A novel Markov Model for the computation of the continuity risk in Maritime applications

Giulio Franzese [1*], Ilaria Martini [2], Letizia Lo Presti[1]

[1]Dipartimento di Elettronica e Telecomunicazioni, Politecnico di Torino
Corso Duca degli Abruzzi 24, 10129, Torino, Italy
[2]Deutsches Zentrum fur Luft- und Raumfahrt, DLR
Munchener Strasse 20 , 82234, Wessling, Germany

*E-mail: giulio.franzese@studenti.polito.it

BIOGRAPHY

Giulio Franzese is a Telecom engineer graduate student at the Politecnico di Torino (Italia). During his undegraduate degree he started a collaboration with Prof. Letizia Lo Presti in the research field of signal processing applied to GNSS navigation. He finished his master thesis at DLR (Oberpfaffenhofen) in November 2016 under the supervision of Prof. Letizia Lo Presti and Dr. Ilaria Martini (DLR) and is currently working at the Politecnico di Torino as a researcher.

ABSTRACT

In this paper we focused on the analysis of the continuity risk for a maritime user and on the derivation of the receiver implementation scheme that fulfills IMO [2] requirements. We started our analysis by considering the model derived in [6], which is accurate for aviation applications, while it is not guaranteed to work in other environments. To take into account the time evolution of the continuity risk, we propose in this paper to introduce Markov models. The derived models can be used to compute the continuity risk when a receiver not implementing exclusion is used, as well as when a receiver implements both snapshot and sequential exclusion. The derived conclusion is that without exclusion, it is not possible to achieve the required performance. It is shown that, considering only satellite faults, over the 3 hours of operation, in principle a snapshot exclusion mechanism is sufficient. When implementing sequential exclusion, at the cost of an increased complexity, the continuity risk is furthermore reduced. We also showed that, under some temporal limitations and with some assumptions on the exclusion mechanism, the results provided by the Markov models and by the model derived in [6] coincide. This final consideration shows that the proposed approach can be seen as a natural extension of the state of the art model from the avionic environment to the maritime environment.

INTRODUCTION

In the Maritime navigation environment it is under the attention of the scientific and technological community the possibility of standardizing the use of multiconstellation navigation systems that are compliant with the IMO (International Maritime Organization) requests. Preliminary studies on the required performance show that one of the critical issues is the requirement related to continuity. According to the definition taken from [1]:

"The Continuity of a system is the ability of the total system (comprising all elements necessary to maintain aircraft position within the defined airspace) to perform its function ***without interruption*** *during the intended operation. More specifically, Continuity is the probability that the specified system performance will be* ***maintained*** *for the duration of a phase of operation, presuming that the system was available at the beginning of that phase of operation and was predicted to operate throughout the operation. "*

The continuity requirement for the maritime user is of 99.97% (0.9997) over 3 hours (10800s) of continuous operation. A preliminary attempt to criticism of the requirement imposed by the IMO in the maritime navigation environment was presented by Klepvsvik et al. in [3], where a simplified model of the system performance is proposed, taking into account the presence of possible failures of one or more satellites during the three hours of the maritime operations. The reached conclusion is that, at the current state of the system, the requirements are fulfilled only for some of the operational phases.

One of the fundamental questions the researcher communities are trying to address is whether exclusion of faulty satellites is necessary or not. In this paper we answer to this question. In particular we evaluate the probability of loss of continuity with and without exclusion, by adopting a new mathematical model of the mechanism of integrity monitoring. The model is quite general and can be applied to different application scenarios. When applied to the maritime environment the conclusion we obtain is that exclusion is necessary. We started our analysis by considering the model derived in [6], which is accurate for aviation applications, while it is not guaranteed to work in other environments. Infact, while in the avionic environment some assumptions can be made without degrading to much the ability of the model to describe the real world scenario, in the maritime applications, instead, these assumptions are too restrictive, mostly because of the different timescales (15s in avionic vs 3 hours in maritime applications). One of the key aspect of this work is the analysis of the difference between the true definition of Continuity (where the time evolution of the system is considered) and the popular average sense approximation (done in avionic environment), that leads to inaccurate estimations of the continuity risk when used in the maritime environment.

To take into account the time evolution of the continuity risk, we propose in this paper to introduce Markov models of the system state, since they are realistic tools for analyzing the system performance in terms of probability of loss of continuity, using as input parameters the number of satellites and the probabilities related to the detection and exclusion mechanism. The derived models can be used to compute the continuity risk when a receiver not implementing exclusion is used, as well as when a receiver implements both snapshot and sequential exclusion. The derived conclusion is that without exclusion, it is not possible to achieve the required performance. It is shown that, considering only satellite faults, over the 3 hours of operation, in principle a snapshot exclusion mechanism is sufficient. However the required efficiency of the exclusion mechanism is extremely high and if we consider also other error sources, such as multipath, clock drifts, scintillations, that have much higher rates of occurrences with respect to the satellite faults (the mean time between failures for a single satellite is roughly 10 years) snapshot exclusion is not sufficient. A possible alternative solution is a sequential exclusion mechanism.

Moreover, it can be shown that, under some temporal limitations and with some assumptions on the exclusion mechanism, the results provided by the Markov models and by the model derived in [6] coincide. This final consideration shows that the proposed approach can be seen as a natural extension of the state of the art model from the avionic environment to the maritime environment.

CONTINUITY RISK ANALYSIS IN AVIONIC ENVIRONMENT

As already mentioned in the intoduction, the model derived in [6] is accepted as simple and effective in describing the true continuity risk for the avionic environment. In this work, a set of assumptions and approximations are made and, while for the avionic user this model is good for describing the Continuity performance, in maritime environment, instead, these assumptions are too restrictive mostly because of the different timescales (15s in avionic vs 3 hours in maritime applications). Moreover, as ICAO suggests, in the avionic environment the Continuity problem is solved in an average sense. To introduce this "average sense" concept, let us suppose that a certain standard requires that, using a standard compliant receiver, the probability ($\overline{P}$) of a certain event to occur (for example, losing continuity or integrity) is fixed over a period T (a measurement frequency, specified in Hertz, f_m is also provided). In this standard it is thus implicitly assumed that the receiver has to fulfill the requirements considering a sequence of $N = Tf_m$ consecutive measurements, with an interval between measurements of $\frac{1}{f_m}$ (in avionic environment, for example, $N = 15$ and $f_m = 1Hz$). The receiver designer knows that the probability P of the occurrence of the considered event is a function of both the receiver characteristics (the design space) and the considered time window, i.e. $P = f(\text{characteristics}, N)$. The function $f(\cdot)$, that relates the time window and the characteristics to the probability of event occurrence, can be analytically very complex and hardly numerically computable. The problem is solved using an average sense interpretation if the following design procedure is applied: instead of designing the system such that

$$P = f(\text{characteristics}, N) \leq \overline{P}$$

we can instead design the system constrained to

$$P_{inst} = f(\text{characteristics}, 1) \leq \overline{P}_{inst} = \frac{\overline{P}}{N}$$

Basically the requested probability over the period T is scaled to an equivalent single measurement requested probability P_{inst}. The receiver design considering the function $f(\text{characteristics}, 1)$ is generally greatly simpler than the design analyzing $f(\text{characteristics}, N)$.

With this assumption the calculations are greatly simplified but the the ability to well represent the original problem can be degraded. In fact continuity is defined over a continuous time window, and not in an average sense. We will see in a successive section that while the average sense approximation is reasonable for a 15 seconds time window, it is not adequate when dealing with a 10800 seconds, as in the maritime environment.

With this interpretation of the average sense approach we can now analyze the model derived in [6]. The computation of the continuity risk presented in [6] is based on a set of assumptions and on the average time between failures of satellites and the average time to alert (so making the fault not dangerous from an integrity point of view). The considered fault rate is assumed to be of 3 SV faults per year (with a total of 24 satellites) and Mean Time to Alert (MTTA) of 1 hour. The set of assumptions is the following:

- all 3 faulted satellites are visible to the aircraft when the fault occurs
- all 3 faults are detected during an approach (they affect continuity)
- once a fault is detected, it will rarely impact the aircraft again

The probability of loss of Continuity (LOC) due to a detected SV fault in the time window of 15 seconds (avionic) computed in [6] using the average sense interpretation is then:

$$P_{F_i} = \frac{3}{24SV \times 1year} \times 15sec = 6 \times 10^{-8}/SV \tag{1}$$

We propose now a different derivation of the continuity risk. Notice that the purpose of this different formulation is not to invalidate the model for the avionic environment, where the assumptions are reasonable, but only to set a basis for understanding the need of a different model when considering maritime operations. We start by considering that, if we have 3 satellite faults per year, and a total of 24 satellites considered for this statistics, we can derive that a satellite breaks approximately every $M_{TBF} = 8$ years (the failure rate per satellite is thus $R_{TBF} = \frac{3}{24 \times year} = 3.96 \times 10^{-9} Hz$). However, in GNSS applications we are not directly interested in the failure rate, but rather in the fraction of the total time in which a single satellite is in Faulty state. Since the mean time to alert is 1 hour ($M_{TTA} = 3600s$), the user will receive the faulty satellite for one hour (on the average). Therefore the average percentage of time in which a satellite is in a faulty state is

$$\eta = \frac{M_{TTA}}{M_{TTA} + M_{TBF}} \simeq \frac{M_{TTA}}{M_{TBF}} = 3.96 \times 10^{-9} Hz \times 3600s = 1.42 \times 10^{-5} \tag{2}$$

The probability of losing Continuity due to a detected satellite fault is thus 1.42×10^{-5}. Having the probability of the satellite to be in Faulty state in a single time instant, for computing the continuity risk, using the average sense interpretation, we have to multiply η by a factor 15. The result is that the computed continuity risk due to a single satellite detected failure, that is $P_{LOC} = 2.14 \times 10^{-4}$, is different from the value computed in (1). This difference is due to an important assumption made in the derivation of (1): when a fault occurs it is immediately detected and from the successive time instant it will not cause anymore a loss of Continuity. This is mainly justified by the authors due to the fact that the pseudorange error introduced by satellite faults is generally monotonically increasing in time, and thus a fault detected cannot cause in a subsequent time instant a loss of Continuity. This is clearly depicted in (1) where the number of harmful instants due to a detected satellite fault is considered to be equal to one. In fact we can interpret (1) as

$$P_{F_i} = \underbrace{\frac{3}{24SV \times 1year}}_{\text{fault rate}} \times \overbrace{1}^{\text{number of harmful instants}} \times \underbrace{15s}_{\text{scaling factor}}$$

where it is underlined the fact that the number of harmful instants for the receiver, due to a satellite fault, is one: if we are assuming that only the first time instant in which a fault occurs can cause a loss of Continuity this is the case. In this work we moved away from this simplified analyisis, because it is not accurate for the maritime application environment, since no time evolution of the continuity risk is considered in the computation of the continuity equations.

