Intrinsic Fault Resistance for Nonlinear Filters with State-Dependent Probability of Detection

Gunner S. Fritsch and Kyle J. DeMars
Texas A&M University, College Station, TX 77843, USA

ABSTRACT

The probability of detection is an element that can be thought of as the probability that a given measurement scan contains a valid, target-oriented return. In most systems, the detection probability is inherently a function of the state, which can make forming closed-form filtering solutions exceedingly difficult. Oftentimes, closed-form filters will either neglect the probability of detection outright, or in some cases may approximate it as state-independent. Both assumptions simplify calculations, yet bar the filter from extracting any state-information from the probability of detection. This work seeks to reevaluate current estimation practices by testing and comparing several probability of detection models of varying fidelity. This is done by proposing a filter update with intrinsic fault resistance capable of processing multiple sensor returns contained within a single measurement scan. Three different methods of detection probability modeling are described, which are subsequently used to form three distinct Gaussian mixture filters. To test the filters, Monte Carlo results are taken from two different simulations: the first a simple falling body tracking scenario, and the second a more complex orbit determination scenario. The findings of both simulations are discussed, and the behaviors of the various methods are commented on.

1. INTRODUCTION

From spaceflight navigation to space situational awareness (SSA), the ability to accurately estimate the characteristics of a physical object of interest is a powerful tool and often will be a significant factor in the success (or, at times, the failure) of a mission [1, 2]. Since the advent of modern estimation in the mid-to-late 1950s, the important, widely-applicable nature of the field has spawned many advancements and improvements across many different disciplines. SSA, in particular, has seen much growth in the way of various multi-target tracking (MTT) methodologies like the joint probabilistic data association (JPDA) and multiple hypothesis tracking (MHT) filters [3], as well as nonlinear-Bayesian approaches to orbit determination [4], to mention a few. Of course, the number of developments in estimation to date is numerous; there exists far more than can be mentioned by this work alone. However, one fundamental consideration precedes the conception or augmentation of any estimation framework, which is deciding what physical aspects *should* be modeled by the estimator machinery. It is importunate to differentiate between this decision and determining which physical attributes *can* be modeled, as the ability to model a certain element does not account for the practicality of its inclusion within the model, which is a product of limitations and objectives such as computational capacity and desired model accuracy. For example, it is customary in estimation practices to model sensor noises and biases, as well as the temporal dynamics of the system, but many factors may be left unmodeled to reduce computations and ease implementation [5]. Thus, deciding which aspects should be modeled is most often a complicated process that is heavily dependent upon the resources and goals at the time. As resources and goals are constantly changing, especially when regarding the progressional leaps in computational technologies [6, 7], this is a question that should be constantly revisited. With this motivation in mind, the work herein reexamines the way by which the probability of detection is modeled and explores the potential benefits that arise from enhancing said model.

The probability of detection is, effectively, a term that quantifies the likeliness that a valid (i.e. target-oriented) sensor return will be generated by a given measurement model [8]. It naturally arises in the update of Bayesian filters which model both false alarms and valid measurements separately, such as the probability hypothesis density filter [9, 10]. The inherent state-dependency of the detection probability is a significant obstacle when attempting to seek closed-form filtering solutions and thus is typically assumed to be state-independent by way of a zeroth-order approximation. However, the legitimacy of this approximation does not always hold, especially under large fluctuations in the detection probability, such as when objects are transitioning across a sensor's field of view [11]. Therefore, this work investigates possible alternatives to manage the state-dependent nature of the probability of detection using a proposed

nonlinear filter inspired by earlier work in intrinsic fault resistance [12]. Existing work on state-dependent probability of detection focuses primarily on application specific implementations or is presented in a more theoretical sense without much analysis [13, 14]. In response, this work seeks to investigate the effectiveness of different modeling fidelities by comparing the performance of three separate algorithms, which, in essence, helps identify an appropriate amount of modeling.

As filtering algorithms are generally considered the core of estimation schemes, much of this paper is dedicated to the background and derivation of the proposed filter, which is constructed via a single-target, fault-cognizant model that is capable of processing multiple measurements simultaneously within the update, as well as accounting for both valid and faulty sensor returns. To achieve improved nonlinear, non-Gaussian operations, the Bayesian filtering solution is actualized via Gaussian mixture approximations of the state distribution. As false alarm modeling is not of primary interest to this work, the typical assumptions are taken wherein faulty returns are independent events temporally Poisson distributed and spatially uniform distributed, an assumption legitimized in [12]. The filter presented here is particularly useful for any sensor where one valid measurement is returned by the sensor at most, but any number of false alarms may be returned as well.

The efficacy of the proposed filter is evaluated with several Monte-Carlo simulations, where it is tested against a Gaussian mixture extended Kalman filter (GMEKF), which is used as a performance baseline. The purpose of the GMEKF is to provide an algorithmic equivalent that operates with the least amount of required modeling. Additionally, a special version of the proposed filter, wherein the probability of detection is approximated as state-independent, is also presented for use as an additional comparison and can be considered a medium fidelity framework. To examine the filter's behavior more clearly, a simulation with linear dynamics and a one-dimensional sensor model is included, as to eliminate any artifacts resulting from nonlinearities. The results indicate that including a state-dependent probability of detection improves the accuracy of the filtering solution and promotes increased robustness against large variations in the detection probability. A second, more complex, simulation is also provided, where an in-depth model for probability of detection is created for an orbit determination scenario of a satellite in a cluttered environment. The performance of all three filters is analyzed and compared against one another, followed by corresponding discussion.

