

# Performance of State Estimate Fusion in Long-Haul Sensor Networks with Message Retransmission

Qiang Liu and Xin Wang

SUNY at Stony Brook

Stony Brook, NY 11794

Email: {qiangliu,xwang}@ece.sunysb.edu

Nageswara S. V. Rao

Oak Ridge National Laboratory

Oak Ridge, TN 37831

Email: raons@ornl.gov

Katharine Brigham and B. V. K. Vijaya Kumar

Carnegie Mellon University

Pittsburgh, PA 15213

Email: {kbrigham,kumar}@ece.cmu.edu

**Abstract**—We consider a long-haul network in which sensors are remotely deployed and tasked to send state estimates of a dynamic target to the fusion center via satellite links. Due to severe loss and delay inherent in satellite channels, the number of estimates successfully arriving at the fusion center can be quite limited. A certain retransmission-based transport protocol can be applied so that lost messages can be recovered over time. However, excess delay may be incurred with message retransmission that can potentially violate the deadline for reporting the estimate. In this work, we analyze the recovery rate and the arrival time performances of the state estimates with message retransmission and time cutoff. We are particularly interested in exploring the extent to which message retransmission can potentially improve the system performance. Results of simulation studies of a ballistic target tracking application are shown in the end to complement our analysis.

**Index Terms**—Long-haul sensor networks, state estimate fusion, message retransmission, mean-square-error performance.

## I. INTRODUCTION

In a long-haul sensor network, sensors are deployed to cover a vast geographical area, which may be a continent or even the entire globe depending on the specific application. We consider a class of such networks in which state estimates (e.g., position and velocity) of certain dynamic targets – such as aircrafts or ballistic missiles [4] – are sent from the remote sensors to a fusion center so that a global estimate can be obtained by fusing the individual estimates. In this case, satellite links may well be the only type of cost-effective medium for such long-range communications because of the prohibitive cost of extending submarine and terrestrial fiber connections extensively to often rough terrain and sparsely populated areas. There are challenges in satellite link-based monitoring and tracking applications. Because of the long distance (that could be tens of thousands of miles long), the propagation time of signals is significant. For example, the round-trip time (RTT) for signal propagation with a geostationary earth orbit (GEO) satellite is more than a half second [11]. More importantly, communication over satellite links is characterized by sporadic high bit-error rates (BERs) and burst losses<sup>1</sup>. Losses either incurred during transmission or resulting from the high BERs could further effectively reduce the number of messages available at the fusion center. It is well known that fusion of estimates from different sensors is

a viable means of reducing the estimation error; with high loss rates, however, only a portion of the potential fusion gain could be actually achieved and the accuracy of the fused estimate output thus obtained may be deemed unacceptable by the system operator. Apparently, all the above-mentioned drawbacks of the satellite links could work against the very purpose of the underlying task – to promptly and accurately report state estimates – and may violate the requirement on the worst-case estimation error.

One way to counteract the effect of the lossy transmission link is to adopt certain transport protocols in which message retransmissions are implemented. Some lost packets can be recovered after one or multiple retransmissions. However, we should not neglect another aspect of the system requirement – the delay performance. Owing to the often near real-time requirement of the monitoring/tracking task, the system often allows for small time gap from the time instant of interest to the time when the global estimate should be finally obtained and reported. This often comes as a predefined reporting deadline before which an estimate must be reported by the fusion center. Message retransmission, albeit easy to implement, may nevertheless exacerbate the reporting delay performance by incurring extra time on top of the already relatively large propagation and transmission latency. When needing to receive a message with greater certainty, the fusion center may nonetheless have to increase its reporting time significantly, even at the risk of violating the stipulated reporting deadline.

We consider state estimation with certain transport protocols being implemented that include message retransmission mechanism. The transmission control protocol (TCP) implemented in wired Internet and wireless local area networks (WLANs) is still garnering research efforts that are too numerous to list. Analysis of TCP-like transport protocols over satellite links can be found in studies such as [1], [6]. Commonly acknowledged are the difficulties in applying “conventional” TCP protocols to transmission over satellite links, mainly because of the very large propagation delay not encountered in other networks. The specificity of the application in our study also somewhat distinguishes it from the analysis geared toward the voice- and video-based broadcasting and data-based Internet access, both of which have continuous data in flight. Also of note is that in our settings, state estimates from the remote sensors are generally intermittently sent over a wide-band satellite channel – with the interval possibly ranging from a few times within

<sup>1</sup>The satellite link is subject to different sources of degradation. For example, the Ka and Ku bands are subject to heavy rain attenuation, whereas communication in the C band is more likely to experience interferences from terrestrial microwave links [11].

a second to once every few minutes – and thus congestion is not as much a concern as in conventional TCP applications. Hence, we assume a simplified transport protocol in which retransmission is performed on the message-level basis.

State estimation under imperfect communications has been studied in the literature. Packet delay in the order of multiple sampling periods have been addressed and state augmentation [12] is often the default solution for fixed delay. In [5], the authors have derived an upper bound of the packet loss rate above which the estimation error will go unbounded. A dynamic selective fusion method is proposed in [9] so that fusion is deferred till enough estimates have arrived at the fusion center. We emphasize that design of advanced transport protocols for state estimate fusion applications lies beyond the scope of this paper. Nevertheless, to the best of our knowledge, this work is among the first to link transmission schemes with state estimation performance and is our first-step effort toward further design of transport protocols and filtering algorithms in long-haul sensor networks.

The major contributions of this work are as follows: By quantifying the performance of delivery probability and finite moments of arrival latency under variable link conditions and retransmission parameter setups, we demonstrate the trade-off between meeting the estimation accuracy and reporting delay requirements. Simulation results from a ballistic target tracking application are presented in the end to complement our analytical study.

The paper is organized as follows: In Section II, the message retransmission mechanism in a long-haul sensor network is briefly introduced. In Section III, we provide detailed analysis of the delivery rate of a message as well as its arrival time distribution. We then present simulation results of a ballistic target tracking application in Section IV before we conclude the paper in Section V.