MARKOV MODELS

The first motivation behind the derivation of the Markovian model is the necessity to have a continuity risk model for maritime application where continuity is defined on long time intervals, in which also the time evolution of the continuity risk is considered. With a solid mathematical background, the model is used to compute the continuity risk over the 3 hours as a function of the average number of in view satellites, the detection probability and, if exclusion is implemented, the exclusion mechanism performance. The derived models will be used in the next Section to argue about the necessity of implementing exclusion. Moreover, the new models will be compared with the model derived in [6].From this comparison we will see that the model [6] is a particular case of one of the new models, under some temporal limitations and with some assumptions on the exclusion mechanism.

MARKOV MODEL: SATELLITE DOMAIN

The first simple model we introduce describes a single satellite behavior and consists of a two states Markov Chain (Healthy and Faulty) where the transition is ruled by the Mean Time Between Failures (MTBF) and the Mean Time To Repair (MTTR). The starting point of our derivation is to analyze the random exiting time from the two states (H =Healthy, F =Faulty) for the single satellite. The exiting time of a state is defined as the time difference between the instant in which the satellite enters a given state and the time instant in which it exits that state. We assume the two random time intervals to be exponentially distributed (a common and reasonable assumption in system fault analysis [4]) with a given mean (MTBF and MTTR). Defining the random interval from the Healty to the Faulty state transition as T_H and the random interval from Faulty to Healthy as T_F their probability density functions (pdfs) are defined as follows

$$\begin{cases} f_{T_H}(t) = \lambda \exp(-\lambda t) u(t) \\ f_{T_F}(t) = \mu \exp(-\mu t) u(t) \end{cases} \tag{3}$$

where $\lambda = \frac{1}{MTBF}$ is the fault rate and $\mu = \frac{1}{MTTA}$ is the alert rate. These rates have the following numerical values: $\lambda = \frac{3SV}{24SV\ \ year} = 3.96 \times 10^{-9}$Hz and $\mu = \frac{1}{hr} = 2.77 \times 10^{-4}$Hz. We can derive the equivalent discrete time Markov chain that is the result of observing the continuous time system (the real satellite states evolve continuously with time) with a fixed interval of T seconds (in our case $T = 1s$).

Finding the discretized version of a continuous Markov chain is, per se, a field of study. However, when the continuous time transition rates are order of magnitude greater than the observation rates a simplified approach can be taken, where we derive the discrete time transition probabilities simply integrating the continuous time transition probabilities over a period equal to T, thus neglecting the probability of multiple transitions in a single second. We consider for example the derivation of the probability of transiting from a Healthy state in a given time instant t_0 to the Faulty state at time instant $t_0 + T$. Defining $\Phi(T)$ as the number of transitions from one state in the time interval $(t_0, t_0 + T)$, given that at time t_0 the system is in an Healthy state, the probability that at time $t_0 + T$ the system is in a Faulty state is

$$P(F(t_0+T)|H(t_0)) = P(\bigcup_{n=0}^{\infty} \text{2n+1 transitions in } (T)) = P(\bigcup_{n=0}^{\infty} \Phi(T) = 2n+1)$$

$$= \sum_{n=0}^{\infty} P(\Phi(T) = 2n+1) - \sum_{n_1<n_2} P(\Phi(T) = 2n_1+1, \Phi(T) = 2n_2+1) - \sum_{n_1<n_2<n_3} \ldots$$

Since the probability of the intersections is equal to 0 ($\Phi(T)$ cannot have simultaneously two different values), we have that

$$P(F(t_0+T)|H(t_0)) = \sum_{n=0}^{\infty} P(\Phi(T) = 2n+1) \simeq P(\Phi(T) = 1)$$

where the last approximation is done because, given our system transition rates, we have that

$$\forall n > 0 \quad P(\Phi(T) = 2n+1) << P(\Phi(T) = 1)$$

This approximation can be mathematically proved but it is also heuristically easily understandable considering the following: if the mean time between breaks is roughly 10 years and the mean time to repair is 1 hour, the probability of having three transitions ($H \to F, F \to H, H \to F$) is a random variable with mean equal to roughly 20 years (10 years+1hour +10 years), and it is easily understandable that this probability is completely negligible with respect to the single transition probability. In the same way the probability of having 5,7,9,... transition is infinitely smaller than the probability of having a single transition. A specular consideration can be done also for the transition sequence ($F \to H, H \to F, F \to H$).

The probability of having a single transition from healthy to faulty state in the time interval $[0, T]$ is equal to the probability that the exiting time T_H from the healthy state occurs in the interval $[0, T]$ (since T_H is positive it is equivalent to say that $T_H < T$) and that the system exits the faulty state after T, that is equivalent to say that the exiting time from the faulty state T_H must be greater than $T - T_H$. Given that $f_{T_H}(t) = \lambda \exp(-\lambda t)u(t)$ and that $f_{T_F}(t) = \mu \exp(-\mu t)u(t)$ and that the probability $P(\Phi(T) = 1)$ can be expressed as

$$P(\Phi(T) = 1) = P(T_H < T, T_F > T - T_H) = P(T_H < T, T_H + T_F > T) = \int_{t_1=0}^{T} \int_{t_2=T-t_1}^{\infty} f_{T_H}(t_1) f_{T_F}(t_2) dt_2 dt_1 = \frac{\lambda}{\mu - \lambda}(\exp(-\lambda T) - \exp(-\mu T)) \tag{4}$$

We can finally derive the discrete transition probability from Healthy to Faulty state p_{01} as

$$p_{01} = P(F(t_0 + T)|H(t_0)) \simeq P(\Phi(T) = 1) = \frac{\lambda}{\mu - \lambda}(\exp(-\lambda T) - \exp(-\mu T)) \simeq \lambda T \tag{5}$$

Figure 1 shows the Markov chain rappresentation of the system state with their two possibile states for a single satellite (Healthy and Faulted). The matrix transition rate of the system is defined as

$$\boldsymbol{T} = \begin{bmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{bmatrix} \tag{6}$$

where it is possible to approximate the four probabilities as

$$\begin{cases} p_{00} \simeq 1 - \lambda T \\ p_{01} = 1 - p_{00} \simeq \lambda T \\ p_{10} = 1 - p_{11} \simeq \mu T \\ p_{11} \simeq 1 - \mu T \end{cases} \tag{7}$$

Moreover, solving the following equation (the steady state equation of the system)

$$\boldsymbol{\pi} = \boldsymbol{\pi} \boldsymbol{T} \tag{8}$$

it is possible to derive the steady state probability distribution of the system (where $\boldsymbol{\pi} = [\pi_0 \ \pi_1]$ are the steady state probabilities for Healthy and Faulty state respectively). Rearranging (8) as

$$\boldsymbol{\pi}(\boldsymbol{I} - \boldsymbol{T}) = \boldsymbol{0}$$

we can solve for π_0 through the following system of equations

$$\begin{cases} \pi_0(p_{00} - 1) + \pi_1 p_{10} = 0 \\ \pi_1 = 1 - \pi_0 \end{cases} \tag{9}$$

where the second equation is due to the total law probability theorem ($\pi_0 + \pi_1 = 1$). The derivation is the following:

$$\begin{aligned} &\pi_0(p_{00} - 1) + (1 - \pi_0)p_{10} = 0 \\ &\pi_0(p_{00} - 1 - p_{10}) = -p_{10} \\ &\pi_0 = \frac{-p_{10}}{p_{00} - 1 - p_{10}} \end{aligned}$$

and, finally,

$$\pi_0 = \frac{p_{10}}{p_{10} + 1 - p_{00}} \simeq \frac{\mu}{\mu + \lambda} \simeq 1 - \frac{\lambda}{\mu} \tag{10}$$

$$\pi_1 \simeq \frac{\lambda}{\mu} \tag{11}$$

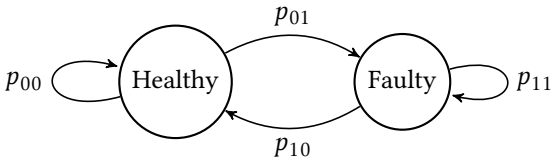


Figure 1: Single satellite model

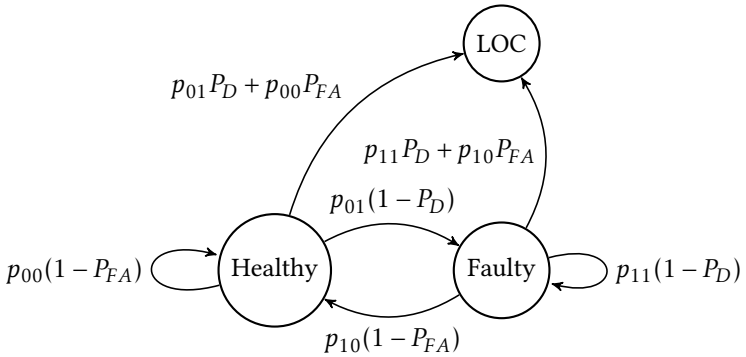


Figure 2: Single Satellite and Receiver model

MARKOV MODEL: POSITION DOMAIN

Until now, we have just considered the probabilistic evolution of the state of a single satellite, regardless the presence of a receiver. The next step is to include in our model the receiver behaviour for calculating the continuity risk.

USER RECEIVER AND SINGLE SATELLITE (NO EXCLUSION)

The first simple model we consider is a single satellite system with a receiver that implements fault detection but not exclusion. It is obvious that such a system cannot exist because of practical reasons (we need at least $3+M$ satellites to have a positioning, where M is the number of constellations), but it is useful for introducing the problem. A receiver implementing detection but not exclusion when detects a satellite fault in the position domain declares the system as not available. This leads the system to a Loss of Continuity (LOC) state. It is crucial to understand that the LOC state is an absorbing state: in fact if we consider continuity in its true sense, and not in the average sense, we understand that once continuity is lost during our operation time window, it is lost for all the mission. This is intrinsically hidden in the meaning of the word "Continuity": the system is defined continuous if and only if for all the mission the system is continuously declared as available.

We can introduce now the three state (Healthy, Faulty, LOC) model, shown in Figure 2, where p_{00}, p_{01}, p_{10}, and p_{11} are the probabilities indicated in figure 1, P_D is the detection probability, and P_{FA} is the false alarm probability. The transition probabilities indicated in figure 2 are explained hereafter.