2. MOTIVATION

The notion that the receipt (or lack thereof) of a valid, target-oriented detection, should the probability of said detection be state-dependent, contains information of the state itself is by no means a novel concept [14]. While historically well-established, finding closed-form solutions for Bayesian filters can be an arduous task, even without the inclusion of detection probability. It is necessary to use restrictive linear and distribution assumptions in order to arrive at the Kalman filter, one of the simplest estimators [15, 16]. Therefore, it is natural to expect that the insertion of additional state-dependent terms, in this case the probability of detection, will only further complicate any closed-form filtering derivations. Concurrently, if modeled correctly, a state-dependent probability of detection will also directly improve the filter's state estimate, as, again, the Bayesian update naturally seeks to optimally fuse all available state information. Thus, while approximating the detection probability as state-independent [17], or even discounting it outright, can significantly ease finding closed-form filtering solutions, it also denies the estimator access to any state information inherent to the probability of detection. It is therefore not only important to examine how a state-dependent probability of detection might improve filtering operations, but also to determine if any benefits that result from the ancillary modeling are worthwhile, or if the complications incurred by the state-dependency should be avoided.

3. METHODOLOGY

This section details the creation of a fault-cognizant likelihood, similar to the one presented in [12], which is then used to create an intrinsically fault resistant (IFR) update. Key assumptions are then defined that allow the development of closed-form, nonlinear filtering algorithms, where multiple treatments of the probability of detection are considered. Note that the nomenclature IFR refers to filtering schemes that explicitly model faulty measurements and is used to distinguish such methods from those that rely on more *ad hoc* mechanisms (i.e. residual editing) to ensure robust filtering operations.

3.1 Fault-Cognizant Likelihood

Consider the typical sensor model, where a single valid sensor return $\boldsymbol{z}_v$ is drawn from the likelihood

$$\boldsymbol{z}_v \sim \ell_v(\boldsymbol{z}_v|\boldsymbol{x}) \ , \tag{1}$$

where $\boldsymbol{x}$ is the state vector. In an ideal sensing environment with perfect equipment, where all measurements are generated according to a single, known model, the classic measurement likelihood of Eq. (1) is theoretically all that is required for successful filtering operations. However, in many applications, there is a clear need for estimators that not only model measurements that are considered valid sensor returns, but also account for faulty, or undesirable, sensor returns. Such scenarios can include detecting objects in cluttered environments, accounting for false alarms triggered by background sensor noise [18], and even safeguarding against unlikely sensor malfunctions [2]. As such, it is often good practice to include some level of fault tolerance within the estimation framework. In this work, the concepts expressed in [12] are expanded upon, wherein a fault-cognizant likelihood is derived that accounts for both faulty and valid returns. The key difference in this work is that, instead of allowing only a single measurement per sensor scan, as is done in [12], the likelihood proposed here permits multiple measurements to be included in each sensor scan. The reason for this alteration is, while the model presented in [12] is simpler, as it is intended to be a direct replacement for residual editing [5], the probabilities of detection and false alarm are necessarily interdependent when a scan is restricted to a single measurement, such that they are both quantified by a single function, termed the probability of validity. While the method in [12] is useful for certain applications, it also prevents the independent modeling of valid and invalid detection occurrences. As a result, it disallows the use of the probabilities of detection and false alarm, which are two fairly prevalent topics of research having models that are readily available for various sensors. This interdependency can be obviated by permitting multiple returns per sensor scan, as the receipt of valid and faulty measurements are no longer mutually exclusive events. Thus, this work assumes sensor operations abide by the following constraints:

1. a scan can contain, at most, one valid measurement;
2. a scan can contain any number of invalid, or faulty, measurements; and
3. a scan can contain no measurements.

Consider that, upon each activation, a sensor returns a measurement set of m measurements as $\boldsymbol{Z} = \{\boldsymbol{z}_1, \boldsymbol{z}_2, \ldots, \boldsymbol{z}_m\}$. Furthermore, it is assumed that a single target is present, and the target can, at most, produce a single valid measurement return $\boldsymbol{z}_v \in \mathbb{V}$ pursuant to some unspecified likelihood function $\ell_v(\boldsymbol{z}|\boldsymbol{x})$. Accordingly, faulty measurements are generated via likelihood $\ell_f(\boldsymbol{z}|\boldsymbol{x})$, are denoted by $\boldsymbol{z}_f \in \mathbb{F}$, and are assumed to be independent of one another. The total likelihood of this model (in the event of m measurements) is

$$\ell(\boldsymbol{Z}|\boldsymbol{x}) = \Pr\left(\boldsymbol{Z} \cap m \cap \left[(\boldsymbol{z}_v \subseteq \boldsymbol{Z}) \cup (\boldsymbol{z}_v \nsubseteq \boldsymbol{Z})\right] \middle| \boldsymbol{x}\right) \ , \tag{2}$$

where $(\boldsymbol{z}_v \subseteq \boldsymbol{Z})$ denotes the event that the measurement set contains a valid target generated measurement, and the event where the detection is missed is $(\boldsymbol{z}_v \nsubseteq \boldsymbol{Z})$. As the target is either detected or is not, these are mutually exclusive events, such that Eq. (2) becomes

$$\ell(\boldsymbol{Z}|\boldsymbol{x}) = \Pr\left(\boldsymbol{Z} \cap m \cap (\boldsymbol{z}_v \subseteq \boldsymbol{Z}) \middle| \boldsymbol{x}\right) + \Pr\left(\boldsymbol{Z} \cap m \cap (\boldsymbol{z}_v \nsubseteq \boldsymbol{Z}) \middle| \boldsymbol{x}\right) \ . \tag{3}$$