## II. MESSAGE RETRANSMISSION OVER A LONG-HAUL CONNECTION

In the long-haul sensor network, the message retransmission process evolves as follows. A remote sensor sends out a message containing the state estimate; upon successful reception of this message, the fusion center sends back an acknowledgment (ACK) message to the sensor. Failure of the ACK message to arrive before the expiration of the timeout  $T_{TO}$  – due to loss and/or long delay of the message itself or the ACK – will prompt the sensor to retransmit the message. Typically,  $T_{TO}$  could be several times the RTT of the connection, and over long-haul connections it could be of the order of seconds. Setting  $T_{TO}$  too long could reduce the maximum number of retransmissions, thereby limiting the potential to recover the lost message; on the other hand, a short  $T_{TO}$  may incur many rounds of retransmission (often unnecessarily) as the sensor sometimes could have waited a bit longer to receive the ACK. Such retransmission continues till the acknowledgment is received by the sensor, or the retransmission window  $W$  expires. This window should ideally contain multiple  $T_{TO}$  periods so that under adverse link conditions, it's likely that the message can eventually be recovered after multiple tries.

The reporting deadline may limit the potential gain from retransmission as the overall time before reporting can be very short. The cutoff time  $T_{CO}$  on the one hand limits the total number of retransmissions, and on the other may prevent certain messages from being eventually delivered due to the randomness of the delay. Of necessity is statistical analysis of this randomness in arrival delay that would allow us to take a closer look at the retransmission process and then the fusion performance. In what follows, we derive the message delivery probability and the statistical distribution of the arrival time.

## III. MESSAGE DELIVERY PERFORMANCE

The message-level loss and delay characteristics are directly determined by the condition of the long-haul link. We assume that each message sent by a sensor is lost during transmission with probability  $p$  independently of other messages. Normally, the latency that a message experiences before arriving at the fusion center consists of the initial detection and measurement delay, data processing delay by both the sensor and the fusion center, propagation delay, and transmission delay, among others. These are collectively considered as the minimum delay that a message must undergo to reach the fusion center, which is often bounded by characteristics of the physical link, such as the distance of the satellite link, and those of the transmission, such as the transmission data rate and length of the message. The extra random delay, due to link conditions such as weather and terrain, could be anywhere from zero upward. Suppose a pdf  $f(t)$  can model the overall delay  $t$  that a message experiences to be successfully delivered to the fusion center. One typical example is that of the shifted exponential distribution:

$$f(t) = \frac{1}{\mu} \exp^{-\frac{t-T}{\mu}}, \text{ for } t \geq T. \quad (1)$$

in which  $T$  serves as the common link and processing delay, and  $\mu$  is the mean of the random delay besides  $T$ . In a real system, the empirical values of the message delay can be measured over time and thus an approximate function  $\hat{f}$  can be estimated.

### A. Message Delivery Probability

We are interested in the average probability of a message being successfully delivered by a certain time, that is, by the cutoff time  $T_{CO}$ . An estimate is only counted once even if it arrives multiple times due to retransmission. The duplicate messages received by the FC can simply be ignored as they will not contribute further to the fusion performance. Since we focus on one instance of the transmission between the sensor and FC, we let  $t = 0$  denote the time of interest for which an estimate is obtained; all the time measures in this work are relative to this zero time.

The maximum number of retransmissions before the cutoff time  $T_{CO}$  is

$$K_{retx} = \left\lceil \frac{\min\{T_{CO}, W\}}{T_{TO}} \right\rceil - 1. \quad (2)$$

From the definition,  $K_{retx} + 1$  is the total rounds of transmission, including the original and subsequent retransmissions. This number is measured from the perspective of the sensors; due to the link and processing delay, it's possible that the

message re-sent during the last round will not be delivered by the cutoff time. Nevertheless, the sensor and the FC should be coordinated in time so that any round of retransmission is likely to arrive before the cutoff time. In other words, the TCP window size  $W$  should be commensurate with the range of cutoff time  $T_{CO}$ .

We define  $p_{del,t}^k$  as the probability that a message is delivered by time  $t$  after  $k$  rounds of retransmissions, and

$$T_{retx,k} = T_{CO} - kT_{TO}, \text{ for } k = 0, 1, \dots, K_{retx} \quad (3)$$

as the duration of the period  $[kT_{TO}, T_{CO}]$  in which the  $k$ -th retransmitted message is in flight and could be potentially delivered to the fusion center.

When there is no retransmission within  $[0, t]$ , the probability of a message being delivered by time  $t$  is

$$p_{del,t}^0 = (1 - p)F(t), \quad (4)$$

in which  $F(t) = \int_0^t f(u) du$  is the cdf of the arrival delay. Its complement, the probability that the original message is unavailable at time  $t$ , is denoted as

$$p_{loss,t}^0 = 1 - p_{del,t}^0 = p + (1 - p)\bar{F}(t), \quad (5)$$

in which  $\bar{F}(t) = 1 - F(t)$  is the tail distribution. With these two probabilities, we can derive the message delivery rate  $p_{del,T_{CO}}^{K_{retx}}$ .

The original message is delivered by  $T_{CO}$  with probability

$$p_{del,T_{CO}}^0 = p_{del,T_{retx,0}}^0 = (1 - p)F(T_{retx,0}). \quad (6)$$

And with the first round of retransmission, the delivery probability totals

$$p_{del,T_{CO}}^1 = p_{del,T_{CO}}^0 + p_{loss,T_{retx,0}}^0 p_{del,T_{retx,1}}^0. \quad (7)$$

In general, for the  $k$ -th ( $0 < k \leq K_{retx}$ ) round of message retransmission, we have

$$p_{del,T_{CO}}^k = p_{del,T_{CO}}^{k-1} + p_{del,T_{retx,k}}^0 \left( \prod_{i=0}^{k-1} p_{loss,T_{retx,i}}^0 \right). \quad (8)$$

In other words, the extra delivery rate from the  $k$ -th round is realized when all the previous  $k - 1$  retransmissions and the original message are not available by  $T_{CO}$ . Subsequently, we can obtain the overall message delivery probability within time  $[0, T_{CO}]$  by summing up all such probabilities:

$$\begin{aligned} p_{del,T_{CO}}^{K_{retx}} &= \sum_{k=0}^{K_{retx}} p_{del,T_{retx,k}}^0 \left( \prod_{i=0}^{k-1} p_{loss,T_{retx,i}}^0 \right) \\ &= (1 - p) \sum_{k=0}^{K_{retx}} F(T_{retx,k}) \left\{ \prod_{i=0}^{k-1} [1 - (1 - p)F(T_{retx,i})] \right\}. \end{aligned} \quad (9)$$

### B. Arrival Time: One-way Communication Analysis

So far we have combined the message-level loss rate and one-way arrival delay distribution to obtain the message delivery probability with a certain deadline requirement as specified by  $T_{CO}$ . In practice, the fusion center should be afforded some flexibility in deciding its actual cutoff time that does not violate the systemic value  $T_{CO}$ . First, sometimes reporting an abnormal change promptly is more crucial than initially

providing an accurate estimate because the alert level can be increased immediately that facilitates further investigation. As the maximum allowable delay,  $T_{CO}$  may nevertheless be too large for such rare but time-critical incidences. On the other hand, owing to the dynamic environment of the field, sometimes the fusion center may decide to reduce its waiting time for the retransmitted messages because of increased computational effort to obtain the final estimate, due to an increased number of sensors or increased state dimensionality when multiple closely positioned targets are in clutter [3].