- If, at time instant 0, we are in Healthy state, the next time instant we can be in
  - Healthy state again, with probability $p_{00}(1-P_{FA})$, i.e. nothing happened and no false alarm occurred
  - in undetected Faulty state with probability $p_{01}(1-P_D)$
  - in LOC state, due to a detected fault, since we are not implementing exclusion, or due to a false alarm, with probability $p_{01}P_D + p_{00}P_{FA}$
- If, at time instant 0, we are in Faulty state, the next time instant we can be in
  - Faulty state again, with probability $p_{11}(1-P_D)$, i.e. the fault has not been repaired and again we have not detected the fault
  - Healthy state, with probability $p_{10}(1-P_{FA})$: the fault has been repaired and no false alarm occurred

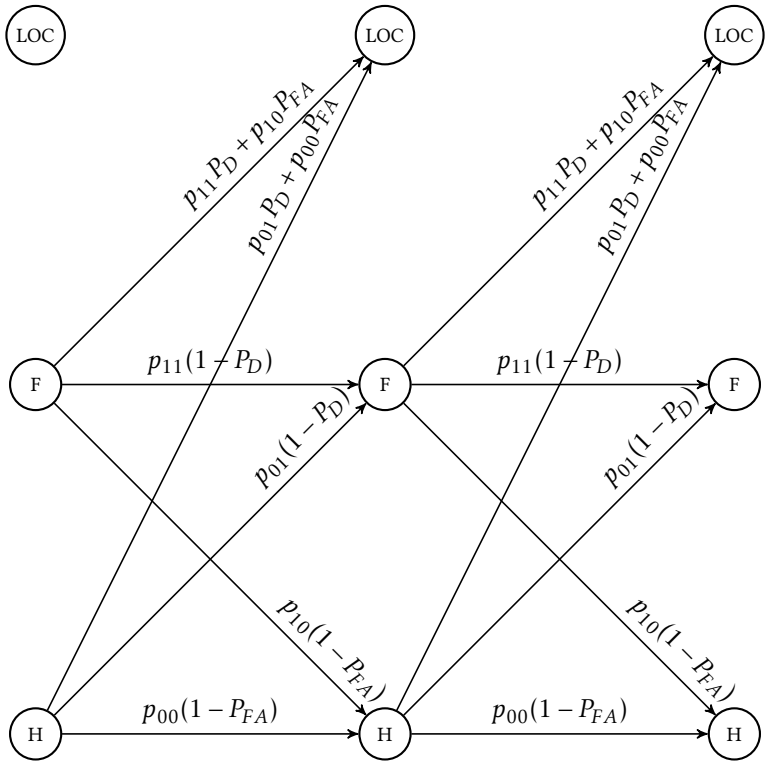


Figure 3: Trellis rappresentation of Single Satellite and Receiver model

  - LOC state with probability $p_{11}P_D + p_{10}P_{FA}$, if the fault persists and we detect it or the satellite returns Healthy but we have a false alarm event. Again the transition due to a detection happens because we are not implementing exclusion

- If, at time instant 0, we are in LOC state, the next time instant we can be in

  - only in LOC state again, since at the previous time step we were in an absorbing state LOC, and by definition the system cannot exit an absorbing state. Once Continuity is lost in our observation time window (15s in aviation, 3 hours in maritime application) it is lost for all the time window.

Figure 3 shows the trellis representation of the evolution of our Markov chain. The multi satellite system transition diagram is a simple generalization of the model for a single satellite.

USER RECEIVER AND MULTIPLE SATELLITES (NO EXCLUSION)

Instead of considering a single satellite system we consider now the more realistic case of multiple satellites. The assumption is that every satellite evolves independently from the other satellites, and from the receiver point of view it is sufficient to detect a fault for declaring the system unavailable (if no exclusion is implemented). The theoretical background that links the model for the single satellite evolution to the multiple satellite evolution is presented in the following.

Suppose that we have two independent separate Markov chains, with known properties, and we want to study the simultaneous evolution of the two systems. To describe the new aggregate system, we consider two independent Markov chains and define as $X(n)$ the random variable that describes the state of the first chain at a generic instant and $Y(n)$ as the variable describing the second chain state. The sets $\Omega_x = \{\omega_x^{(1)}, \omega_y^{(1)}, ..., \omega_N^{(1)}\}$, $\Omega_y = \{\omega_x^{(2)}, \omega_y^{(2)}, ..., \omega_N^{(2)}\}$ are the state spaces of X, Y and the matrices $\boldsymbol{T_X}, \boldsymbol{T_Y}$ are the transition matrices of the two systems respectively, where

$$\boldsymbol{T_X}(i \to j) = P(X(n) = \{\Omega_x\}_j | X(n-1) = \{\Omega_x\}_i) \tag{12}$$

$$\boldsymbol{T_Y}(i \to j) = P(Y(n) = \{\Omega_y\}_j | Y(n-1) = \{\Omega_y\}_i) \tag{13}$$

The purpose is to merge the two Markov chains and derive the properties of the new unique system. We start our analysis by observing that the couple $X(n), Y(n)$ (as well as $X(n-1), Y(n-1)$) belongs to a space that is the cartesian product of Ω_x, Ω_y.

$$(X(n), Y(n)) \in \Omega_x \times \Omega_y = \{(x, y) \quad | x \in \Omega_x, \quad y \in \Omega_y\}$$

Obviously, the cardinality of $\Omega_x \times \Omega_y$ is the product of the original cardinalities. If we focus on a particular case, in which the two independent systems are two independent copies of the same system (as in the case of considering the mutual evolution of two satellites) then the couple $(X(n), Y(n))$ belongs to

$$(X(n), Y(n)) \in \Omega \times \Omega = \{(x, y) \quad | x \in \Omega, \quad y \in \Omega\}$$

and the two transition matrices are equal, then

$$\boldsymbol{T_X} = \boldsymbol{T_Y} = \boldsymbol{T}$$

If we write the composition of $\Omega \times \Omega$ as

$$\Omega \times \Omega = \{(\omega_x, \omega_x), (\omega_x, \omega_y), ..., (\omega_x, \omega_N), (\omega_y, \omega_x), (\omega_y, \omega_y), ...(\omega_y, \omega_N), ...(\omega_N, \omega_N)\}$$

we can easily derive that the generic couple $(\omega_{i_A}, \omega_{i_B})$ maps to the $(i_A - 1)N + i_B$ element of $\Omega \times \Omega$. Notice that this correspondence is bijective (i.e. $\forall k \in \{1, ..., N\}^2 \quad \exists!(i_A, i_B) : (i_A - 1)N + i_B = k \quad i_A, i_B \in \{1, ..., N\}$). The first question we want to answer is: what is the probability of transition from one state to another of our new state space. Or, in equations,

$$P(Z(n) = \{\Omega \times \Omega\}_{j_C} | Z(n-1) = \{\Omega \times \Omega\}_{i_C}) \tag{14}$$

where $Z(n)$ is the new time dependent random variable that describes the state of the overall system. Existing a biunique relationship between i_C, j_C and $(i_A, i_B), (j_A, j_B)$ we can rewrite (14) as

$$\begin{aligned}
&P(Z(n) = \{\Omega \times \Omega\}_{j_C} | Z(n-1) = \{\Omega \times \Omega\}_{i_C}) = \\
&P(X(n) = \{\Omega_x\}_{j_A}, \quad Y(n) = \{\Omega_y\}_{j_B} | X(n-1) = \{\Omega_x\}_{i_A}, \quad Y(n-1) = \{\Omega_y\}_{i_B}) = \\
&P(X(n) = \{\Omega_x\}_{j_A} | X(n-1) = \{\Omega_x\}_{i_A}) P(Y(n) = \{\Omega_x\}_{j_B} | Y(n-1) = \{\Omega_x\}_{i_B})
\end{aligned}$$

where the last step of our chain of equalities is due to the independence of the two systems described by X, Y. It is also straightforward to demonstrate that the random variable $Z(n)$ has a Markovian property, i.e.

$$P(Z(n) = \{\Omega \times \Omega\}_{j_C} | Z(n-1) = \{\Omega \times \Omega\}_{i_C}, Z(n-2) = \{\Omega \times \Omega\}_{l_C}) = P(Z(n) = \{\Omega \times \Omega\}_{j_C} | Z(n-1) = \{\Omega \times \Omega\}_{i_C}) \tag{15}$$

If we define a new matrix $\boldsymbol{T}_Z$ as an $N \times N$ matrix whose generic entries (i_C, j_C) is equal to

$$\boldsymbol{T}_Z(i_C, j_C) = P(Z(n) = \{\Omega \times \Omega\}_{j_C} | Z(n-1) = \{\Omega \times \Omega\}_{i_C})$$

then we can derive that

$$\begin{aligned}
&\boldsymbol{T}_Z(i_C, j_C) = \boldsymbol{T}_Z((i_A - 1)N + i_B, (j_A - 1)N + j_B) = \\
&P(X(n) = \{\Omega_x\}_{j_A} | X(n-1) = \{\Omega_x\}_{i_A}) P(Y(n) = \{\Omega_x\}_{j_B} | Y(n-1) = \{\Omega_x\}_{i_B}) \\
&\boldsymbol{T}(i_A, j_A) \boldsymbol{T}(i_B, j_B)
\end{aligned}$$

where, focusing on

$$\boldsymbol{T}_Z((i_A - 1)N + i_B, (j_A - 1)N + j_B) = \boldsymbol{T}(i_A, j_A) \boldsymbol{T}(i_B, j_B)$$

we notice that the new transition matrix is exactly the Kroenecker Product of the first transition matrix with itself, that is

$$\boldsymbol{T}_Z = \boldsymbol{T} \otimes \boldsymbol{T} \tag{16}$$

In conclusion, the new system is a Markov chain whose state space is the cartesian product of the original state spaces and the transition matrix is the Kroenecker product of the starting matrices.