As a direct result of conditional probability, Eq. (3) becomes

$$\ell(\boldsymbol{Z}|\boldsymbol{x}) = \Pr\left(\boldsymbol{Z} \cap m \middle| (\boldsymbol{z}_v \subseteq \boldsymbol{Z}) \cap \boldsymbol{x}\right) \Pr\left((\boldsymbol{z}_v \subseteq \boldsymbol{Z}) \middle| \boldsymbol{x}\right) + \Pr\left(\boldsymbol{Z} \cap m \middle| (\boldsymbol{z}_v \nsubseteq \boldsymbol{Z}) \cap \boldsymbol{x}\right) \Pr\left((\boldsymbol{z}_v \nsubseteq \boldsymbol{Z}) \middle| \boldsymbol{x}\right) \ , \tag{4}$$

where $\Pr\left((\boldsymbol{z}_v \subseteq \boldsymbol{Z}) \middle| \boldsymbol{x}\right) = p_D(\boldsymbol{x})$ is the state-dependent probability that the target will be detected by the sensor. Therefore, the probability that the target is not detected must be complimentary such that $\Pr\left((\boldsymbol{z}_v \nsubseteq \boldsymbol{Z}) \middle| \boldsymbol{x}\right) = 1 - p_D(\boldsymbol{x})$. Therefore, Eq. (4) can be expressed as

$$\ell(\boldsymbol{Z}|\boldsymbol{x}) = p_D(\boldsymbol{x}) \Pr\left(\boldsymbol{Z} \cap m | (\boldsymbol{z}_v \subseteq \boldsymbol{Z}) \cap \boldsymbol{x}\right) + [1 - p_D(\boldsymbol{x})] \Pr\left(\boldsymbol{Z} \cap m | (\boldsymbol{z}_v \nsubseteq \boldsymbol{Z}) \cap \boldsymbol{x}\right) \ . \tag{5}$$

Attention is now directed towards $\Pr\left(\boldsymbol{Z} \cap m \middle| (\boldsymbol{z}_v \subseteq \boldsymbol{Z}) \cap \boldsymbol{x}\right)$, which is the probability that (in this case) exactly one valid measurement is returned by the target. As a result, $m-1$ measurements must clearly be faulty. For now, consider

the event where the ith measurement is valid, denoted by $(\boldsymbol{z}_i = \boldsymbol{z}_v)$, where the collection of events for $i = 1, 2, \ldots, m$ is mutually exclusive, such that the probability becomes

$$\Pr\left(\boldsymbol{Z} \cap m \middle| (\boldsymbol{z}_v \subseteq \boldsymbol{Z}) \cap \boldsymbol{x}\right) = \sum_{i=1}^{m} \Pr\left(\boldsymbol{Z} \cap m \cap (\boldsymbol{z}_i = \boldsymbol{z}_v) \middle| (\boldsymbol{z}_v \subseteq \boldsymbol{Z}) \cap \boldsymbol{x}\right) . \tag{6}$$

The properties of conditional events can again be utilized such that Eq. (6) becomes

$$\Pr\left(\boldsymbol{Z} \cap m \middle| (\boldsymbol{z}_v \subseteq \boldsymbol{Z}) \cap \boldsymbol{x}\right) = \sum_{i=1}^{m} \Pr\left(\boldsymbol{Z} \cap m \middle| (\boldsymbol{z}_i = \boldsymbol{z}_v) \cap (\boldsymbol{z}_v \subseteq \boldsymbol{Z}) \cap \boldsymbol{x}\right) \Pr\left((\boldsymbol{z}_i = \boldsymbol{z}_v) \middle| (\boldsymbol{z}_v \subseteq \boldsymbol{Z}) \cap \boldsymbol{x}\right) . \tag{7}$$

Here, if it is assumed that measurements are unordered, and that any $\boldsymbol{z}_i$ has an equal probability of being the valid measurement, then the probability of the event $(\boldsymbol{z}_i = \boldsymbol{z}_v)$, given $\boldsymbol{Z}$ contains a valid measurement, is uniformly distributed over the number of measurements as

$$\Pr\left((\boldsymbol{z}_i = \boldsymbol{z}_v) \middle| (\boldsymbol{z}_v \subseteq \boldsymbol{Z}) \cap \boldsymbol{x}\right) = \begin{cases} \frac{1}{m} , & \text{if } m > 0 \\ 0 , & \text{otherwise} \end{cases} , \tag{8}$$

such that Eq. (7) becomes

$$\Pr\left(\boldsymbol{Z} \cap m \middle| (\boldsymbol{z}_v \subseteq \boldsymbol{Z}) \cap \boldsymbol{x}\right) = \frac{1}{m} \sum_{i=1}^{m} \Pr\left(\boldsymbol{Z} \cap m \middle| (\boldsymbol{z}_i = \boldsymbol{z}_v) \cap (\boldsymbol{z}_v \subseteq \boldsymbol{Z}) \cap \boldsymbol{x}\right) . \tag{9}$$