In this subsection, we aim to derive the distribution of the arrival time, and more particularly, the cdf of the first instance of arrival before  $T_{CO}$ . This provides us a view of the internal structure of the arrival process within  $[0, T_{CO}]$ , which could be explored for the above flexible scheduling of early cutoff.

We again treat loss and delay as two independent processes, although a lost message can be regarded as having an arrival delay of infinity. In the last subsection, only the one-way delay characterized by the pdf  $f(t)$  is considered because we noticed the equivalence of the final delivery probability no matter how acknowledgments are actually received. In deriving the arrival time, however, we need to consider two-way delay as well: loss and latency of ACKs would affect the total number of retransmissions, which in turn decides the distribution of the arrival time. For ease of explication though, we first work on the case in which there are exactly  $K_{retx}$  retransmissions – as if no ACKs were ever sent back by the fusion center – and later extend the results to an arbitrary number of retransmissions.

First, we define the cdf of a truncated nonnegative random variable  $Y_T$  with the upper truncation point  $b > 0$  as

$$F_T^b(y) = \frac{F(y)}{F(b)}, \text{ for all } 0 \leq y \leq b. \quad (10)$$

And the associated pdf is

$$f_T^b(y) = \frac{f(y)}{F(b)}, \text{ for all } 0 \leq y \leq b. \quad (11)$$

In our study, we are interested in a series of truncated cdfs and pdfs corresponding to different retransmission cycles. More specifically, we consider the  $k$ -th round of retransmitted message has a truncated cdf by time  $T_{retx,k}$  as

$$F_T^{T_{retx,k}}(t) = \frac{F(t)}{F(T_{retx,k})}, \text{ for all } 0 \leq t \leq T_{retx,k}. \quad (12)$$

Let  $D$  denote the arrival time of the message, and more specifically,  $D_k = d_k + kT_{TO}$  the arrival time of the  $k$ -th retransmitted message (or the original message when  $k = 0$ ). Apparently,  $d_k$  denotes the arrival delay of the  $k$ -th retransmission.

We are interested in deriving the distribution of  $D_{(1)}$  – the time of the first arrival – before the cutoff time  $T_{CO}$ . When there are a maximum number of  $K_{retx}$  retransmissions before  $T_{CO}$ , we let

$$D_{(1),T_{CO}}^{K_{retx}} = \min_{k=0,\dots,K_{retx}} D_{k,T_{CO}}^{K_{retx}} = \min_{k=0,\dots,K_{retx}} \{d_{k,T_{CO}}^{K_{retx}} + kT_{TO}\} \quad (13)$$

be the time of the first arrival among all  $K_{retx} + 1$  messages sent out by the sensor. We note

$$\Pr\{D_{(1),T_{CO}}^{K_{retx}} \leq T_{CO}\} = p_{del,T_{CO}}^{K_{retx}}, \quad (14)$$

where the right-hand side of the equation was given in Eq. (9). Our goal is to derive the distribution of  $D_{(1),T_{CO}}^{K_{retx}}$ , that is,  $\Pr\{D_{(1),T_{CO}}^{K_{retx}} \leq t\}$  for any  $0 < t < T_{CO}$ . For ease of presentation, in the remainder of this section, we assume that a certain  $T_{CO}$  value has been specified along with the resulting  $K_{retx}$  and drop them from the notations unless otherwise specified.

One may first be tempted to directly apply the result of the distribution of the minimum of  $n$  random variables, which is a special case of the order statistics [10]. Despite the seemingly similar relationship, the problem at hand is more complicated. First, from Eq. (13), we need to find the minimum of  $D_k$  for  $k = 0, 1, 2, \dots, K_{retx}$ , which are from different distributions for different  $k$ . Although results for independently and non-identically distributed random variables have been studied in the literature [2], the operations involve substantial use of matrix manipulation, and one must enumerate all  $2^{K_{retx}+1}$  possible arrival patterns since each would yield a distinct result for the distribution of the minimum. To circumvent the issue, we follow another approach by finding the probability that the  $k$ -th retransmitted message (and the original message when  $k = 0$ ) is the earliest to arrive, denoted as  $\Pr\{\mathbb{I}_{D_{(1)}} = k\}$ , in which  $\mathbb{I}$  is the indicator for the earliest arriving message. As stated above, we have

$$\Pr\{\mathbb{I}_{D_{(1)}} = k\} = \Pr\{D_k \leq D_j\} \text{ for all } j \neq k. \quad (15)$$

1)  $K_{retx} = 0$ : For the simplest case where there is no retransmission before the cutoff time  $T_{CO}$ , that is, if  $K_{retx} = 0$ , the cdf of the first instance of arrival is simply the truncated cdf as shown in Eq. (12) for  $k = 0$ . To show this, we first have (a) the delivery probability in Eq. (9) is  $(1-p)F(T_{retx,0})$ ; and (b) the associated probability that the arrival time – which happens to be the delay of the original message as well in this case – is no greater than  $t$  is  $(1-p)F(t)$ . And the cdf can thus be obtained by having (b) divided by (a):

$$F_{D_{(1)}}^0(t) = \Pr\{D_{(1)}^0 \leq t\} = \frac{(1-p)F(t)}{(1-p)F(T_{retx,0})} = F_T^{T_{retx,0}}(t). \quad (16)$$

2)  $K_{retx} = 1$ : When  $K_{retx} = 1$ , the sensor can retransmit the message at most once before the cutoff time. There are basically two scenarios for the first instance of arrival, namely:

- a) the original message arrives first before  $T_{CO}$ ;
- b) the retransmitted message arrives first before  $T_{CO}$ .

The first case can be further subdivided into

- a1) the retransmitted message is not available by  $T_{CO}$ ;
- a2) the retransmitted message is also delivered by  $T_{CO}$ , but its random arrival delay  $d_1$  must satisfy  $d_0 \leq d_1 + T_{TO}$ .