Once the extension from a single satellite to multiple satellites is clear, we can include in the system description also the receiver behaviour, similarly to the single satellite case, introducing the Loss of Continuity state (L_{oc}). In this case both the probability to detect an existing fault and the probability to have a false alarm depend on the number of satellites. For this

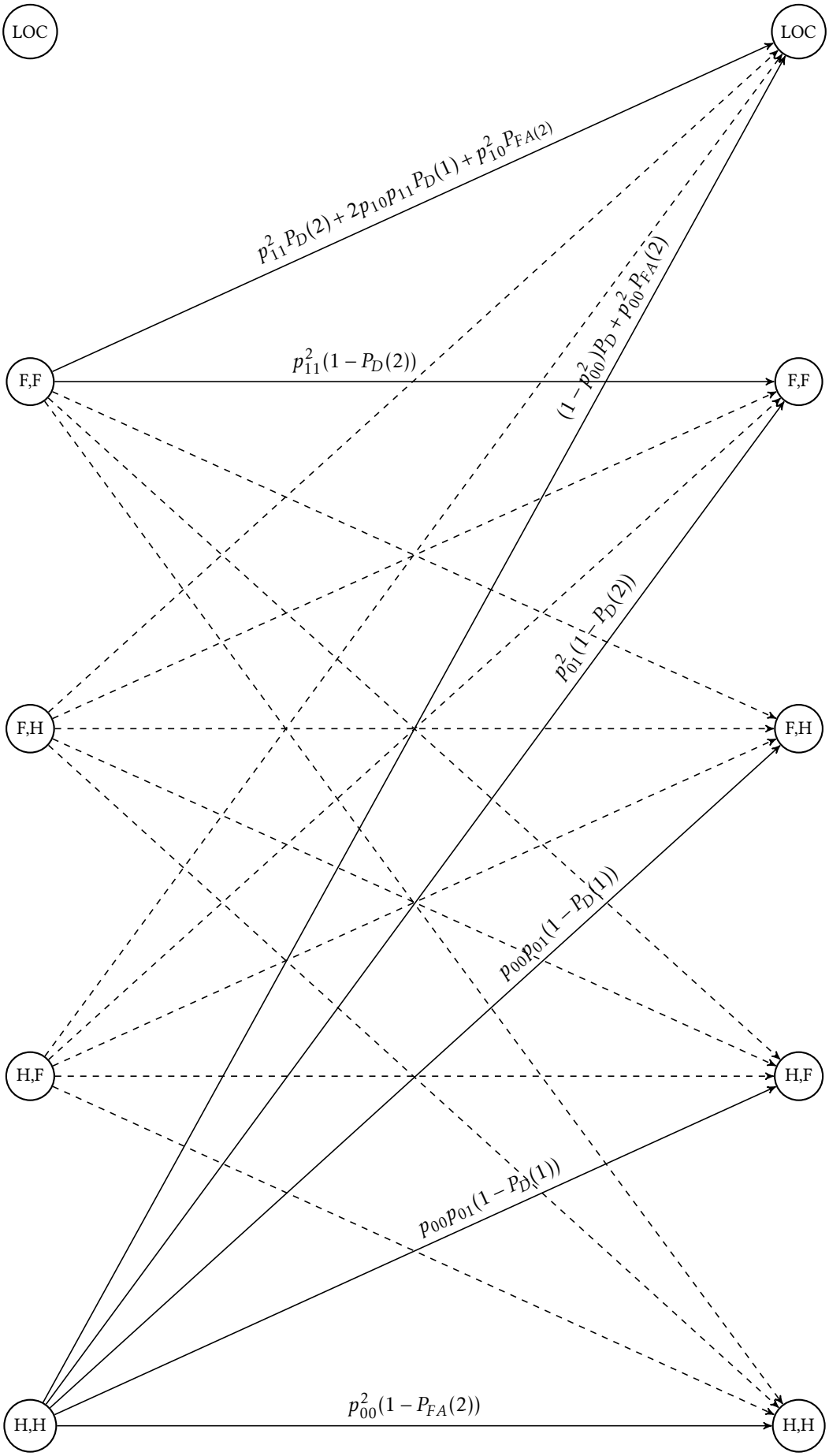


Figure 4: Multi Satellite and Receiver Model trellis

reason they are indicated respectively as $P_D(N_{\text{sat in fault}})$ and $P_{FA}(N_{sat})$. Notice that the detection probability is a function of the number of faulty satellites $N_{\text{sat in fault}}$, while the false alarm probablity depends only on the overall number of satellites N_{sat}. In the following we will consider $P_D(N_{\text{sat in fault}})$ a constant with respect to the number of faulty satellites. However, this hypothesis is not a necessary for the work to be consistent. The new overall trellis description of the system is depicted in Figure 4 for the case $N_{sat} = 2$.

In the general case, to derive the transition matrix, we order the $2^{N_{sat}} + 1$ possible states of the system such that the first state is the all satellites Healthy state, the states from the second to the second to last are the states were at least one satellite is Faulty (but non fault is detected), and the last state is the Loss of Continuity state (the set of states is: $\{(H,H,H,H,...H)$, $(F,H,H,H,H....H)$, $(H,F,H,H,H,..H),....(F,F,F,F,...H),(F,F,F,F,...F),L_{oc}\}$).
The transition matrix $\boldsymbol{M}$ of the system can be then easily shown to be

$$\boldsymbol{M} = \begin{bmatrix} \boldsymbol{T}_{\boldsymbol{N_{sat}}} & \boldsymbol{t}_{\boldsymbol{N_{sat}}} \\ \boldsymbol{0}^T & 1 \end{bmatrix} \tag{17}$$

where

$$\boldsymbol{T}_{\boldsymbol{N_{sat}}} = (\otimes^{N_{sat}} \boldsymbol{T})\boldsymbol{D}$$

and

$$\boldsymbol{D} = \text{diag}\{[(1-P_{FA}),(1-P_D),(1-P_D),(1-P_D),(1-P_D)....]\}$$

Notice that the fact that the matrix $\boldsymbol{D}$ is present is due to the fact that we can transit from the all satellites Healthy state to the L_{oc} state due to a False Alarm or from a state in which one or more satellites are Faulty and we detect the fault. This information is also included in the vector $\boldsymbol{t}_{N_{sat}}$ that describes the probability of transition from a generic non absorbing state to the absorbing (L_{oc}) state. Infact, being $\boldsymbol{M}$ a transition matrix, $\boldsymbol{t}_{N_{sat}}$ is easily derived as $\boldsymbol{t}_{N_{sat}} = \boldsymbol{1} - \boldsymbol{T}_{N_{sat}}\boldsymbol{1}$ where $\boldsymbol{1} = \begin{bmatrix} 1 & 1 & 1 & 1... & 1 \end{bmatrix}^T$.
The initial state probability is instead described by the vector $\boldsymbol{\tau}_{\boldsymbol{N_{sat}}}$:

$$\boldsymbol{\tau}_{\boldsymbol{N_{sat}}} = \otimes^{N_{sat}} \boldsymbol{\pi}$$

where $\boldsymbol{\pi}$ is the steady state distrubution of the state of the single satellite.

The cumulative distribution function that gives us the information about the probability that at a certain time instant k we are in the absorbing state, that in our case coincide with having lost Continuity is

$$F(k) = 1 - \boldsymbol{\tau}\boldsymbol{T}^k_{\boldsymbol{N_{sat}}}\boldsymbol{1} \tag{18}$$

Thanks to the introduced model and using equation (18) we can thus compute the continuity risk considering also the time evolution of the system.

USER RECEIVER AND MULTIPLE SATELLITES (SNAPSHOT EXCLUSION)

In case we want to consider also a snapshot exclusion mechanism we can simply modify the system described previously. When the receiver detects a fault it tries to perform exclusion by isolating the Faulty satellites. This operation can be successful, i.e. exclusion is correctly performed, or unsuccessful. An unsuccessful exclusion happens when there are Faulty satellites that triggered the Integrity Risk threshold (see the concept of Failed Exclusion [5]), but the receiver was not able to identify the Faulty satellites. In this case the system is declared unavailable and continuity is lost. The probability of failing an exclusion is defined as P_{NEX} and has to be specified considering IMO requirements. We can define a new probability $\overline{P} = P_D P_{NEX}$, that is the probability of detecting the presence of a fault and failing the exclusion. Notice that in a snapshot exclusion mechanism, when a satellite is excluded at a given epoch it is reinserted in the set of possible satellites used for positioning in the successive epochs, and a new test is performed for determining whether it has to be re-excluded or not.

Also in this case, we can derive the complete trellis, shown in Figure 5, for a user that is using N_{sat} satellites. The difference with respect to the previous case is the fact that implementing exclusion, we improve the performance of the system in terms of Continuity because we lose Continuity only when we detect a fault and we are not able to exclude the Faulty satellites. In numerical terms we have as weight of the edges of the transition to the L_{OC} state the probability $\overline{P} = P_D P_{NEX}$ instead of P_D, where it is easy to understand that since the failed exclusion probability must be a small number $\overline{P} << P_D$. Notice that in the proposed model we consider also the case in which exclusion is not possible. This happens for all the system states

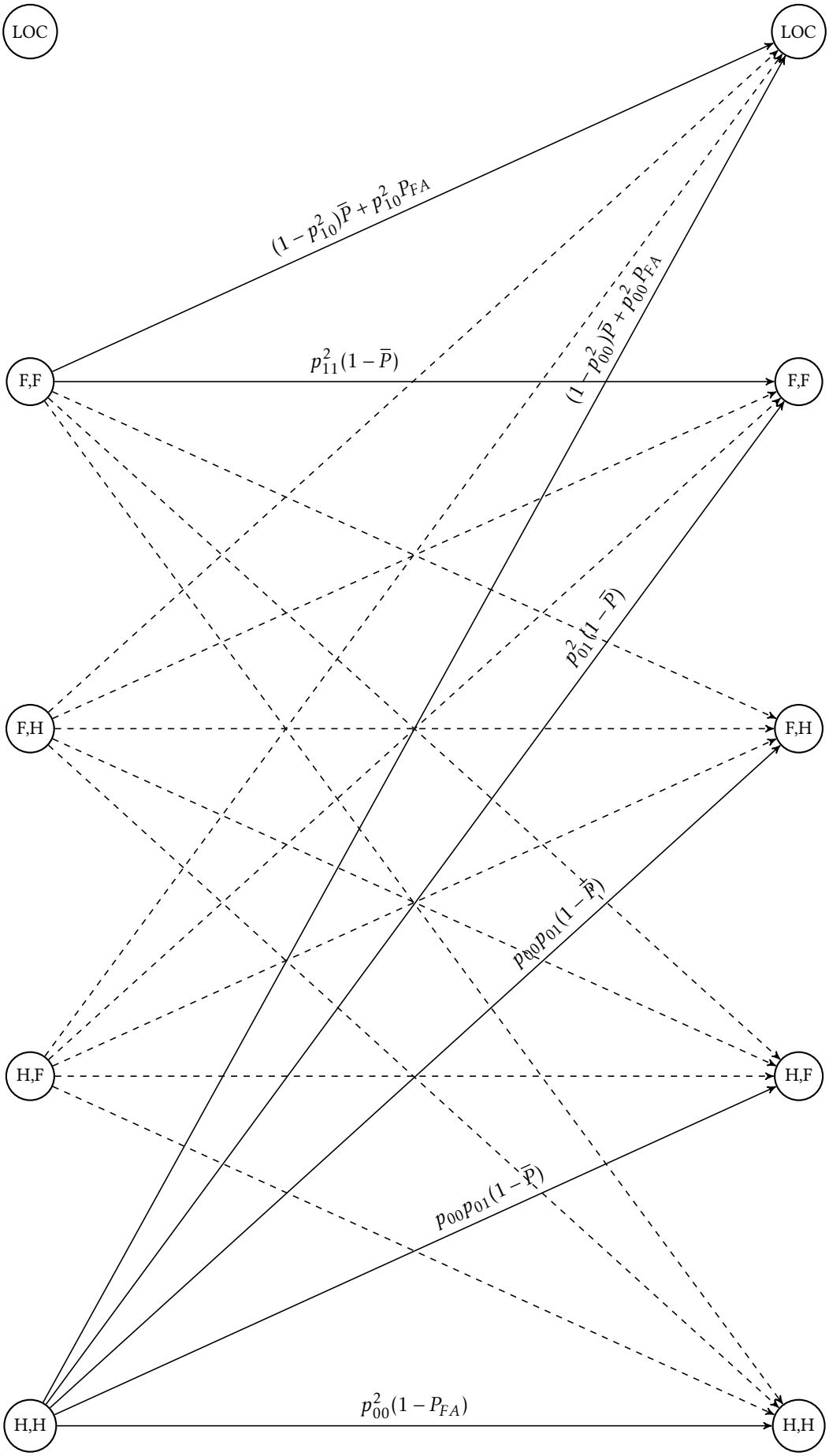


Figure 5: Multi Satellite and Receiver implementing exclusion Model trellis

in which the number of Healthy satellites is smaller or equal to $M + 3$ Concerning the algebraic details, also in this case the transition matrix of the system is in the form

$$\boldsymbol{M} = \begin{bmatrix} \boldsymbol{T}_{\boldsymbol{N_{sat}}} & \boldsymbol{t}_{\boldsymbol{N_{sat}}} \\ \boldsymbol{0}^T & 1 \end{bmatrix} \tag{19}$$

and the initial state distribution is $\boldsymbol{\tau}_{\boldsymbol{N_{sat}}}$. As in the previous case

$$\boldsymbol{T}_{\boldsymbol{N_{sat}}} = (\otimes^{N_{sat}} \boldsymbol{T})\boldsymbol{D}$$
$$me\boldsymbol{\tau}_{\boldsymbol{N_{sat}}} = \otimes^{N_{sat}} \boldsymbol{\tau}$$