Yet again making use of conditional probabilities, Eq. (9) takes the form

$$\Pr\left(\boldsymbol{Z} \cap m \middle| (\boldsymbol{z}_v \subseteq \boldsymbol{Z}) \cap \boldsymbol{x}\right) = \frac{1}{m} \sum_{i=1}^{m} \Pr\left(\boldsymbol{Z} \middle| m \cap (\boldsymbol{z}_i = \boldsymbol{z}_v) \cap (\boldsymbol{z}_v \subseteq \boldsymbol{Z}) \cap \boldsymbol{x}\right) \Pr\left(m \middle| (\boldsymbol{z}_i = \boldsymbol{z}_v) \cap (\boldsymbol{z}_v \subseteq \boldsymbol{Z}) \cap \boldsymbol{x}\right) . \tag{10}$$

The term $\Pr\left(m \middle| (\boldsymbol{z}_i = \boldsymbol{z}_v) \cap (\boldsymbol{z}_v \subseteq \boldsymbol{Z}) \cap \boldsymbol{x}\right)$ is the probability that m measurements are generated, provided one of the measurements is valid. Equivalently, this can be considered the probability that $m-1$ measurements are faulty, or

$$\Pr\left(m \middle| (\boldsymbol{z}_i = \boldsymbol{z}_v) \cap (\boldsymbol{z}_v \subseteq \boldsymbol{Z}) \cap \boldsymbol{x}\right) = p_F(m-1, \boldsymbol{x}) ,$$

where $p_F(m-1, \boldsymbol{x})$ signifies the state-dependent probability of $m-1$ false alarms occurring. Thus, Eq. (10) becomes

$$\Pr\left(\boldsymbol{Z} \cap m \middle| (\boldsymbol{z}_v \subseteq \boldsymbol{Z}) \cap \boldsymbol{x}\right) = \frac{p_F(m-1, \boldsymbol{x})}{m} \sum_{i=1}^{m} \Pr\left(\boldsymbol{Z} \middle| m \cap (\boldsymbol{z}_i = \boldsymbol{z}_v) \cap (\boldsymbol{z}_v \subseteq \boldsymbol{Z}) \cap \boldsymbol{x}\right) . \tag{11}$$

At this point it is recalled that faulty measurements are independently and identically distributed according to the faulty likelihood function $\ell_f(\boldsymbol{z}|\boldsymbol{x})$ and valid target generated measurements are distributed according to the valid likelihood function $\ell_v(\boldsymbol{z}|\boldsymbol{x})$. Therefore, it can be shown that the remaining term inside the summation of Eq. (11) becomes

$$\begin{aligned} \Pr\left(\boldsymbol{Z} \middle| m \cap (\boldsymbol{z}_i = \boldsymbol{z}_v) \cap (\boldsymbol{z}_v \subseteq \boldsymbol{Z}) \cap \boldsymbol{x}\right) &= \Pr\left(\boldsymbol{z}_1 \cap \boldsymbol{z}_2 \cap \cdots \cap \boldsymbol{z}_m \middle| m \cap (\boldsymbol{z}_i = \boldsymbol{z}_v) \cap (\boldsymbol{z}_v \subseteq \boldsymbol{Z}) \cap \boldsymbol{x}\right) \\ &= \prod_{j=1}^{m} \Pr\left(\boldsymbol{z}_j \middle| m \cap (\boldsymbol{z}_i = \boldsymbol{z}_v) \cap (\boldsymbol{z}_v \subseteq \boldsymbol{Z}) \cap \boldsymbol{x}\right) \\ &= \Pr\left(\boldsymbol{z}_i \middle| m \cap (\boldsymbol{z}_i = \boldsymbol{z}_v) \cap (\boldsymbol{z}_v \subseteq \boldsymbol{Z}) \cap \boldsymbol{x}\right) \prod_{\substack{j=1 \\ j \neq i}}^{m} \Pr\left(\boldsymbol{z}_j \middle| m \cap (\boldsymbol{z}_i = \boldsymbol{z}_v) \cap (\boldsymbol{z}_v \subseteq \boldsymbol{Z}) \cap \boldsymbol{x}\right) \\ &= \ell_v(\boldsymbol{z}_i|\boldsymbol{x}) \prod_{\substack{j=1 \\ j \neq i}}^{m} \ell_f(\boldsymbol{z}_j|\boldsymbol{x}) . \end{aligned} \tag{12}$$

Substituting the result of Eq. (12) into Eq. (11) yields

$$\Pr\left(\boldsymbol{Z} \cap m \middle| (\boldsymbol{z}_v \subseteq \boldsymbol{Z}) \cap \boldsymbol{x}\right) = \frac{p_F(m-1, \boldsymbol{x})}{m} \sum_{i=1}^{m} \left\{ \ell_v(\boldsymbol{z}_i|\boldsymbol{x}) \prod_{\substack{j=1 \\ j \neq i}}^{m} \ell_f(\boldsymbol{z}_j|\boldsymbol{x}) \right\} . \tag{13}$$