Likewise, for the second scenario, we have

- b1) the original message is not available by  $T_{CO}$ ;
- b2) the original message is also delivered by  $T_{CO}$  with its random arrival delay  $d_0$  satisfying  $d_1 + T_{TO} < d_0$ .

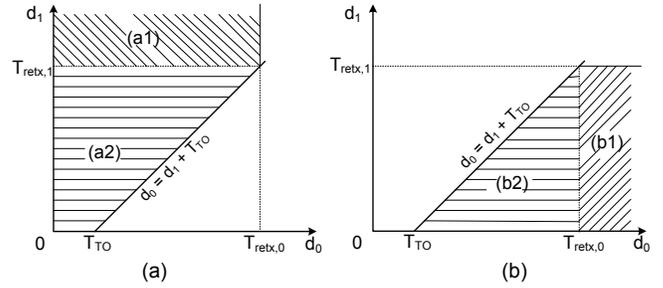


Fig. 1. distributions of  $d_0$  and  $d_1$

All different scenarios are illustrated in Fig. 1. Their probabilities are calculated as

$$\Pr\{a1\} = p_{loss,T_{retx,1}} p_{del,T_{retx,0}}^0 = (1 - (1-p)F(T_{retx,1}))(1-p)F(T_{retx,0}). \quad (17)$$

$$\begin{aligned} \Pr\{a2\} &= (1-p)^2 \Pr\{d_0 - T_{TO} \leq d_1 < T_{retx,1}\} \\ &= (1-p)^2 \int_0^{T_{retx,1}} \left( \int_0^{d_1+T_{TO}} f(d_0) dd_0 \right) f(d_1) dd_1 \\ &= (1-p)^2 \int_0^{T_{retx,1}} F(t+T_{TO})f(t) dt. \end{aligned} \quad (18)$$

Similarly, we have

$$\Pr\{b1\} = p_{loss,T_{retx,0}} p_{del,T_{retx,1}}^0 = (1 - (1-p)F(T_{retx,0}))(1-p)F(T_{retx,1}). \quad (19)$$

$$\begin{aligned} \Pr\{b2\} &= (1-p)^2 \Pr\{d_1 + T_{TO} < d_0 < T_{retx,0}\} \\ &= (1-p)^2 \int_{T_{TO}}^{T_{retx,0}} \left( \int_0^{d_0-T_{TO}} f(d_1) dd_1 \right) f(d_0) dd_0 \\ &= (1-p)^2 \int_0^{T_{retx,1}} F(t)f(t+T_{TO}) dt. \end{aligned} \quad (20)$$

Note that the sum of Eqs. (18) and (20) is  $p_{del,T_{retx,0}}^0 p_{del,T_{retx,1}}^0 = (1-p)F(T_{retx,0}) \cdot (1-p)F(T_{retx,1})$ , the probability that both the original and the retransmitted messages are available at the cutoff time.

The resulting cdf of  $D_{(1)}^1$  is hence

$$\begin{aligned} F_{D_{(1)}^1}(t) &= \Pr\{D_{(1)}^1 \leq t\} = \\ &= \frac{\Pr\{\mathbb{I}_{D_{(1)}^1} = 0\} F_T^{T_{retx,0}}(t) + \Pr\{\mathbb{I}_{D_{(1)}^1} = 1\} F_T^{T_{retx,1}}(t - T_{TO})}{p_{del,T_{CO}}^1}, \end{aligned} \quad (21)$$

in which  $\Pr\{\mathbb{I}_{D_{(1)}^1} = 0\}$  and  $\Pr\{\mathbb{I}_{D_{(1)}^1} = 1\}$  are the sums of Eqs. (17) and (18), and Eqs. (19) and (20), respectively. From Eq. (9), we have the total delivery probability  $p_{del,T_{CO}}^1$  for “normalizing” the probabilities to obtain the cdf.

3)  $K_{retx} > 1$ : When  $K_{retx}$  is any number greater than one, thanks to the independence of different retransmissions, we can carry out the above pairwise comparison for any arbitrary pair

of arrivals. In fact, if we generalize the above results, we have for any pair  $i$  and  $j$  ( $i < j$ ) of effective retransmissions

$$\Pr\{D_i \leq D_j\} = (1-p)^2 \int_0^{T_{retx,j}} F(t + (j-i)T_{TO}) f(t) dt + (1 - (1-p)F(T_{retx,j})) (1-p)F(T_{retx,i}), \quad (22)$$

and

$$\Pr\{D_i > D_j\} = (1-p)^2 \int_0^{T_{retx,j}} F(t)f(t + (j-i)T_{TO}) dt + (1 - (1-p)F(T_{retx,i})) (1-p)F(T_{retx,j}). \quad (23)$$

Repeating for all possible pairs, we have the probability of  $D_k$  being the minimum, that is, the  $k$ -th retransmitted message is received first, as

$$\Pr\{\mathbb{I}_{D_{(1)}^{K_{retx}}} = k\} = \prod_{\substack{j=0 \\ j \neq k}}^{K_{retx}} \Pr\{D_k \leq D_j\} = (1-p) \cdot \prod_{\substack{j=0 \\ j \neq k}}^{K_{retx}} \left\{ (1-p) \int_0^{T_{retx, \max\{k,j\}}} F(t + \max\{0, (j-k)T_{TO}\}) \cdot f(t + \max\{0, (k-j)T_{TO}\}) dt + (1 - (1-p)F(T_{retx,j})) F(T_{retx,k}) \right\}. \quad (24)$$

This leads to the overall distribution of the  $d_{(1)}^{K_{retx}}$ :

$$F_{D_{(1)}^{K_{retx}}}(t) = \Pr\{D_{(1)}^{K_{retx}} \leq t\} = \frac{\sum_{k=0}^{K_{retx}} \Pr\{\mathbb{I}_{D_{(1)}^{K_{retx}}} = k\} F_T^{T_{retx,k}}(t - kT_{TO})}{p_{del,TCO}^{K_{retx}}}. \quad (25)$$

### C. Arrival Time: Two-way Communication Analysis

Eq. (25) describes the distribution of the earliest arrival time under the condition that all  $K_{retx}$  rounds of retransmissions are sent out after the original message. In reality, though, the total number of retransmissions can be anywhere from 0 to  $K_{retx}$ . In this subsection, we consider the two-way communications that determines the number of the actual retransmissions, which in turn affects the overall distribution of the earliest arrival time before the cutoff.

