pursuer i with control policy γ_i is

$$\mathbf{u}_i(t) = \gamma_i(\mathbf{Z}^i_t, \mathbf{Y}^i_t)$$

where

$$\mathbf{Y}^i_t = \{\mathbf{Y}^i(\tau_0), \mathbf{Y}^i(\tau_1), \dots, \mathbf{Y}^i(\tau_j)|\} \quad \tau_j \leq t < \tau_{j+1}$$

and

$$\mathbf{Z}^i_t = \{\mathbf{z}_i(0), \dots, \mathbf{z}_i(t)\}$$

In particular, we limit the discussion to greedy control policies with constrained motion, i.e.,

$$\mathbf{u}_i(t) = \arg \max_{v \in \mathcal{N}(s_i) \subset \mathcal{U}} p^i_{t+1|t}(v, \bar{x}_e, Z_t, Y_t)$$

where $\mathcal{N}(s_i)$ are all neighboring cells of the current position of pursuer i. Thus, at each time instant t the control policy γ_i moves pursuer i to a neighboring cell v that maximizes the conditional probability of finding the evader at time $t+1$, given the measurements and communicated data up to that time. Note that the greedy policy does not maximize the local probabilistic map $\tilde{p}^i_{t+1|t}$, but the probabilistic map $p^i_{t+1|t}$ which depends on the data received through the network. The probabilistic map $p^i_{t+1|t}$ is the fusion of the local probabilistic map with all received probabilistic maps. (Fusion of probabilistic maps is discussed in [15].) The fusion in this paper is computed as a normalized product of all the maps, i.e., as an independent opinion pool [16]. This is motivated by the assumption that the pursuers are far from each other so measurements are approximately independent.

III. Entropy-triggered synchronization

In this section we introduce an event-triggered synchronization scheme based on the map entropy.

Definition III.1 *The map entropy of a probabilistic map $M(t) = p_{t+1|t}$ is equal to*

$$H(M(t)) = -\frac{1}{\log |\mathcal{X}|} \sum_{x_e \in \mathcal{X}} p_{t+1|t} \log p_{t+1|t} \tag{1}$$

The *relative entropy* [9], $D(p||q)$, gives a measure of the distance between two probability distributions, p and q defined over the set $\mathcal{X}$. It is defined as

$$D(p||q) = \sum_{x \in \mathcal{X}} (p \log (p/q))$$

In particular, the relative entropy between the probabilistic map $p^i_{t+1|t}(x_e, \overline{x}_e, Z_t, Y_t)$ and the uniform probabilistic map $p^i_{0|-1}(x_e, \emptyset, \emptyset, \emptyset)$ is

$$D(p^i_{t+1|t}||p^i_{0|-1}) \propto 1 - H(M^i(t)) \tag{2}$$

Where $D(p^i_{t+1|t}||p^i_{0|-1})$ thus provides a measure of how much the probabilistic map have changed since initial time. If we assign to the probabilistic map $p^i_{0|-1}$ zero information content, then the map entropy $H(M^i(t))$ is a measure of information

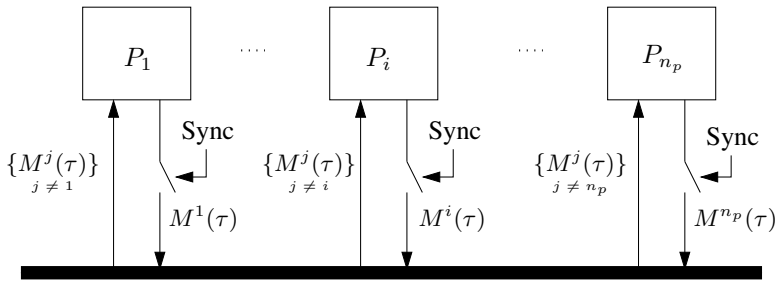


Fig. 1. At a synchronization time $\tau \in \mathcal{T}$, each pursuer P_i broadcasts the data $M^i(\tau)$ to the network and receives the data $\{M^j(\tau)\}_{j\neq i}$ from the other pursuers.

content of a probabilistic map as expressed by (1). In the following we consider two event-triggered synchronization schemes based on the map entropy.