However, this time the diagonal matrix $\boldsymbol{D}$ is different, due to the fact that when a fault is detected, when possible, exclusion is attempted. The first element of the matrix is the same as in the previous case and reflects the effect of False Alarms $\boldsymbol{D}|_{\{1,1\}} = 1 - P_{FA}$. The other elements of the diagonal will have the two different values depending on the number of Faulty satellites, i.e. the generic diagonal component of the matrix will be defined as

$$\boldsymbol{D}|_{\{1,1\}} = \begin{cases} 1 - \bar{p} & \text{if the number of healthy satellites >M+3} \\ 1 - P_D & \text{otherwise} \end{cases} \tag{20}$$

There will be respectively N_1 elements of the diagonal matrix with value $1 - \bar{p}$ and N_2 with value $1 - P_D$, where

$$N_1 = \sum_{i=1}^{N_{sat}-4-M} \binom{N_{sat}}{i}$$

and, obviously

$$N_2 = (2^{N_{sat}} - 1) - N_1$$

Again, the probability of being in the absorbing state (L_{oc}) at the generic time instant k is equal to

$$F(k) = 1 - \boldsymbol{\tau}\boldsymbol{T}_{\boldsymbol{N_{sat}}}^{k}\boldsymbol{1} \tag{21}$$

USER RECEIVER AND MULTIPLE SATELLITES (SEQUENTIAL EXCLUSION)

In this section we introduce the Markov model for a receiver implementing sequential exclusion. The proposed model is depicted in Figure 6, where a single satellite system is rappresented. The generalization to the multi satellite case will be easily done in the same way as for the previous cases.

Considering a sequential exclusion mechanism we introduced a new state, the excluded satellite state. When a Faulty satellite is detected and correctly excluded it will remain excluded as long as the successive epoch detection and exclusion mechanism will work correctly. We defined $P_{D,2}$ as the secondary detection probability of a fault and $P_{ex,2}$ as the secondary correct exclusion probability in a successive epoch to the first in which the fault is correctly detected and excluded. The obvious definition of secondary failed exclusion probability follows $P_{nex,2} = 1 - P_{ex,2}$. If a fault is continuously going on for a given satellite, it will be easier to keep track of that fault (having a continuous detection) and correctly excluding it. In the proposed system model to rappresent the fact that detection and exclusion are easier after the first epoch in which a fault is correctly excluded, the new detection probability and exclusion probability are greater than the first time instant probabilities, i.e.

$$P_{D,2} > P_D, \quad P_{ex,2} > P_{ex}$$

In the result section $P_{D,2}$ as well as P_D will be fixed while the design will take place for the two correct exclusion probabilities ($P_{ex}, P_{ex,2}$) and the false alarm probability (P_{FA}). Concerning the correct exclusion mechanism the following assumption has been made: when more than one satellite is Faulty exclusion is either correctly performed on all the Faulty satellites or failed, i.e. it is not possible to correctly exclude only one of the Faulty satellites and fail the exclusion of the other satellite(s).

HYPOTHESES ON THE INITIAL STATE DETECTION MECHANISM: BLIND WAKE UP (H_0) OR PAST INFORMATION AVAILABLE (H_1)

In this brief section we will understand the difference between two different assumptions on the detection mechanism when

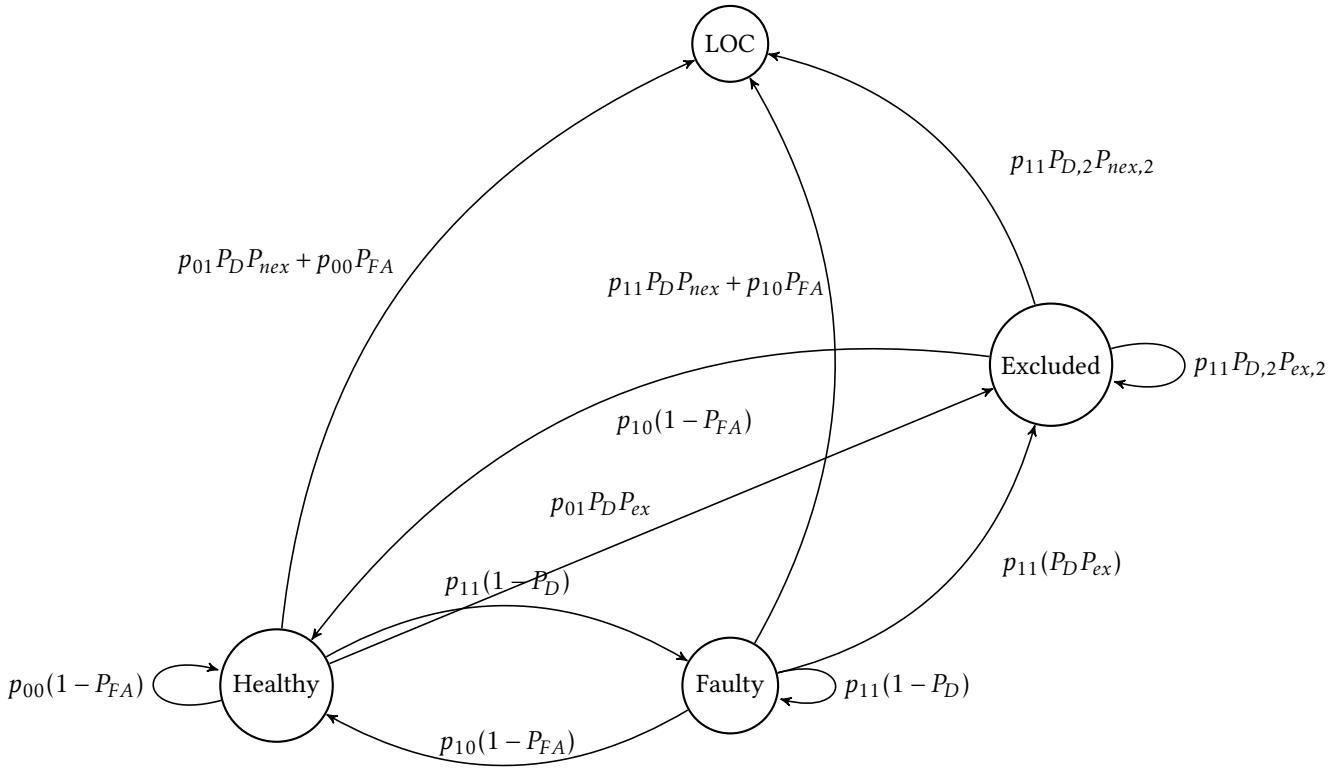


Figure 6: Single Satellite and Sequential exclusion Receiver model

the receiver starts an operation: the past information available wake up or the blind wake up. The past information available wake up is the hypothesis H_1, done in [6], that when the receiver starts an operation has the past information about the Faulty satellites. If this is the case when an operation is started all the considered satellites in view are Healthy (the Faulty ones have been excluded previously). The blind wake up hypothesis H_0 is instead the one assumed in the derivation of our model: when the receiver starts an operation, some of the satellites can be in Faulty state according to the initial distribution probability vector $\boldsymbol{\tau}_{\boldsymbol{N_{sat}}}$ since no information about the Faulty satellites in the previous epochs is available.

However, our model is flexible with respect to this hyphotesis: infact, we can analyze the continuity risk time evolution under the past information available wake up hypothesis simply by imposing that the initial state of the system is constrained to be in the all satellites Healthy state, or, in equations, that the vector $\boldsymbol{\tau}_{\boldsymbol{N_{sat}}}$ is equal to

$$\boldsymbol{\tau}_{\boldsymbol{N_{sat}}} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \end{bmatrix}^T$$

Independently on the fact whether information about the past is a reasonable hypothesis and when it is applicable, the formalism introduced for the initial state distribution provides flexibility in the analysis of different systems.

PERFORMANCE ANALYSIS

In the previous sections we developed Markovian models for the three cases of a receiver not implementing exclusion, implementing snapshot exclusion and implementing sequential exclusion. In this section we will see that exclusion is always required in maritime environment. We must warn the reader that in our analysis the only fixed parameter is the Detection Probability (P_D) that is driven by integrity requirement and is fixed to the value $1-10^{-2}$. This value is derived by [5] where it is stated that: *" The EMT test prevents faults that are not large enough to ensure detection from creating vertical position errors greater than 15 m more often than 0.00001% of the time"* The probability that a fault that creates a great vertical error is undetected is smaller than 10^{-7} . This event is the joint occurrence of a fault (with probability P_{fault}) and a miss detection (with probability P_{MD}). If we consider the apriori probability of having a fault equal to $P_{fault} = \pi_0 = 10^{-5}$, then the probability of miss detection

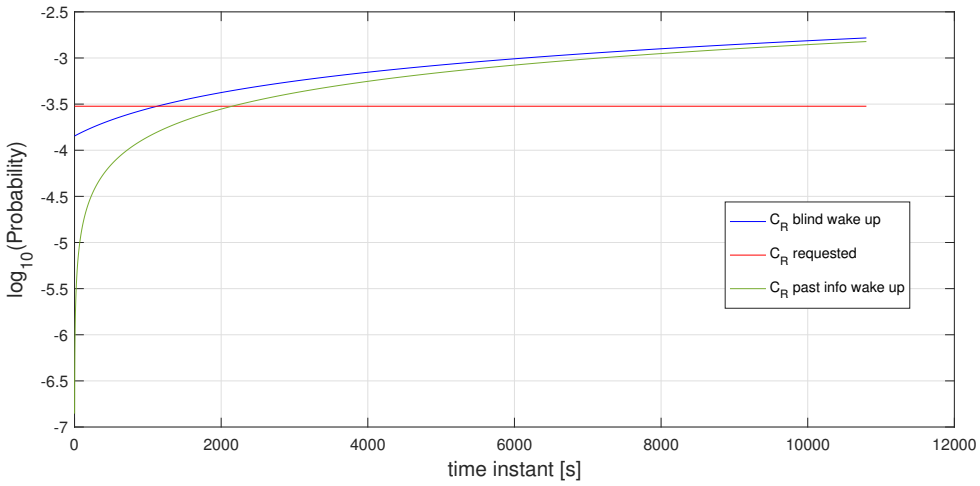


Figure 7: Time evolution of continuity risk

is equal to 10^{-2}. All the analyses for the three different cases (no exclusion, snapshot exclusion and sequential exclusion) will be carried out considering both the hypothesis of having satellites past information available or the blind wake up hypothesis.