Next, Eq. (5) is revisited to address the term $\Pr\left(\boldsymbol{Z}\cap m\middle|(\boldsymbol{z}_v\nsubseteq\boldsymbol{Z})\cap\boldsymbol{x}\right)$. By conditional probability

$$\Pr\left(\boldsymbol{Z}\cap m\middle|(\boldsymbol{z}_v\nsubseteq\boldsymbol{Z})\cap\boldsymbol{x}\right)=\Pr\left(\boldsymbol{Z}\middle|m\cap(\boldsymbol{z}_v\nsubseteq\boldsymbol{Z})\cap\boldsymbol{x}\right)\Pr\left(m\middle|(\boldsymbol{z}_v\nsubseteq\boldsymbol{Z})\cap\boldsymbol{x}\right)\ , \tag{14}$$

where $\Pr\left(m\middle|(\boldsymbol{z}_v\nsubseteq\boldsymbol{Z})\cap\boldsymbol{x}\right)$ is the probability that m measurements are generated, provided that none of them are valid. This is also the probability that all m measurements are faulty, denoted by

$$\Pr\left(m\middle|(\boldsymbol{z}_v\nsubseteq\boldsymbol{Z})\cap\boldsymbol{x}\right)=p_F(m,\boldsymbol{x})\ , \tag{15}$$

where, at this point, $p_F(\cdot)$ can be state-dependent. Similar to the developments of Eq. (12), the expression $\Pr\left(\boldsymbol{Z}\middle|m\cap(\boldsymbol{z}_v\nsubseteq\boldsymbol{Z})\cap\boldsymbol{x}\right)$ can be transformed as

$$\begin{aligned}\Pr\left(\boldsymbol{Z}\middle|m\cap(\boldsymbol{z}_v\nsubseteq\boldsymbol{Z})\cap\boldsymbol{x}\right)&=\Pr\left(\boldsymbol{z}_1\cap\boldsymbol{z}_2\cap\cdots\cap\boldsymbol{z}_m\middle|m\cap(\boldsymbol{z}_v\nsubseteq\boldsymbol{Z})\cap\boldsymbol{x}\right)\\&=\prod_{j=1}^{m}\Pr\left(\boldsymbol{z}_j\middle|m\cap(\boldsymbol{z}_v\nsubseteq\boldsymbol{Z})\cap\boldsymbol{x}\right)\\&=\prod_{j=1}^{m}\ell_f(\boldsymbol{z}_j|\boldsymbol{x})\ ,\end{aligned} \tag{16}$$

assuming, again, that the faulty measurements are independently and identically distributed according to the faulty likelihood. Equations (15) and (16) are enough to show that Eq. (14) becomes

$$\Pr\left(\boldsymbol{Z}\cap m\middle|(\boldsymbol{z}_v\nsubseteq\boldsymbol{Z})\cap\boldsymbol{x}\right)=p_F(m,\boldsymbol{x})\prod_{j=1}^{m}\ell_f(\boldsymbol{z}_j|\boldsymbol{x})\ . \tag{17}$$

Accordingly, the results of Eqs. (13) and (17) allow Eq. (5) to ultimately be expressed as

$$\ell(\boldsymbol{Z}|\boldsymbol{x})=p_D(\boldsymbol{x})\frac{p_F(m-1,\boldsymbol{x})}{m}\sum_{i=1}^{m}\left\{\ell_v(\boldsymbol{z}_i|\boldsymbol{x})\prod_{\substack{j=1\\j\neq i}}^{m}\ell_f(\boldsymbol{z}_j|\boldsymbol{x})\right\}+[1-p_D(\boldsymbol{x})]p_F(m,\boldsymbol{x})\prod_{j=1}^{m}\ell_f(\boldsymbol{z}_j|\boldsymbol{x})\ . \tag{18}$$

Special Case 1 [$m=0$]: It should be noted that due to a division by m, the case where there are no measurements seems to produce a singularity, at a glance. While many filters will not perform an update unless at least one measurement is received, the fact that no measurement is returned can provide information in and of itself. Therefore, in situations where $\boldsymbol{Z}=\{\emptyset\}$, Eq. (18) becomes

$$\ell(\emptyset|\boldsymbol{x})=[1-p_D(\boldsymbol{x})]p_F(0,\boldsymbol{x})\ , \tag{19}$$

as the first term vanishes—since the probability of -1 false alarms is zero as well as the probability of Eq. (8)—and by definition $\prod_{j=1}^{0}\ell_f(\boldsymbol{z}_j|\boldsymbol{x})=1$. In typical single-target filtering applications that do not account for state-dependent probabilities, when no measurements are produced, no update is performed. Leveraging the developments here, however, an update can be performed due to the state-dependent probabilities of detection and false alarm.

Special Case 2 [$m=1$]: In the instance where exactly a single measurement is received, or $\boldsymbol{Z}=\{\boldsymbol{z}\}$, it can be shown that the likelihood of Eq. (18) becomes

$$\ell(\boldsymbol{z}|\boldsymbol{x})=p_D(\boldsymbol{x})p_F(0,\boldsymbol{x})\ell_v(\boldsymbol{z}|\boldsymbol{x})+[1-p_D(\boldsymbol{x})]p_F(1,\boldsymbol{x})\ell_f(\boldsymbol{z}|\boldsymbol{x})\ , \tag{20}$$

which looks very similar to the fault-cognizant likelihood derived using the probability of validity in [12]. The main difference between the two likelihoods is that the probabilities of detection and false alarm are modeled independently of one another as two distinct parameters, whereas the probability of validity attempts to achieve a similar likelihood using only a single parameter.