For satellite systems with conventional bent pipe type of transponders [11], one uplink (sensor  $\rightarrow$  satellite) and downlink (satellite  $\rightarrow$  FC) pair is used for the forward link, and the reverse link similarly consists of the uplink (FC  $\rightarrow$  satellite) and downlink (satellite  $\rightarrow$  sensor) pair. Depending on specific channel allocation schemes (e.g., TDMA- or FDMA-based), that is, whether the forward and reverse channels are assigned the same frequency band, the delay distribution of the ACK could vary from that of the messages<sup>2</sup>. Regardless, we have the

<sup>2</sup>Also the initial delay could be quite different too, owing to the usually much smaller size of the ACK messages.

pdf of the sum of two random delay values being expressed as the convolution of their respective pdfs:

$$h(t) = f(t) \star g(t) = \int_{t=0}^{\infty} f(u)g(t-u) du, \quad (26)$$

in which  $f$  and  $g$  are the distributions of the forward and reverse links, respectively. Meanwhile, if the ACK message is lost over the reverse link with a probability  $p_{ACK}$ , the overall probability that the ACK message can be eventually delivered is  $(1-p)(1-p_{ACK})$ , and its complement

$$p_T = 1 - (1-p)(1-p_{ACK}) \quad (27)$$

is the loss rate of the ‘‘super-message’’ that includes both the estimate message and ACK. With this loss rate and  $h(t)$  function, we can derive a general form of the arrival time.

1) *Probability of Having  $k_1$  ( $0 \leq k_1 \leq K_{retx}$ ) Retransmissions*: First, we have the trivial case in which  $K_{retx} = 0$  as  $T_{CO} \leq T_{TO}$ , then with probability one, there is no retransmission. Next we focus on the cases where  $K_{retx} \geq 1$ .

Having exactly  $k_1$  ( $0 \leq k_1 < K_{retx}$ ) retransmissions means that the earliest reception of the ACK message by the sensor occurs in the interval  $[k_1 T_{TO}, (k_1 + 1)T_{TO})$ . In other words, the first instance of the ACK arrival at the sensor

$$D_{T,(1)}^{k_1} = \min_{k=0, \dots, k_1} \{D_{T,k}\} \quad (28)$$

must satisfy

$$k_1 T_{TO} \leq D_{T,(1)}^{k_1} < (k_1 + 1)T_{TO}, \quad (29)$$

in which  $D_{T,(1)}^{k_1}$  is similarly defined as in Eq. (13), with the subscript  $T$  specifying that this is the arrival time accounting for the total delay from both forward and reverse links. Meanwhile, we define the delivery rate for the ‘‘super-message’’ described earlier as  $p_{T,del,t}^{K_{retx}}$  with the maximum number of retransmissions  $K_{retx}$ . Then we have

$$\Pr\{\text{There are exactly } k_1 \text{ retransmissions, } 0 \leq k_1 < K_{retx}\} = p_{T,del,(k_1+1)T_{TO}}^{K_{retx}} - p_{T,del,k_1 T_{TO}}^{K_{retx}}. \quad (30)$$

The delivery probabilities can be similarly calculated as in Eq. (9), with  $p$  being replaced by  $p_T$  and  $F(t)$  by  $H(t) = \int_0^t h(u) du$ , respectively.

On the other hand, having  $K_{retx}$  retransmissions means that none of the ACKs have been received by  $K_{retx}T_{TO}$ , and we have

$$\Pr\{\text{There are exactly } K_{retx} \text{ retransmissions}\} = \prod_{k=0}^{K_{retx}-1} \Pr\{D_{T,k}^{K_{retx}} > K_{retx}T_{TO}\} = \prod_{k=0}^{K_{retx}-1} [1 - (1-p_T)H((K_{retx}-k)T_{TO})]. \quad (31)$$

2) *Distribution of the Arrival Time for a Given  $T_{CO}$* : With  $K_{retx}$  in Eqs. (9) and (24) being replaced by any  $k_1$  ( $0 \leq k_1 \leq K_{retx}$ ), we can easily find the delivery rate  $p_{del,TCO}^{k_1}$  after  $k_1$  rounds of retransmissions and the probability  $\Pr\{\mathbb{I}_{D_{(1)}^{k_1}} = k\}$  that the  $k$ -th retransmission marks the earliest arrival among

all  $k_1 + 1$  sent out messages. And then we have the cdf with exactly  $k_1$  retransmissions as

$$F_{D_{(1)}^{k_1}}(t) = \Pr\{D_{(1)}^{k_1} \leq t\} = \frac{\sum_{k=0}^{k_1} \Pr\{\mathbb{I}_{D_{(1)}^{k_1}} = k\} F_T^{T_{retx},k}(t - kT_{TO})}{p_{del,TCO}^{k_1}}. \quad (32)$$

Finally, we can combine Eqs. (30), (31), and (32) to obtain the distribution of the arrival time for any given  $T_{CO}$ :

$$F_{D_{(1)}}(t) = \sum_{k_1=0}^{K_{retx}} F_{D_{(1)}^{k_1}}(t) \Pr\{\text{There are } k_1 \text{ retransmissions}\}. \quad (33)$$

#### IV. ANALYTICAL AND SIMULATION STUDIES OF ESTIMATE FUSION PERFORMANCE

Having studied the message recovery process, we demonstrate the effectiveness of the message retransmission mechanism in improving the estimation accuracy by (1) analysis of the mean square error (MSE) performance with message retransmission and time cutoff; and (2) simulations of the tracking of a coasting ballistic target. In particular, we are interested in exploring the trade-offs between estimation accuracy and latency.