A. Synchronization based on threshold

Let the event map be equal to the map entropy: $\Phi = H$. Then a synchronization event is triggered whenever $H(M^i(\tau)) < \lambda(\tau)$, where λ is the synchronization threshold. We consider two cases: synchronization based on a fixed threshold and synchronization based on a dynamic threshold. In the latter case, the time-varying threshold $\lambda(\tau)$ is given by the difference equation

$$\lambda(\tau+1) = \alpha\lambda(\tau)$$
$$\lambda(0) = \lambda_0$$

where λ_0 is the initial threshold and $0 < \alpha < 1$. Since the map entropy H in most experiments is decreasing, a fixed threshold, $\lambda(\tau) = \lambda$, leads to that the synchronization event $H < \lambda$ is triggered for all $t > \tau \in \mathcal{T}$, where τ is the first synchronization instance. The synchronization will hence take place also when there is no new information in the probabilistic map M^i. Therefore, it is natural to introduce a decreasing dynamic threshold.

B. Synchronization based on relative entropy

Synchronization based on relative entropy triggers a synchronization event when the local probabilistic map differs sufficiently much from the previously broadcasted map. The difference is measured through the relative entropy D_i. Note that in deriving D, we need to neglect the zero elements corresponding to the pursuers positions. The synchronization is carried out when

$$D(p^i_{t+1|t}||p^i_{\tau-1|\tau-2}) > \lambda \qquad \tau - 1 \leq t < \tau$$

that is, synchronization is performed when the relative entropy between the current probabilistic map and the last synchronized probabilistic map is larger than $\lambda > 0$.

IV. Bandwidth limitations

In order to cope with communication bandwidth limitations, it is natural to send only a part of the probabilistic map M that contains most of the information.[2] The idea is to transform the map M into a new map K, denoted reproduction probabilistic map. That map should contain almost all information in M but it should require less bits to be encoded. We consider a vector quantization

$$Q : \mathbb{R}^{n_c\times n_c} \to \mathbb{R}^{n_c\times n_c} : M \mapsto K = Q(M)$$

where Q defines a partition of the matrix M into square sub-matrices $M_1, \ldots, M_N$ of dimension $n_1, \ldots, n_N$ such that $\sum_{i=1}^N n_i = n_c$. The reproduction probabilistic map K is block partitioned correspondingly into $K_1, \ldots, K_N$ with

$$K_i = \frac{1}{n_i^2}\,\mathbb{1}_{n_i\times 1} M_i \mathbb{1}_{1\times n_i}\mathbb{1}_{n_i\times n_i} \tag{3}$$

where $\mathbb{1}_{k\times\ell}$ defines a $k\times\ell$ matrix with unit elements. Each element of K_i is thus given by the average of the elements of M_i. Associated with the quantization Q, we define a distortion measure

$$d(M,K) = |H(M) - H(K)| \tag{4}$$

where $H(M)$ and $H(K)$ are the map entropies of M and K, respectively. The choice of granularity in the block partition, i.e., the size $n_i\times n_i$ of the sub-matrices of the partition, should be chosen such that $d(M,K)$ is small. This corresponds to a small loss of information in the quantization. That choice of quantization granularity is not always possible due to limited communication bandwidth. Trade-off between quantization granularity and distortion is treated by the rate distortion theory [17]. Analytical solutions seem to be hard to obtain in our case. We therefore consider two heuristic approaches next.

A. Uniform quantization

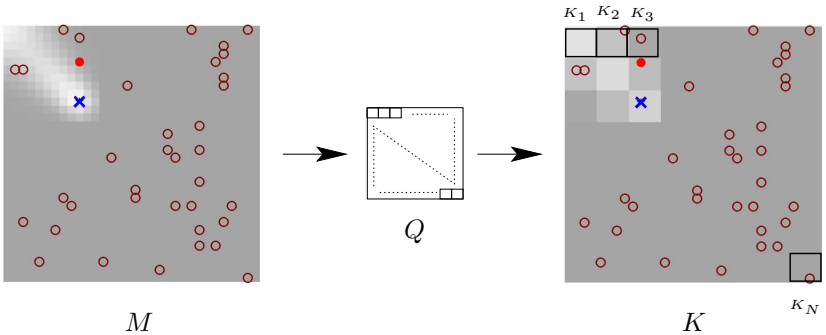
For uniform quantization, the block partition of Q is such that the dimensions of all square blocks $M_1, \ldots, M_N, K_1, \ldots, K_N$ are equal to n. An illustrative example is shown to the right in Figure 2(a). The reproduction probabilistic map K contains $O(n_c^2/n^2)$ values (one for each sub-matrix) instead of $O(n_c^2)$ as is the case for (the whole map) M.