3.2 Intrinsically Fault Resistant Bayesian Update

Many, if not most, filtering frameworks predicate their update on Bayes' rule [16]. The proposed filtering architecture of this work is no exception, and by using the fault-cognizant likelihood presented in Eq. (18), an intrinsically fault resistant (IFR) update, similar to the one proposed in [12], can be reached. It is well-known that Bayes' rule seeks an optimal posterior distribution $p^+(\boldsymbol{x})$ by fusing previous knowledge of $\boldsymbol{x}$ contained in the prior distribution, $p^-(\boldsymbol{x})$, with the incoming information made available by the likelihood $\ell(\boldsymbol{Z}|\boldsymbol{x})$. Mathematically, this is expressed as

$$p^+(\boldsymbol{x}) \propto p^-(\boldsymbol{x})\ell(\boldsymbol{Z}|\boldsymbol{x}) \ . \tag{21}$$

Substituting the likelihood of Eq. (18) into Bayes' rule of Eq. (21) yields a single-target, multi-measurement version of the IFR filter presented in [12], given by

$$p^+(\boldsymbol{x}) \propto p_D(\boldsymbol{x})\frac{p_F(m-1,\boldsymbol{x})}{m}p^-(\boldsymbol{x})\sum_{i=1}^{m}\left\{\ell_v(\boldsymbol{z}_i|\boldsymbol{x})\prod_{\substack{j=1\\ j\neq i}}^{m}\ell_f(\boldsymbol{z}_j|\boldsymbol{x})\right\} + [1-p_D(\boldsymbol{x})]p_F(m,\boldsymbol{x})p^-(\boldsymbol{x})\prod_{j=1}^{m}\ell_f(\boldsymbol{z}_j|\boldsymbol{x}) \ . \tag{22}$$

3.3 System Description

Although a valid Bayesian update is reached in Eq. (22), it is useful, if not necessary, to define a relevant system such that functional, closed-forms of the update can be produced. It is important to note that the restrictive assumptions made in this section do not embody all possible variations of the proposed IFR update and that different assumptions yield unique manifestations of the filter that may be just as practical, depending on the particular system at hand. In any case, this section characterizes specific forms for the probability of false alarm, the valid and invalid measurement distributions, and the prior state distribution that are deemed appropriate for this specific work.

3.3.1 Dynamics

The dynamics of the system refer to the way by which the state $\boldsymbol{x}$ evolves through time. In this work, a discrete form is taken, such that the current state $\boldsymbol{x}_k$ at some time t_k is recursively found as

$$\boldsymbol{x}_k = \boldsymbol{f}_k(\boldsymbol{x}_{k-1}) + \boldsymbol{w}_{p,k} \ , \tag{23}$$

where $\boldsymbol{f}_k(\cdot)$ is some nonlinear function of the state, and $\boldsymbol{w}_{p,k}$ is a process noise drawn from a multivariate, zero-mean, Gaussian distribution with covariance matrix $\boldsymbol{Q}_k$. As such, the state $\boldsymbol{x}_k$ can alternatively be expressed as

$$\boldsymbol{x}_k \sim p_g\left(\boldsymbol{x}_k\middle|\boldsymbol{f}_k(\boldsymbol{x}_{k-1}),\boldsymbol{Q}_k\right) \ , \tag{24}$$

where $p_g(\boldsymbol{a}|\boldsymbol{b},\boldsymbol{C})$ denotes a Gaussian distribution of $\boldsymbol{a}$ with mean $\boldsymbol{b}$ and covariance $\boldsymbol{C}$. Note that the time index k will be omitted for brevity throughout the majority of the paper, as the primary focus of this work is in the development of an update, the elements of which exist at a single time step.

3.3.2 Probability of False Alarm

Typically, false alarms are temporally modeled as state-independent events that occur at a rate that is described by a Poisson distribution, such that probability of m false alarms is

$$p_F(m) = \frac{\lambda^m}{m!\mathrm{e}^{\lambda}} \ , \tag{25}$$

where λ is the average number of false alarms expected to be contained in sensor return $\boldsymbol{Z}$. As the credibility of this assumption is well-established in [8], this work is justified in utilizing this assumption. Note that the Poisson probability of false alarm is used frequently in other clutter-related literature [19, 20].

3.3.3 Faulty Measurement Distribution

While false alarms (or faulty measurements) may be temporally Poisson, the manner in which they are spatially distributed remains equivocal. For instance, under some circumstances, false alarms may be returned nearer to the true state $\boldsymbol{x}$, and thus be state-dependent. Other cases may involve false alarms generated by some unknown model. For the most part, false alarms are inherently difficult to model, since they are relatively infrequent and, by definition, are a product of undesirable sensor behavior. The investigation in [12] helps address this, where assuming faulty measurements are spatially uniform across the sensor's field of view is suggested as a good practice, as the assumption is fairly robust to the effects of mismodeling. Therefore, for this work, the faulty likelihood of Eq. (20) becomes

$$\ell_f(\boldsymbol{z}) = \frac{1}{V} \ , \tag{26}$$

where V is the sensor volume given by

$$V = \prod_{\gamma=1}^{m} \left(b_\gamma - a_\gamma\right) \ , \quad \forall\, z_\gamma \in [a_\gamma, b_\gamma] \ ,$$