DETECTION (NO EXCLUSION) SYSTEM ANALYSIS AND DESIGN

When no exclusion mechanism is implemented, the only degree of freedom the receiver designer has is the determination of the False Alarm probability, since as already explained the detection probability is fixed and driven by integrity requirements. We will see for both hypotheses that sometimes, even with a zero false alarm probability, the continuity risk is not met. Remember that the continuity risk is the maximum allowed probability of losing Continuity during the 3 hours of operation ([2]) and is equal to 3×10^{-4}.

As already explained before, the continuity risk is computed thanks to (18), where the transition matrix is a function of the number of satellites and the detection and false alarm probabilities. Figure 7 shows the time evolution of the continuity risk for the case of $N_{sat} = 10$ and $P_{FA} = 10^{-7}$. As we can see in both cases at the end of the 3 hours the continuity risk requirement is not met since the actual continuity risk is higher, implying that the False alarm probability is too high (actually we will see that for a constellation of 10 satellites even a zero false alarm probability is not sufficient for meeting the IMO requirements). Moreover we can see that the difference between the continuity risks for the two hypothesis gets smaller as time passes: the advantage we have at the start of the operations when we are using information about past is that the number of average Faulty satellites is zero, but as time passes the average number of Faulty satellites starts to reach again is steady state value, as we can see from figure 8. Somehow counterintuitively this is another difference we can underline with respect to the avionic case due to the large time scales in maritime operation. In fact, while the difference in terms of performance in avionic field under the two different hypotheses is great, in the maritime field it is small.

We analyzed the required False Alarm probabilities as a function of the number of in view satellites (from 7 to 14) and for both the hypotheses. The result is that only in the case of 7 satellites in view, under the hypothesis of past information, and with a False Alarm probability smaller than $10^{-10.45}$ (that is an unrealistically low number considering that $P_D = 1 - 10^{-7}$) it is possible to fulfill the IMO requirements. In all the other cases, even a zero false alarm probability is not sufficient for having a continuity risk smaller than 3×10^{-4}. It is then deduced that the only alternative is an exclusion mechanism implementation.

DETECTION AND SNAPSHOT EXCLUSION SYSTEM ANALYSIS AND DESIGN

As we have seen, a detection mechanism without exclusion is not sufficient for fulfilling the IMO requirements. Introducing an exclusion mechanism, with an associated failed exclusion probability, we can analyze for different numbers of in view satellites and for the two hypotheses the required couple of False Alarm probability and Failed Exclusion probability. Figure 9 reports the unallowed operating point for the receiver for different number of satellites in view (from 7 to 14) . The blue

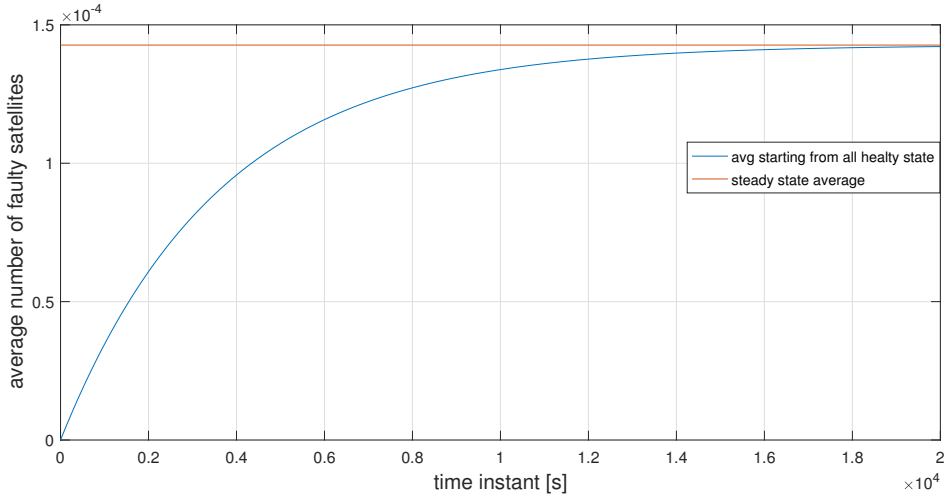


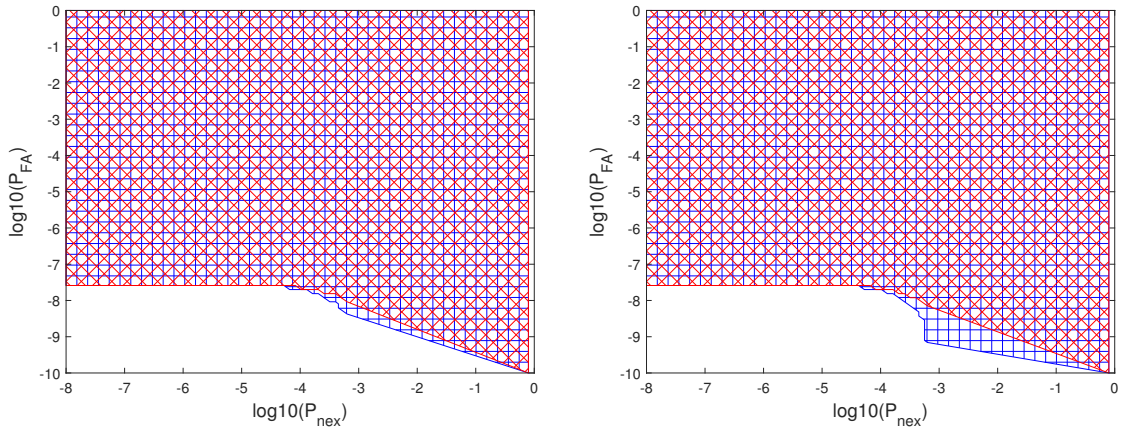
Figure 8: Average number of Faulty satellites

grid represents the unallowed zone when considering hypothesis H_0, the red grid instead the unallowed operating zone when considering hypothesis H_1. The first straightforward comment we can make is that as the number of satellites increases we need to have a smaller false alarm probability as well as a smaller failed exclusion probability. This is due to the fact that a higher number of satellites implies a higher probability that at least one satellite is Faulty and thus a higher probability of failing an exclusion. Moreover, as expected, with the past information available hypothesis the requirements on the false alarm probability and failed exclusion probability are looser, even if, as previously discussed, a long time scale decreases the gap between the probabilities of losing Continuity under the two hypotheses. The results presented in figures 9 show that, besides having a very small false alarm probability, also a very efficient exclusion mechanism is required. Repeating the simulations also for different values of detection probability P_D, it is observed that when exclusion is performed the dependency of the continuity risk on the detection probability is loose. When considering a large number of satellites and very low fault rates such as in this paper, the continuity risk is principally composed by the false alarm events and by the failed exclusion events. The failed exclusion event, qualitatively speaking, depends on the quantity $\bar{p} = P_D P_{nex} = P_{nex} - P_{MD} P_{nex}$ and when the miss detection probability is small, we can easily approxiamate $\bar{p}$ as $\bar{p} \simeq P_{nex}$, showing that the dependency of the continuity risk component due to the failed exclusion il slowly varying with respect to the detection probability. Notice that by increasing the fault rates the probability of having a critical number of satellites is not negligile and the dependency of the continuity risk on the detection probability is noticeable.

DETECTION AND SEQUENTIAL EXCLUSION SYSTEM ANALYSIS AND DESIGN

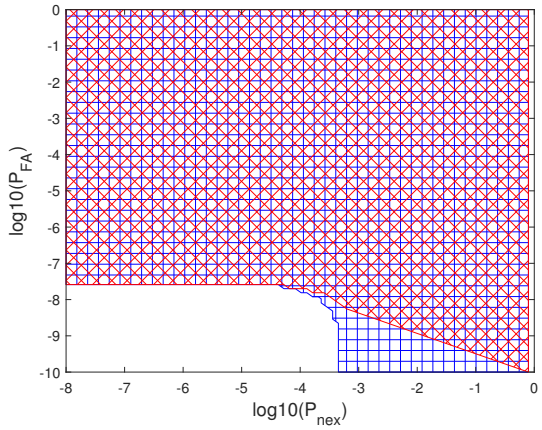
The required performance of a receiver if a sequential exclusion mechanism is adopted are hereafter analyzed. Obviously the continuity risk is decreased when a sequential exclusion mechanism is adopted. In designing a receiver implementing sequential exclusion we have two extra degrees of freedom that are the secondary detection probability and the secondary failed exclusion probability $P_{D,2}, P_{nex,2}$. We decided to fix the new detection probability $P_{D,2}$, that is a less critical parameter, and studied the performance in terms of Continuity Risk versus the triplet $P_{FA}, P_{nex}, P_{nex,2}$.

The main result is that introducing a sequential exclusion mechanism we can afford to reduce the required false alarm and failed exclusion probabilities at the cost of an increased complexity. In figure 10 are reported the required False alarm probabilities and primary Failed exclusion probabilities P_{nex} for different values of the secondary failed exclusion probability $P_{nex,2}$ for hypothesis H_0 and H_1 respectively. We decided to show the required performance for the cases when the secondary exclusion probability is equal to the primary exclusion probability times a constant $\gamma < 1$, i.e. $P_{nex,2} = \gamma P_{nex}$. In particular the following cases have been analyzed $P_{nex,2} = [1 \quad 0.5 \quad 0.01] \times P_{nex}$, corresponding respectively to the purple, blue and orange grids.