B. Non-uniform quantization

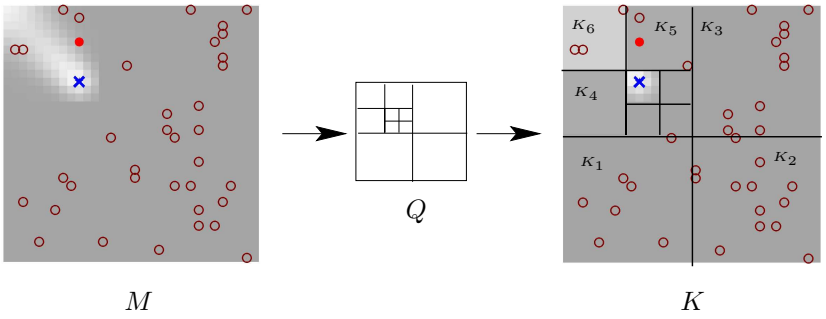
A possible non-uniform quantization is illustrated in Figure 2(b). This corresponds to a "divide-and-conquer" scheme,

[2] In this section, the pursuer index i and the time dependence are suppressed.

,



(a) Uniform quantization



(b) Non-uniform quantization

Fig. 2. Vector quantization of probabilistic map M into $K = Q(M)$.

which is known as vector quantization with QuadTree map [18]. The partition $M_1, \ldots, M_N$ imposed by the quantization Q is in this case carried out recursively, such that

$$\begin{aligned} \dim M_1 &= \frac{1}{4} \dim M \\ \dim M_{i+1} &= \begin{cases} \dim M_i, & \text{if } i \bmod 3 = 0; \\ \frac{1}{4} \dim M_i, & \text{otherwise} \end{cases} \end{aligned} \tag{5}$$

where mod is the modulus operation. In each recursion step, the current block is divided into four sub-matrices. Three of them are quantized using (3), while the remaining sub-matrix is partitioned in four smaller sub-matrices, and so on. The recursion stops when the smallest block has reached the preassigned dimension n.

Compared with uniform quantization, one advantage of the proposed non-uniform quantization is the possibility of an on-line termination of the quantization if the loss of information is too high, i.e., if the distortion measure (4) is large. Solving the recursion (5), we find that the number of values to transmit is equal to $O(\log n_c^2 + n^2)$.

V. Simulation Results

Sets of 100 Monte Carlo simulations have been performed in order to evaluate the presented synchronization and quantization strategies. The capture time T^* and the mean number of synchronization instances $|\mathbf{T_s}|$ are used as performance indices.

Figures 3 and 4 show the results for a game with two pursuers and one evader on a grid with $n_c^2 = 576$ cells. Three different synchronization schemes are compared: time-triggered, event–triggered based on dynamic threshold, and event-triggered based on relative entropy. The time-triggered synchronization had a synchronization period $\Delta = 20$ s. We see in Figure 3 that the capture time T^* is varying considerably much over the set of experiments. The mean capture time $\overline{T}^*$ is similar for all the synchronization schemes as indicated by the dashed lines (dashed-dotted lines indicate the standard deviations). The values are collected in the following table

Synchronization schemes	$\overline{T}^*$	$\overline{\lvert\mathbf{T}_s\rvert}$
Time-Triggered	68	3.859
Event-Triggered Dynamic Threshold	64	2.454
Event-Triggered Relative Entropy	66	2.576

Note that the mean number of synchronization times $\overline{|\mathbf{T}_s|}$ is much smaller for the event-triggered schemes than for the time-triggered. Hence, event-triggered synchronization allow a more efficient utilization of the communication channel. This fact is also illustrated in Figure 4. The main difference between the two event-triggered schemes is mainly the distribution of synchronization times. We see that when using relative entropy the pursuers tend to communicate more regularly. This is due to that new information is available quite regularly for the pursuers and this information triggers

the synchronization events in this scheme.