##### A. Target Model

The states of a coasting ballistic target are generated using the following state-space model [8]:

$$\dot{\mathbf{x}} \triangleq \begin{bmatrix} \dot{\mathbf{p}} \\ \dot{\mathbf{v}} \end{bmatrix} = \mathbf{f} \left( \begin{bmatrix} \mathbf{p} \\ \mathbf{v} \end{bmatrix} \right) \triangleq \begin{bmatrix} \mathbf{v} \\ \mathbf{a}_G(\mathbf{p}) \end{bmatrix}. \quad (34)$$

The target state vector  $\mathbf{x} = \begin{bmatrix} \mathbf{p}^T & \mathbf{v}^T \end{bmatrix}^T$ , where  $\mathbf{p} = \begin{bmatrix} x & y & z \end{bmatrix}^T$  and  $\mathbf{v} \triangleq \dot{\mathbf{p}} = \begin{bmatrix} \dot{x} & \dot{y} & \dot{z} \end{bmatrix}^T$  are the target position and velocity vectors, respectively.  $\mathbf{a}_G(\mathbf{p})$  is the gravitational acceleration under the spherical Earth model [8]:

$$\mathbf{a}_G(\mathbf{p}) = -\frac{\mu}{p^2} \mathbf{u}_p = -\frac{\mu}{p^3} \mathbf{p}, \quad (35)$$

where  $\mathbf{p}$  is the vector from the Earth's center to the target,  $p \triangleq \|\mathbf{p}\|$  is its length,  $\mathbf{u}_p \triangleq \mathbf{p}/p$  is the unit vector in the direction of  $\mathbf{p}$ , and  $\mu = 3.986012 \times 10^5 \text{ km}^3/\text{s}^2$  is the Earth's gravitational constant. The algorithm used for state propagation can be found in [13]. The initial target state is [7]:  $[113.75 \ 3950 \ 5150 \ 0.94 \ 3.33 \ -6.0125]^T$ , in which the position and velocity values are in the units of km and km/s respectively.

##### B. Sensor Profiles

A total of  $N = 5$  sensors are deployed for reporting their state estimates of the dynamic target defined above. A state estimate  $\hat{\mathbf{x}}_i(k)$  from sensor  $i$  at time  $k$  is generated by adding random Gaussian noise to the true states:

$$\hat{\mathbf{x}}_i(k) = \mathbf{x}(k) + \mathbf{n}_i(k)$$

where  $\mathbf{x}(k)$  is the true target state for a target at time  $k$ , and  $\mathbf{n}_i(k) \sim \mathcal{N}(\mathbf{0}, \Sigma)$ .  $\Sigma$  is a diagonal matrix:

$$\Sigma = \text{diag} \left( \left[ \sigma_x^2 \ \sigma_y^2 \ \sigma_z^2 \ \sigma_{\dot{x}}^2 \ \sigma_{\dot{y}}^2 \ \sigma_{\dot{z}}^2 \right] \right)$$

$\sigma_x^2$ ,  $\sigma_y^2$ , and  $\sigma_z^2$  are the position error variances, and  $\sigma_{\dot{x}}^2$ ,  $\sigma_{\dot{y}}^2$ , and  $\sigma_{\dot{z}}^2$  are the velocity error variances. The following state estimation errors are set commonly for all sensors:  $\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = 1$  and  $\sigma_{\dot{x}}^2 = \sigma_{\dot{y}}^2 = \sigma_{\dot{z}}^2 = 10^{-4}$ .

##### C. Noise Profiles

We consider the bias and correlation characteristics of the process noise that could potentially increase estimation errors. Different combinations include (1) unbiased and uncorrelated noise; (2) biased and uncorrelated noise; (3) unbiased and correlated noise; and (4) biased and correlated noise. The random position bias for each axis is uniformly taken from the range  $[-0.5, 0.5]$ , and the velocity bias in the same manner from  $[-0.005, 0.005]$ . In the cases with correlated noise, the state estimation errors defined above (for the uncorrelated scenario) are split into 80% uncorrelated noise and 20% correlated noise in each dimension.

##### D. Fusion Rule

We apply the linear fuser defined as follows:

$$\mathbf{P}_F = \left( \sum_{i=1}^L \mathbf{P}_i^{-1} \right)^{-1}, \text{ and } \hat{\mathbf{x}}_F = \mathbf{P}_F \sum_{i=1}^L \mathbf{P}_i^{-1} \hat{\mathbf{x}}_i, \quad (36)$$

where  $\hat{\mathbf{x}}_F$  is the fused estimate and  $\mathbf{P}_F$  is its error covariance matrix.  $\mathbf{P}_i$  and  $\hat{\mathbf{x}}_i$  are similarly defined for sensor  $i$ . A total of  $L$  ( $L \leq N$ ) state estimates are combined at the fusion center. This simple fuser is a special form of the track-to-track fuser [3] where the process noise is zero.

##### E. Communication Link Statistics

The default forward link loss rate is set to be  $p = 0.5$ , compared to that of the reverse link  $p_{ACK} = 0.1$ . The arrival delay of both directions satisfies the shifted exponential distribution defined in Eq. (1), with  $\mu_F = 0.3$  s and  $\mu_R = 0.2$  s for the forward and reverse links respectively and a common  $T = 0.5$  s. The default  $T_{TO}$  and  $W$  are set to be 1.5 s and 4.5 s respectively, which are both multiples of the measured RTT at 0.75 s.

##### F. MSE analysis

From Eq. (36), the MSE of the position and velocity estimates of the linearly fused estimate satisfies

$$|\hat{\mathbf{p}}_F - \mathbf{p}|^2 = |\hat{\mathbf{p}}_i - \mathbf{p}|^2/L, \text{ and } |\hat{\mathbf{v}}_F - \mathbf{v}|^2 = |\hat{\mathbf{v}}_i - \mathbf{v}|^2/L, \quad (37)$$

if the sensors have the same estimation error profiles (that is, the same MSEs). Suppose the system imposes its maximum tolerable errors of position and velocity estimates as  $MSE_{max,\hat{\mathbf{p}}}$  and  $MSE_{max,\hat{\mathbf{v}}}$  respectively, with independent transmissions from up to  $N$  sensors, the minimum delivery probability is given by

$$p_{del,min} = \frac{L_{min}}{N} = \frac{1}{N} \max \left\{ \frac{|\hat{\mathbf{p}}_i - \mathbf{p}|^2}{MSE_{max,\hat{\mathbf{p}}}}, \frac{|\hat{\mathbf{v}}_i - \mathbf{v}|^2}{MSE_{max,\hat{\mathbf{v}}}} \right\}, \quad (38)$$

in which  $L_{min}$  is the minimum number of estimates that should be delivered. The actual delivery probability from message retransmission as expressed in Eq. (9) must be checked against this minimum delivery rate to ensure the MSE requirements are met.