which is based on the assumption that each scalar component z_γ of vector measurement $\boldsymbol{z}$ will always exist inside the domain $[a_\gamma, b_\gamma]$. Accordingly, faulty vector measurements $\boldsymbol{z}_f$ are simulated by sampling the scalar components $z_{f,\gamma}$ as uniform random variables of the form

$$z_{f,\gamma} \sim p_u(a_\gamma, b_\gamma) \ .$$

3.3.4 Valid Measurement Distribution

While many different sensor models exist, most typically adhere to the form [21]

$$\boldsymbol{z}_v = \boldsymbol{h}(\boldsymbol{x}) + \boldsymbol{w}_s \ , \tag{27}$$

where $\boldsymbol{z}_v$ is a valid measurement vector, $\boldsymbol{h}(\cdot)$ is a nonlinear observation function of the state $\boldsymbol{x}$, and $\boldsymbol{w}_s$ is a random sensor noise drawn from a multivariate, zero-mean, Gaussian distribution with covariance matrix $\boldsymbol{R}$. In likelihood form, Eq. (27) is expressed as

$$\ell_v(\boldsymbol{z}|\boldsymbol{x}) = p_g\left(\boldsymbol{z}\middle|\boldsymbol{h}(\boldsymbol{x}), \boldsymbol{R}\right) \ . \tag{28}$$

As this observation model is commonplace, the valid measurements herein are assumed to be generated accordingly.

3.3.5 Gaussian Mixture Distribution

In order to realize the IFR update as a nonlinear filter—as opposed to a linear filter such as the EKF—a specific methodology must be selected. It is noted that while many nonlinear filtering frameworks exist, such as particle filtering and variational inference [22, 23], this work elects to use Gaussian mixture models (GMMs) to actualize the proposed update for several reasons. Firstly, Gaussian distributions possess many characteristics that aid in operations, making closed-form solutions easier to accomplish, especially in instances where the likelihood function is assumed to be Gaussian, as in Section 3.3.4. Secondly, while estimate extraction for many nonlinear methods tends to be complex and computationally intensive—such as clustering for particle filtering—there exist straightforward techniques for extracting estimates from Gaussian mixtures, another consequence of Gaussian properties [10].

For the proposed filter in this work, it is assumed that the posterior state distribution of the preceding time step t_{k-1} is available in the form of a Gaussian mixture (GM) comprised of L^+_{k-1} components as

$$p^+(\boldsymbol{x}_{k-1}) = \sum_{\ell=1}^{L^+_{k-1}} w^+_{\ell,k-1} p_g\left(\boldsymbol{x}_{k-1}\middle|\boldsymbol{m}^+_{\ell,k-1},\, \boldsymbol{P}^+_{\ell,k-1}\right) \ , \tag{29}$$

where $w^+_{\ell,k-1}$, $\boldsymbol{m}^+_{\ell,k-1}$, and $\boldsymbol{P}^+_{\ell,k-1}$ are the weight, mean, and covariance of the ℓth GM component, respectively. To be a proper probability density function (pdf), the GM of Eq. (29) must remain non-negative and integrate to unity across the support of $\boldsymbol{x}_{k-1}$, which can easily be enforced as [24]

$$w^+_{\ell,k-1} \geq 0 \; \forall \; \ell = 1,\, 2,\, \ldots,\, L^+_{k-1} \quad \text{and} \quad \sum_{\ell=1}^{L^+_{k-1}} w^+_{\ell,k-1} = 1 \; .$$

3.4 Probability of Detection

As with the system assumptions described in Section 3.3, an exact form of the probability of detection must also be provided before developing the update of Section 3.2 further. The selected characteristics of the probability of detection directly determine the precise form of the update in Eq. (22), and thus several different assumptions of $p_D(\boldsymbol{x})$ are examined.

3.4.1 Complete p_D Neglect

It is not uncommon to implement filters that, instead of accounting for events like missed-detections, elect to ignore such sensor behavior, resulting in updates that are only performed in the presence of some return from the sensor. Such filters commonly rely upon *ad hoc* extensions like residual editing to screen out erroneous sensor data in order to protect filtering operations [5]. While these filters tend to be computationally efficient enough for most real-time, onboard applications, they often exchange the burden of constructing more accurate sensor models for an increase in manual parameter tuning and the requirement for extra machinery that ensures robust filtering operations. Nevertheless, these filters, wherein probabilities such as $p_D(\boldsymbol{x})$ are completely disregarded, remain some of the most proven estimation architecture in existence, and should not be dismissed outright [25, 26, 27]. In this work, the Bayesian update of such filters no longer resembles the form of Eq. (22). Instead, it is performed as m sequential partial updates for each $\boldsymbol{z}_i \in \boldsymbol{Z}$, as each measurement is processed individually. Mathematically, this partial update recursion can be expressed by multiple applications of Bayes' rule as

$$p^{[i]}(\boldsymbol{x}) \propto p^{[i-1]}(\boldsymbol{x}) \ell_v(\boldsymbol{z}_i|\boldsymbol{x}) \quad \forall \; i = 1,\, 2,\, \ldots,\, m \; , \tag{30}$$

which is initialized as $p^{[0]}(\boldsymbol{x}) = p^-(\boldsymbol{x})$. The complete Bayesian posterior yielded by Eq. (30) is thus $p^+(\boldsymbol{x}) = p^{[m]}(\boldsymbol{x})$. Note that the update of Eq. (30) is introduced here as a basis for the residual editing filter in Section 3.5.2.