In Figure 5 the uniform and non-uniform quantization strategies are compared. The results are for a game with two pursuers and one evader on an environment that consists of $n_c^2 = 1024$ cells. The synchronization of the probabilistic maps among pursuers is time-triggered with period $\Delta = 20$ s. The quantization map Q has been chosen so that the dimension of the sub-matrices K_i is $n^2 = 64$. Figure 5 shows T^* for the following cases: no quantization ($n = 1$), uniform quantization with $n = 8$ and non-uniform quantization with $n = 8$. The averaged experimental results are collected in the following table

Quantization	$\overline{T}^*$	$\overline{d}(M, K)$	$\overline{V}$
Uniform ($n = 1$)	90.64 s	(0, 0)	1024
Uniform ($n = 8$)	141.46	(0.09, 0.1)	≈ 16
Non-uniform ($n = 8$)	120.26	0.04, 0.04	≈ 70

Here $\overline{d}(M, K)$ denotes the average distortion over all experiments and $\overline{V}$ the average number of broadcasted values. Notes that $\overline{V}$ is one or two magnitudes smaller for the quantized cases compared with the non-quantized case. Still the mean capture time is only about 50% larger. The uniform quantization compared with the non-uniform quantization has quite high average distortion $\overline{d}(M, K)$ (in the table the average distortion for two probabilistic map of the two pursuers that are playing the game are reported). This implies a relevant loss of information that makes $\overline{T}^*$ larger in this last case. On the other hand the average number of transmitted data $\overline{V}$ is considerably reduced.

VI. Conclusions and Future Work

In this paper, we have presented communication protocols based on time-triggered and event-triggered synchronization for a distributed pursuit–evasion game. The event-triggered schemes were based on the entropy of the probabilistic map. Simulations showed that by limiting the communication to certain events, the utilization of the communication link can be considerably improved compared to conventional time-triggered communication. Two different vector quantization maps are considered in order to cope with bandwidth limitation. A distortion measure based on map entropy is introduced for evaluating the compression of the probabilistic map. The problem of the compression of the probabilistic map M can be also treated considering in the framework of image transmission. Using standard algorithms, like for example JPEG, we would cope with the bandwidth limitation of the channel but it would be difficult to guarantee that the new probabilistic map would contain most of the information of the original one.

The communication schemes developed in the paper can be applied in many cases when a probabilistic map has to be sent through a network channel to a decision maker. These kind of problems are common in robotics, examples are the occupancy grids used for localization of robots in an unknown environment [19], [2] and applications of particle filters [20]. A related problem of our current interest is optimal control of mobile sensors that share a bandwidth limited communication channel.

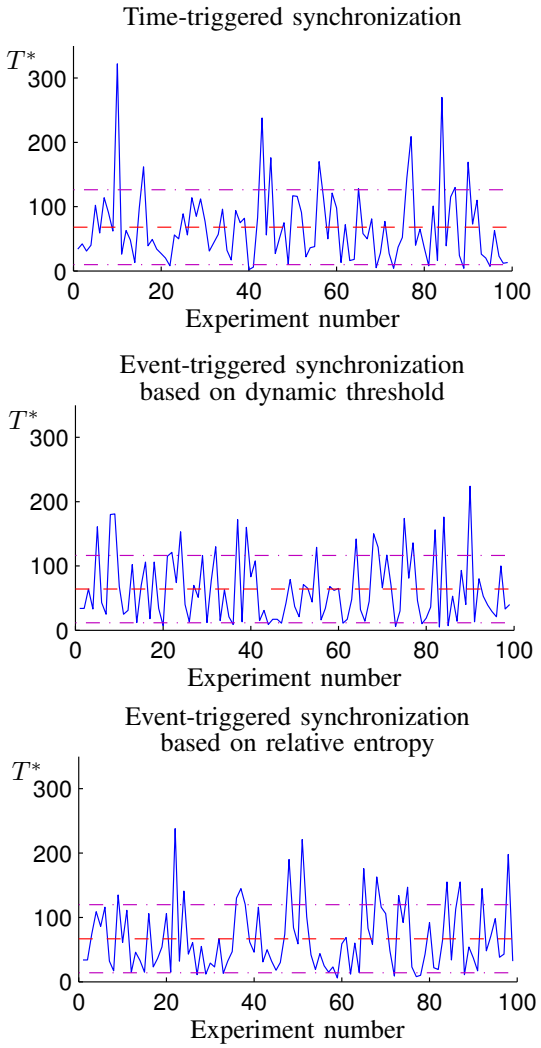


Fig. 3. In the plots are shown the capture time $T^*$for all the 100 Monte Carlo experiments. In dashed line is shown the mean capture times $\overline{T}^*$. In dashed-dotted line the standard deviations. The map size consider in these simulations is $n_c^2 = 576$.