The distribution of the arrival time, as in Eq. (33), can be used to derive an upper bound of the estimation error. For a given cutoff time  $T_{CO}$ , when the fusion center decides to finalize the estimate at an earlier time  $t$ , according to Eq. (36), the expected contribution of the position estimate from the sensor  $i$  is lower bounded by  $\frac{\alpha}{|\hat{\mathbf{p}}_i - \hat{\mathbf{p}}|^2}$ , in which  $\alpha = F_{d(i)}^{T_{CO}}(t)$ . The other part, weighted by  $1 - \alpha$ , depends on how the fuser chooses to substitute the missing estimate, such as using one- or multi-step prediction. If we consider the error performance of the fused estimate, the upper bound – as in the worst-case scenario in which a missing estimate results in zero contribution – is

$$|\hat{\mathbf{p}}_F(t) - \mathbf{p}|_{max}^2 = \frac{|\mathbf{p}_F(T_{CO}) - \mathbf{p}|^2}{F_{D(i)}^{T_{CO}}(t)}, \text{ for } 0 < t \leq T_{CO}, \quad (39)$$

in which  $\mathbf{p}_F(T_{CO})$  is the fused position estimate at time  $T_{CO}$ . A similar result can be found for the velocity estimate.

### G. Simulation Results and Analysis

Monte Carlo simulations were run for each case. Unless otherwise specified, the cutoff time is set as  $T_{CO} = 5$  s and the other parameters take the default values as previously defined.

1) *Timeout  $T_{TO}$  and Cutoff  $T_{CO}$* : Figs. 2 and 3 demonstrate the effect of different retransmission timeout and fusion cutoff. There is no retransmission when  $W = T_{TO} = 4.5$  s. When  $T_{TO}$  is set to be 1.5 s, however, two rounds of retransmissions can effectively reduce the MSE of the estimate. For example, when the loss rate is 60%, the error is reduced by more than 50%. Likewise, for a given  $T_{TO}$ , when the reporting deadline requirement is tightened, as reflected by decreasing cutoff time  $T_{CO}$ , the estimation MSE will increase given the same loss rate. On the other hand, when read horizontally, the plots indicate that to meet the same MSE requirement, with increasing loss rates,  $T_{TO}$  should be reduced and/or  $T_{CO}$  should be increased. A good rule of thumb to determine the MSE performance is the ratio  $T_{CO}/T_{TO}$ , although this rule may at times fail due to the periodicity of the retransmission process, especially when the values being compared are close.

2) *Noise Profiles*: We also simulated the position MSE performance with different noise profiles. From Fig. 4, we observe that correlated and/or biased noise generally leads to increased estimation error; although in our case, the zero-mean random bias is not large enough to significantly degrade the accuracy and the associated curves are fairly close to the ones without bias. Generally, the higher the absolute mean of the bias and the proportion of the process noise that is common to all sensors, the worse the fused estimate will be in terms of accuracy.

3) *Retransmission Performance*: In Fig. 5, we plotted the proportion of different numbers of retransmissions with respect to various loss rates. Note that the last group (labeled as “2”) does not indicate the message will be delivered within this round, but rather this is the last try as  $K_{retx} = 2$ . As expected,

an increased message-level loss rate requires more rounds of retransmissions so that the message can be recovered with the same probability over a longer period of time.

4) *Upper Bound of the MSEs From Early Cutoff*: Finally, in Figs. 6 and 7, the upper bounds of the MSEs resulting from earlier-than-scheduled cutoff are shown. The singular points in these plots indicate the arrival of a new round of retransmission. The concavity of the bounds (excluding the singular points) implies that the “best” time for early fusion is roughly in the middle of each round (accounting for the link delay), where the deepest descent in this round has occurred. In addition, as time inches closer to the cutoff time  $T_{CO}$ , the bound also approaches the actual MSE obtained when the cutoff time is set at that point, and is a good indicator of the actual MSE performance.

## V. CONCLUSION

In this paper, we studied the message retransmission performance in long-haul state estimate fusion applications. In particular, accounting for imperfect bi-directional communications over the long-haul satellite link, we derived the message delivery probability, the arrival time distribution, and the estimation error performance for the linear fuser under different estimation error profiles. Simulation results of a coasting ballistic target demonstrate the effectiveness of the retransmission mechanism and provide us insight into the trade-off between the estimation accuracy and latency.

## REFERENCES

- [1] M. Allman, C. Hayes, H. Kruse, and S. Ostermann. Tcp performance over satellite links. In *Proceedings of Fifth International Conference on Telecommunications Systems*, Nashville, TN, Mar. 1997.
- [2] R. B. Bapat and M. I. Beg. Order statistics for nonidentically distributed variables and permanents. *Sankhy: The Indian Journal of Statistics, Series A*, 51(1):79–93, Feb. 1989.
- [3] Y. Bar-Shalom, P. K. Willett, and X. Tian. *Tracking and Data Fusion: A Handbook of Algorithms*. YBS Publishers, 2011.
- [4] W. Boord and J. B. Hoffman. *Air and Missile Defense Systems Engineering*. CRC Press, 2012.
- [5] A. Chiuso and L. Schenato. Information fusion strategies and performance bounds in packet-drop networks. *Automatica*, 47:1304–1316, Jul. 2011.
- [6] Thomas R. Henderson and Randy H. Katz. Transport protocols for internet-compatible satellite networks. *IEEE Journal on Selected Areas in Communications*, 17:326–344, 1999.
- [7] T. H. Kerr. Streamlining measurement iteration for ekf target tracking. *IEEE Transactions on Aerospace and Electronic Systems*, 27(2):408–421, Mar. 1991.
- [8] X. R. Li and V. P. Jilkov. Survey of maneuvering target tracking. part ii: Motion models of ballistic and space targets. *Aerospace and Electronic Systems, IEEE Transactions on*, 46(1):96–119, Jan. 2010.
- [9] Q. Liu, X. Wang, and N. S. V. Rao. Fusion of state estimates over long-haul sensor networks under random delay and loss. In *Proceedings of 31st IEEE International Conference on Computer Communications (INFOCOM 2012), Mini-conference track*, pages 2968–2972, Orlando, FL, Mar. 2012.
- [10] A. Papoulis and S. U. Pillai. *Probability, random variables, and stochastic processes*. McGraw-Hill series in electrical and computer engineering. McGraw-Hill, 2002.
- [11] D. Roddy. *Satellite Communications*. McGraw-Hill, 2006.
- [12] D. Simon. *Optimal state estimation: Kalman, H [infinity] and nonlinear approaches*. Wiley-Interscience, 2006.
- [13] M. Yeddapanudi, Y. Bar-Shalom, K. R. Pattipati, and S. Deb. Ballistic missile track initiation from satellite observations. *IEEE Transactions on Aerospace and Electronic Systems*, 31(3):1054–1071, Jul. 1995.