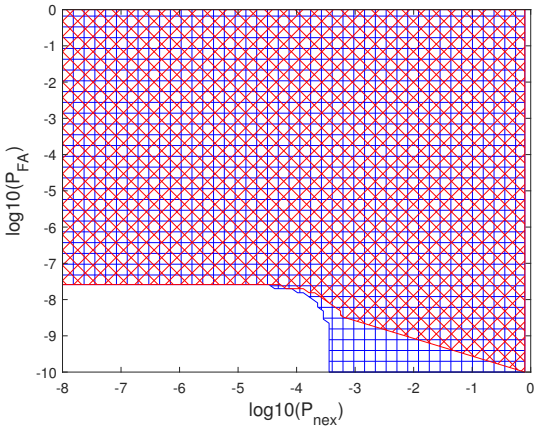


(a) $N_{sat} = 7$

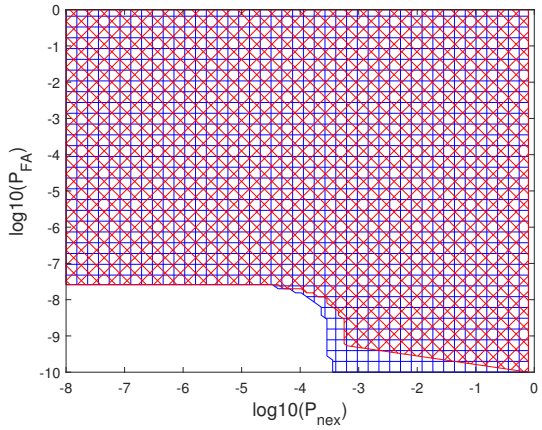
(b) $N_{sat} = 8$



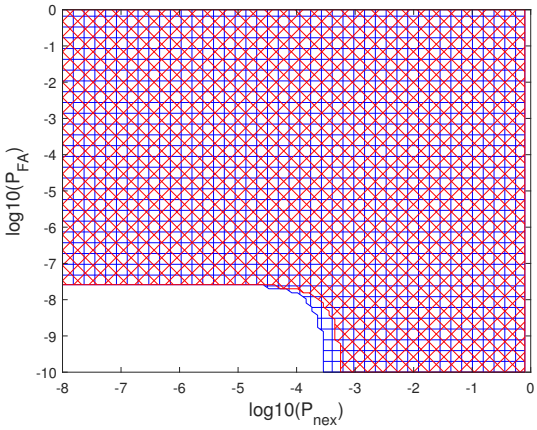
(c) $N_{sat} = 9$



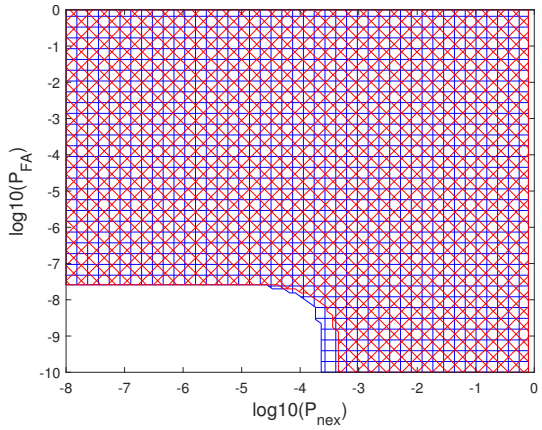
(d) $N_{sat} = 10$



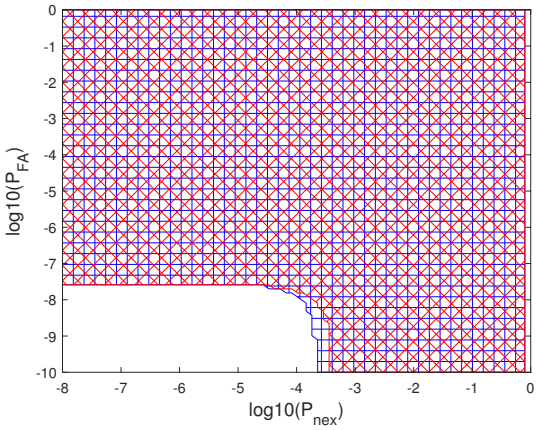
(e) $N_{sat} = 11$



(f) $N_{sat} = 12$

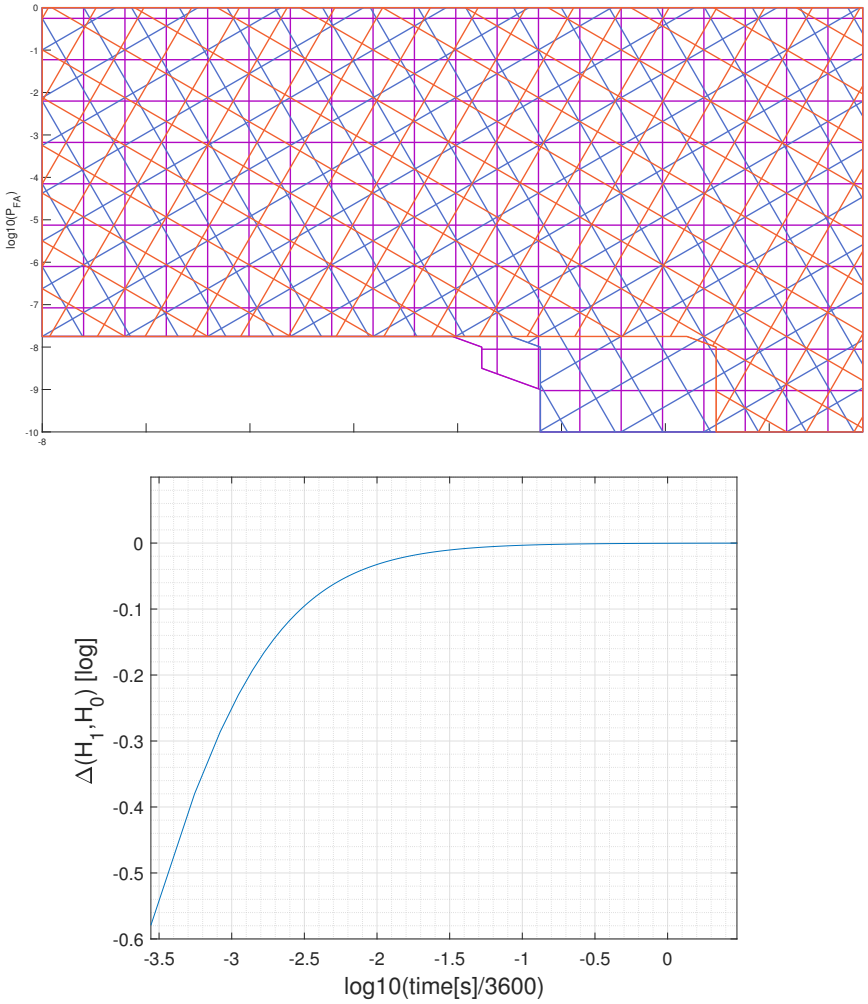


(g) $N_{sat} = 13$



(h) $N_{sat} = 14$

Figure 9: Required False Alarm and Failed Exclusion probabilities



DIFFERENT HYPOTHESES H_0 AND H_1 WHEN DEALING WITH SEQUENTIAL EXCLUSION

In this section we will briefly show that when dealing with sequential exclusion on the long time scale, the Markov models compute the same probability of losing Continuity. This is due to the fact that conceptually, if at the first time instant we are assuming H_0 that we have no information about the past, but we implement a good detection and sequential exclusion mechanism, in the second epoch, we have with an high probability that the Faulty satellites have been excluded in the previous epoch.This is equal to being from the second time instant, with high probability, in the same conditions assumed with hypothesis H_1. The difference between the two cases tend to naturally converges vanishes as time evolves, as we can see from Figure where the quantity

$$\Delta(H_1, H_0) = log_{10}(C_{R_{H_1}}) - log_{10}(C_{R_{H_0}})$$

is depicted.

PROOF OF CONCEPT AND CONSISTENCY WITH RESPECT TO TRADITIONAL APPROACHES

The purpose of this section is to show the differences and the similarities between the approach classically used in the avionic

environment and the approach derived in this paper.

The first part of this analysis will be focused on short time scales (up to 15s) as in avionic Continuity and Integrity requirements. Interestingly, we will show that for short time scales the derived continuity risks using the two models are numerically the same, and give a mathematical explanation of the phenomenon. On the other hand, enlarging the time scale, the two computed continuity risks start to diverge, and we will conclude that the estimation of the continuity risk for long time scales is not correct when using the same approach used in avionic applications.

We will refer to the classical approach as the Bernoulli model, where the probability of losing Continuity is computed by firstly computing the probability of losing Continuity on a single time instant (we will call this probability p) and then the overall continuity risk is computed using the well known formula for N_{inst} repeated Bernoulli trials:

$$P_{LOC,Bernoulli} = 1 - (1-p)^{N_{inst}} \tag{22}$$

As a last comment we should warn the reader that the calculation had been carried out only for the hypothesis of blind wake up of the receiver. However, it is straightforward to demonstrate that the same results hold also for the hypothesis of past information available.

SHORT TIME SCALE ANALYSIS

In this first part we will consider the case of a short time scale (15 s) and a receiver implementing a good exclusion mechanism ($P_{nex} \to 0$). Considering the classical approach, the probability p of losing Continuity in a single time instant is equal to the probability that all the satellites are Healthy and a false alarm event is present or at least one satellite is Faulty, we detect the fault, but we are not able to exclude it. This probability is equal to

$$p = \pi_0^{N_{sat}} P_{FA} + \sum_{i=1}^{N_{sat}} \pi_1^i \pi_0^{N_{sat}-i} P_D P_{nex} \simeq \pi_0^{N_{sat}} P_{FA} + N_{sat} \pi_1^1 \pi_0^{N_{sat}-1} P_D P_{nex} \tag{23}$$

where π_0 and π_1 are defined respectively in (10) and (11). The last approximation in (23) is due to the fact that for the computed probability on a single time instant we can neglect the case of multiple satellite faults. Notice that when running the numerical simulations the exact probability p has been used for the computation of the continuity risk, and that this approximation is done here just for simplicity in showing the calculations. By substituting (23) in (22) we obtain

$$P_{LOC,Bernoulli} = 1 - (1-p)^{N_{inst}} \simeq N_{inst} p = N_{inst} (\pi_0^{N_{sat}} P_{FA} + N_{sat} \pi_1 \pi_0^{N_{sat}-1} P_D P_{nex}) \tag{24}$$

where the approximation is due to the fact that $pN << 1$. Basically for short time scales using the Bernoulli formula or a simple scaling factor is numerically equivalent.