VII. Acknowledgments

The authors gratefully acknowledge Joao P. Hespanha for fruitful discussions and for providing us with a simulation environment. This work is supported by the European Commission though the RECSYS project IST-2001-32515 and the Swedish Research Council.

References

[1] T. Balch and R. C. Arkin, "Communication in reactive multiagent robotic systems," *Autonomous Robots*, pp. 27–52, 1994.

[2] S. Thrun, W. Burgard, and D. Fox, "A probabilistic approach to concurrent mapping and localization for mobile robots," *Machine Learning and Autonomous Robots (joint issue)*, vol. 31, pp. 29–53, 1998.

[3] R. Isaac, *Differential Games*, 2nd ed. Robert E. Kriger Publisher Company, 1967.

[4] T. Basar and J. G. Olsder, *Dynamic noncooperative game theory*, 2nd ed. Academic Press, 1995.

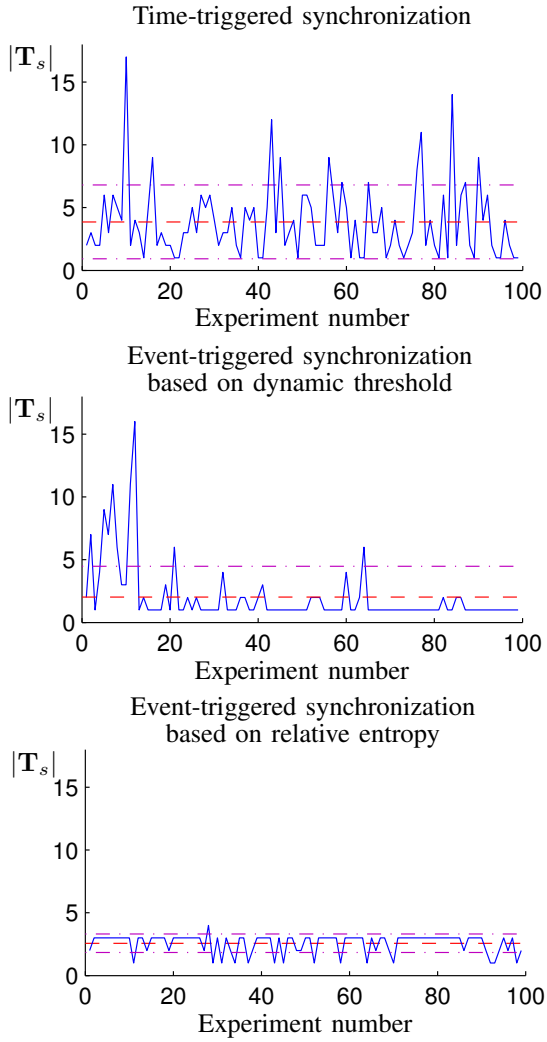


Fig. 4. In the plots are shown the number of synchronization instances $|\mathbf{T}_s|$ for all the 100 Monte Carlo experiment. In dashed-line the mean number of synchronization instances $\overline{|\mathbf{T}_s|}$. In dashed-dotted line the standard deviation. The map size consider in this simulations is $n_c^2 = 576$.