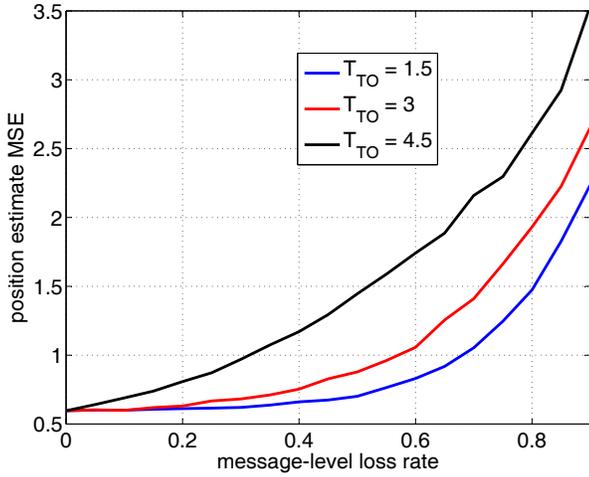


Fig. 2. Position MSE (km<sup>2</sup>) versus message-level loss rates with  $T_{CO} = 5$  s and different  $T_{TO}$

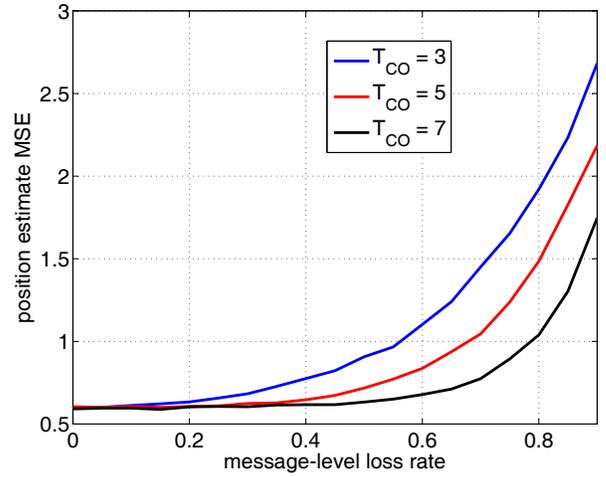


Fig. 3. Position MSE (km<sup>2</sup>) versus message-level loss rates with  $T_{TO} = 1.5$  s and different  $T_{CO}$

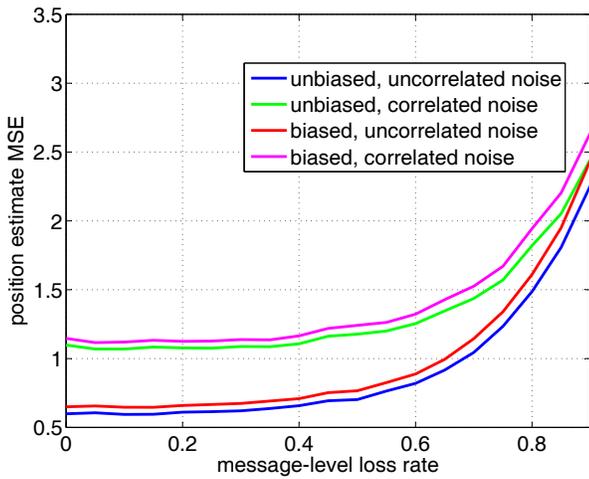


Fig. 4. Position MSE (km<sup>2</sup>) versus message-level loss rates with different noise profiles

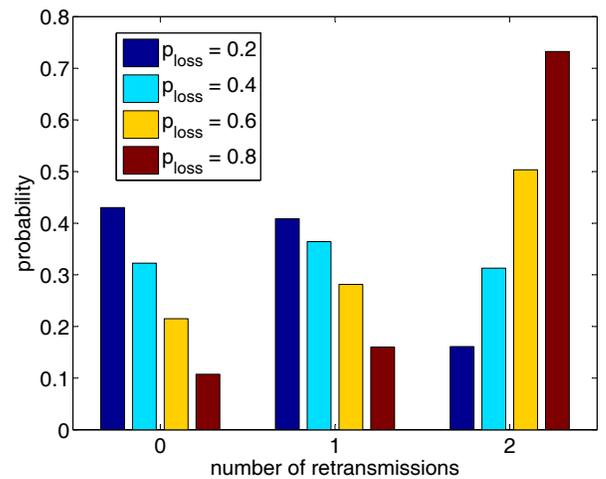


Fig. 5. probability of different number of retransmissions with different loss rates

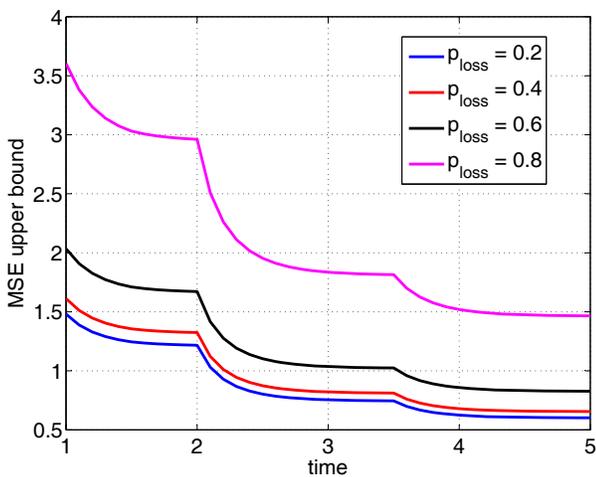


Fig. 6. Upper bound of position MSE (km<sup>2</sup>) with different message-level loss rates

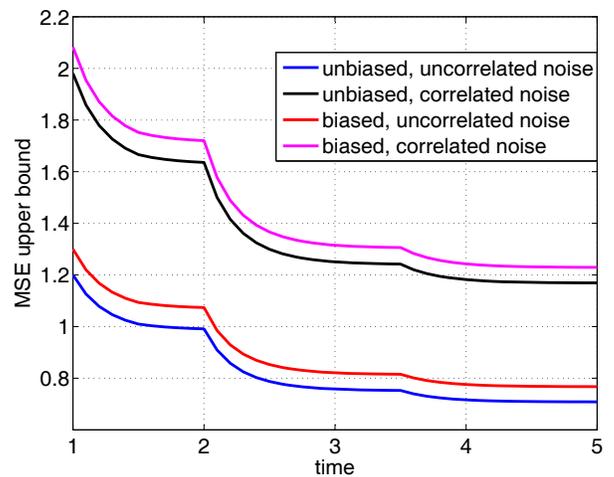


Fig. 7. Upper bound of position MSE (km<sup>2</sup>) with different message-level loss rates: random bias