Instead, if we consider the Markov model the derivation is in the following. First, we notice that since the failed exclusion probability is a very small number ($P_{nex} \to 0$), then also the probability of detecting an error and failing an exclusion approaches zero($\bar{p} \to 0$). For small values of N_{inst} it is then possible to study the continuity risk, defined in (18) using perturbation techniques for eigenvectors:

$$\boldsymbol{\tau}_{\boldsymbol{N_{sat}}} ((\otimes^{N_{sat}} \boldsymbol{T}) \boldsymbol{D})^{N_{inst}} \simeq \boldsymbol{\tau}_{\boldsymbol{N_{sat}}} (\boldsymbol{D})^{N_{inst}} \tag{25}$$

This approximation holds because we can write

$$\begin{cases} \boldsymbol{\tau}_{\boldsymbol{N_{sat}}} (\otimes^{N_{sat}} \boldsymbol{T}) = \boldsymbol{\tau}_{\boldsymbol{N_{sat}}} \\ \boldsymbol{D} \simeq \boldsymbol{I} \end{cases} \tag{26}$$

and

$$\boldsymbol{\tau}_{\boldsymbol{N_{sat}}} \boldsymbol{D} (\otimes^{N_{sat}} \boldsymbol{T}) \simeq \boldsymbol{\tau}_{\boldsymbol{N_{sat}}} \boldsymbol{D}$$

Moreover, as long as $\boldsymbol{D}^{N_{inst}}$ is numerically close to an identity matrix $\boldsymbol{I}$, that is the case for $N_{inst} = 15$, the reasoning can be iterated and equation (25) can be proved to be correct. Remember that the first element of the diagonal matrix $\boldsymbol{D}$ has the value $1 - P_{FA}$ and that the other elements of the diagonal, that are defined by (20), contains either the value $1 - \bar{p}$ or $1 - P_D$ depending on the number of Healthy satellites. The result of a diagonal matrix elevated to the power of N_{inst} is equal to a diagonal matrix whose diagonal entries are the $N_{inst_{th}}$ power of the original diagonal entries. Considering moreover that:

- the power of the first element of the diagonal matrix can be approximated as $(1-P_{FA})^{N}_{inst} \simeq 1 - N_{inst}P_{FA}$
- the power of the elements of the diagonal matrix corresponding to the Faulty cases where exclusion is possible $(1-\bar{p})^{N}_{inst} \simeq 1 - N_{inst}\bar{p}$
- and that the quantity $(1-P_D)^{N_{inst}}$ is completely negligible and can be approximated with the 0 value

we can finally derive that a good approximation for the diagonal matrix $\boldsymbol{D}^{N_{inst}}$ is

$$\boldsymbol{D}^{N_{inst}} = \text{diag}(1-N_{inst}P_{FA}, \quad 1-N_{inst}\bar{p}, \quad 1-N_{inst}\bar{p}, \quad 1-N_{inst}\bar{p}, \quad ...0 \quad 0) = \boldsymbol{I} - N_{inst}\text{diag}(P_{FA}, \quad \bar{p}, \quad \bar{p}, ...0) \tag{27}$$

Finally, thanks to (25) and (27) it is possible to compute the probability of LOC ((18)) at the end of the $N_{inst} = 15$ seconds:

$$\begin{aligned}
&F(N_{inst}) = 1 - \boldsymbol{\tau}_{\boldsymbol{N_{sat}}}\boldsymbol{T}^{N_{inst}}_{\boldsymbol{N_{sat}}}\mathbf{1} = \\
&1 - \boldsymbol{\tau}_{\boldsymbol{N_{sat}}}(\boldsymbol{I} - N_{inst}\text{diag}(P_{FA}, \quad \bar{p}, \quad \bar{p}, ...0))\mathbf{1} = \\
&1 - \boldsymbol{\tau}_{\boldsymbol{N_{sat}}}\mathbf{1} + \boldsymbol{\tau}_{\boldsymbol{N_{sat}}}N_{inst}\text{diag}(P_{FA}, \quad \bar{p}, \quad \bar{p}, ...0)\mathbf{1} = \\
&\boldsymbol{\tau}_{\boldsymbol{N_{sat}}}N_{inst}\text{diag}(P_{FA}, \quad \bar{p}, \quad \bar{p}, ...0)\mathbf{1} \simeq \\
&N_{inst}(\pi_0^{N_{sat}}P_{FA} + N_{sat}\pi_1\pi_0^{N_{sat}-1}P_D P_{nex}) = P_{LOC,Bernoulli}
\end{aligned}$$

For short time scales (avionic environment) and implementing a good exclusion mechanism, the results of the two models are the same.

LONG TIME SCALE ANALYSIS

Since we have just shown that the probability derived with a Markov model and the Bernoulli probability are numerically equivalent, a natural question is the following: why do we have to built a such complicated model? The answer is the following: if it is true that for a short time period and with good exclusion mechanism the results coincide, this is not true anymore for a period of 10800s (3hrs). Moreover, these models gives us the ability to characterize different types of multiple faults (multiple faults with number of Healthy satellites greater or smaller than 4).

Figure 11 shows the time evolution of the continuity risk, fixed all the system probabilities, for the two different hypotheses (blind wake up and past information available) for the Bernoulli method of computation and for the Markov method. We can see that, if for short duration the 3 graphs basically concide (as mathematically proved in the previous Subsection) , they converge to different values at the end of the 3 hours.

Under the hypothesis H_0 of blind wake up, the Bernoulli model is over conservative. To understand why, it is fundamental to have clear in mind the difference between the true Continuity requirement and the average sense Continuity requirement. Consider the following thought experiment: suppose that you have to move from point A to point B , taking 100 steps, and that you have a hole in your pocket.

At every step you take, you lose your wallet with a probability p. To be fully precise, at every step you take, you lose your wallet with probability p if your wallet is still there and with probability 0 if you already lost your pocket. This is why, if you compute the probability that once you reached point B using a Bernoulli distribution you are overestimatng the probability of losing your wallet. Infact, with a Bernoulli like approach you include in the computed probability also multiple falls of the same wallet, while this is not possible. When the probability of losing your wallet is very very small and you have to take few steps, then the Bernoulli computation is an accurate approximation because numerically, the multiple falls event has a negligible probability (maybe serveral order of magnitudes smaller than the total probability) and this is why on short time scales (15 s of avionics) it is reasonable to use an average sense approach while in maritime applications it is not.

On the other hand, when hypothesis H_1 is selected, the Bernoulli model is underconservative. This is due to the fact that under the hypothesis H_1, we are assuming for all the 3 hours of duration of the operation that the perfect knowledge of the Faulty satellites in the previous time instant is available and that the only satellites that can be Faulty are the satellites that transit to a Faulty state between the previous time instant and the considered measurement time. It is obviously an unreasonable assumption for a 3 hours long operation. In terms of equations the probability p used for the computation of the continuity risk is the following:

$$p_{H_1} = p_{00}^{N_{sat}}(1-P_{FA}) + N_{sat}p_{01}\bar{p} \tag{28}$$

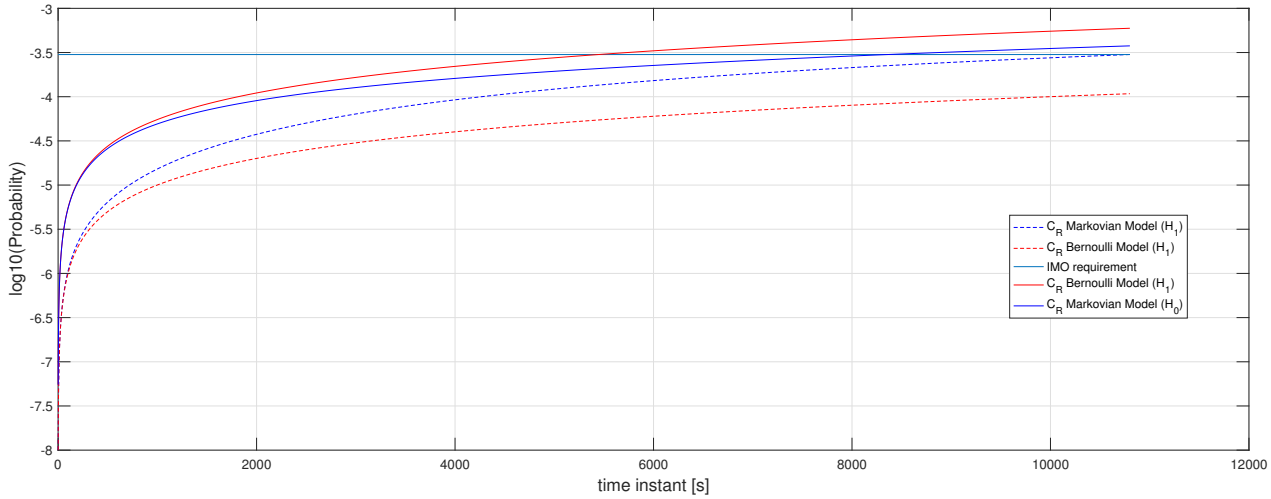


Figure 11: Markov, Bernoulli and scaling models to comparison,$N_{sat=10}$,$P_D = 1 - 10^{-5}$,$P_{nex} = 10^{-3.5}$,$P_{fa} = 10^{-8}$

since we are assuming that all the Faulty satellites have been excluded in the previous time instants. Basically, under the hypothesis H_1 and considering the Bernoulli trials the leading term is the term related to the false alarm probability and assuming

$$p_{H_1} = p_{00}^{N_{sat}}(1 - P_{FA}) \tag{29}$$

numerically is the same for the computation of the continuity risk.

Finally we can conclude that the Bernoulli approach used classicaly in aviation is very accurate for short time scales (we considered 15 second) and coincide with the Markov approach, while for long time scales it is necessary to adopt a more complicated model such as the one proposed in this paper.

CONCLUSIONS

In this paper we proposed a new markovian model for the computation of the continuity risk in maritime environment. The derived model is flexible with respect to different important parameters such as the number of in view satellites, the average fault and repair rates, the detection and false alarm probabilities and the receiver implementation structure. With the derived tool we showed the necessity for exclusion to fullfill the maritime continuity risk requirements. Moreover when implementing sequential exclusion, at the cost of an increased complexity, the continuity risk at the end of the 3 hours is reduced. Furthermore the information that the receiver implementing sequential exclusion has about the faulty satellites when starting an operation has a time decreasing impact on the continuity risk with respect to the case when the receiver has no available information at the start of the operation. We also showed that when the time scale is reduced to 15s, such as in aviation operation, the proposed model and the preceeding models [6] lead to the same numerical results. The proposed model could be considered as a valid tool for the computation of the continuity risk for a wide range of applications with different requirements and time scales.

References

[1] Minimum Operational Performance Standards for GPS WAAS. Technical report.

[2] International Maritime Organization (IMO). *Res. A.953(23) on World Wide Radionavigation System (WWRNS)*, 2002.

[3] M. Baldauf J. O. Klepsvik, P. B. Ober. A critical look at the IMO requirements for GNSS. September 2007.

[4] V. Krivtsov M. Modarres, M. Kaminskiy. *Reliability Engineering and Risk Analysis: A Practical Guide.* CRC Press, 2010.

[5] GPS-Galileo Working Group C ARAIM Technical Subgroup. Technical report.

[6] Joerger Mathieu Pervan Boris Zhai, Yawei. Continuity and Availability in Dual-Frequency Multi-Constellation ARAIM. pages pp. 664–674, Tampa, Florida„ September 2015.