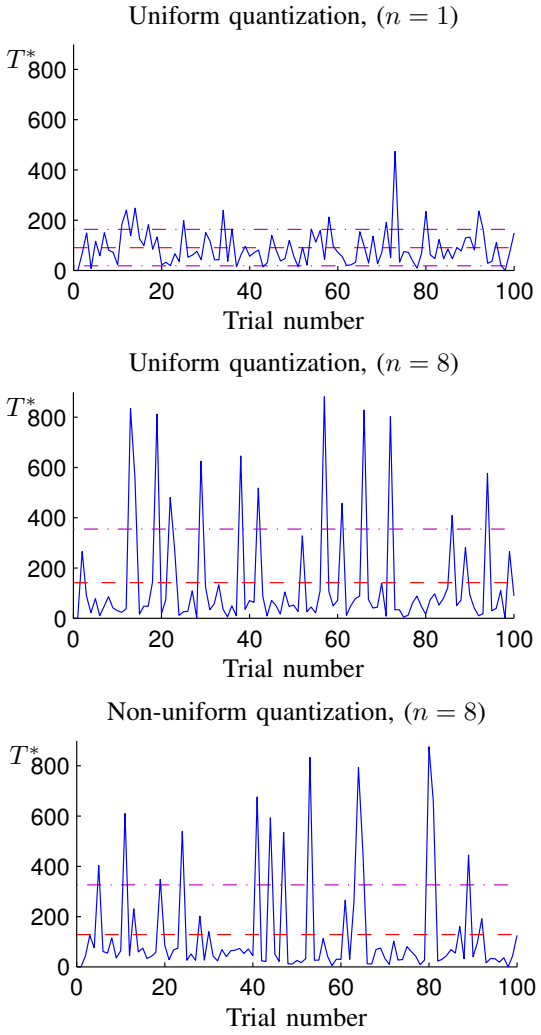


Fig. 5. In the plots are shown the capture time T^* for all the 100 Monte Carlo experiments. In dashed line the mean capture time $\overline{T}^*$. In dashed-dotted the standard deviation. The size of the map used in the simulations is $n_c^2 = 1024$.

[5] J. P. Hespanha, H. J. Kim, and S. Sastry, "Multiple-agent probabilistic pursuit-evasion games." in *In 38th Conf. on Decision and Control*, vol. 3, 1999, pp. 2432–2437.

[6] J. P. Hespanha and M. Prandini, "Optimal pursuit under partial information," in *In Proc. of the 10th Mediterranean Conference on Control and Automation*, 2002.

[7] J. P. Hespanha, H. H. Kizilocak, and Y. S. Ateskan, "Probabilistic map building for aircraft–traking radars," University of Southern California, Los Angeles, CA, Tech. Rep., 2000.

[8] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423,623–656, 1948.

[9] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, D. L. Schilling, Ed. Wiley Interscience, 1991.

[10] H. Koeptz, "Should responsive systems be event triggered or time triggered?" *IEICE Trans. on Information and Systems*, no. 10, pp. 1525–1532, 1993.

[11] K. J. Åström and B. M. Bernhardsson, "Comparison of Riemann and Lebesgue sampling for first order stochastic systems," in *Proceedings of the 41st IEEE Conference on Decision and Control*, 2002.

[12] S. K. Mitter, "Control with limited information: the role of systems theory and information theory," *IEEE Information Theory Society Newsletter*, vol. 4, no. 50, pp. 122–131, 2000.

[13] W. S. Wong and R. W. Brockett, "Systems with finite communication bandwidth constraints ii: stabilization with limited information feedback," *IEEE Trans. on Automatic Control*, vol. 44, no. 5, pp. 1049–1052, 1999.

[14] S. Tatikonda and S. K. Mitter, "Control under communication constraints," in *IEEE Conference on Decision and Control*, 2000.

[15] J. P. Hespanha and H. Kizilocak, "Efficient computation of dynamic probabilistic maps," in *In Proc. of the 10th Mediterranean Conference on Control and Automation*, 2002.

[16] J. Berger, *Statistical Decision Theory and Bayesian Analysis*. Springer–Verlag, Berlin, 1985.

[17] T. Berger, *Rate Distortion Theory*. Prentice-Hall, Inc. Englewood Cliffs, NJ, 1970.

[18] A. Gersho and R. Gray, *Vector Quantization and Data Compression*. Kluwer, 1991.

[19] H. Moravec and A. Elfes, "High resolution maps from wide angle sonar," in *In Proc. of the IEEE International Conference on Robotics and Automation (ICRA)*, 1985.

[20] S. Thrun, "Particle filters in robotics," in *Proceedings of the 17th Annual Conference on Uncertainty in AI (UAI)*, 2002.