3.4.2 Zeroth-Order Approximation of p_D

There are many cases where estimation processes must be performed on a system where the probability of detection is state-dependent, and neglecting this reality outright is deemed overly detrimental to filtering operations. In cases such as these, the most common practice is to repeatedly approximate $p_D(\boldsymbol{x})$ during each update so that it may be treated as state-independent.

Considering that the probability of detection $p_D(\boldsymbol{x})$ is some function of state $\boldsymbol{x}$, it can be expressed via a Taylor series as

$$p_D(\boldsymbol{x}) = \sum_{n=0}^{\infty} \frac{(\boldsymbol{x}-\boldsymbol{m})^n}{n!} \left[\frac{\mathrm{d}^n p_D(\boldsymbol{x})}{\mathrm{d}\boldsymbol{x}^n} \right]_{\boldsymbol{x}=\boldsymbol{m}} .$$

A zeroth-order approximation of $p_D(\boldsymbol{x})$ is made by evaluating the $n = 0$ term of the Taylor series and neglecting higher order terms such that

$$\begin{aligned} p_D(\boldsymbol{x}) &= p_D(\boldsymbol{m}) + \mathscr{O}(\boldsymbol{x}) \\ &\approx p_D(\boldsymbol{m}) \; , \end{aligned} \tag{31}$$

where $\boldsymbol{m}$ can be selected in a number of ways, depending on the desired behavior of the filter [28]. As GMMs are of primary interest to this work, two potential, appropriate choices for $\boldsymbol{m}$ are considered. Firstly, a conservative approach

involves evaluating $p_D(\cdot)$ at the single point $\boldsymbol{m}^-$, which is the overall mean extracted from the prior GM of $p^-(\boldsymbol{x})$. This method produces an approximation of $p_D(\boldsymbol{x})$ before the processing of measurements, while also treating the GM as a single distribution by operating on the overall mean instead of the means of individual components. Alternatively, noting that $p_D(\boldsymbol{x})$ is often directly multiplied with individual GM components, the probability of detection may be approximated such that

$$\sum_{\ell=1}^{L} w_\ell p_D(\boldsymbol{x}) p_g\big(\boldsymbol{x}\big|\boldsymbol{m}_\ell, \boldsymbol{P}_\ell\big) \approx \sum_{\ell=1}^{L} w_\ell p_D(\boldsymbol{m}_\ell) p_g\big(\boldsymbol{x}\big|\boldsymbol{m}_\ell, \boldsymbol{P}_\ell\big) \ . \tag{32}$$

Note that this method allows for the calculation of $p_D(\boldsymbol{x})$ following measurement processing, while also evaluating $p_D(\boldsymbol{x})$ at each GM component across the entire distribution. As such, this second method is selected.

The approximation of Eq. (31) is generally considered acceptable in cases where the probability of detection is relatively constant compared to the state $\boldsymbol{x}$ [11]. In reality, this criteria is frequently violated; the probability of detection often switches between high and low values immediately as an object transitions over a sensor's field-of-view. Furthermore, the state information present in the true probability of detection is no longer accessible by the Bayesian update. As such, estimation operations in some systems may fail under this treatment of $p_D(\boldsymbol{x})$.

3.4.3 Gaussian Model of p_D

In order to derive a state-dependent probability of detection that lends itself to a closed form update, the beneficial properties of Gaussians mentioned in Section 3.3.5 are recalled. Consider the case where the probability of detection is modeled as a Gaussian distribution of the form

$$p_D(\boldsymbol{x}) = p_g\big(p_{D_z}\big|g(\boldsymbol{x}), R_D\big) \ , \tag{33}$$

where $g(\cdot)$ is some nonlinear function of the state, R_D is an associated variance, and p_{D_z} is a random variable drawn from the distribution of $p_D(\boldsymbol{x})$. Immediately, it is clear that some restrictions should be placed upon the elements of Eq. (33), as the probability of detection must exist within the interval [0,1] to be a valid probability. However, as the support of a Gaussian extends into infinity, the assumption of Eq. (33) poses a possible violation to this condition. To address this incompatibility, note that most sensor models are founded on similar fallacies; most sensor returns, while assumed to have Gaussian noise, are in reality bounded to exist within some finite interval pursuant to either some physical constraint of the system or behavior enforced upon the sensor. Therefore, pending careful treatment of Eq. (33), the Gaussian assumption on $p_D(\boldsymbol{x})$ should be comfortable enough for practical use. Even so, it is highly recommended to define $g(\cdot)$ and R_D such that the vast majority of the pdf $p_D(\boldsymbol{x})$ lies within $[0,1]$, as well as enforcing all p_{D_z} realizations to remain within that interval when simulated, even if the filtering algorithms treat p_{D_z} as unbounded.

Next, it is important to note that because a Gaussian random variable will remain a Gaussian under a linear transformation [21], that the probability of missed detection is distributed as

$$\big[1 - p_D(\boldsymbol{x})\big] = p_g\big(\big[1 - p_{D_z}\big]\big|\big[1 - g(\boldsymbol{x})\big], R_D\big) \ . \tag{34}$$

The reality of Eq. (34) is especially useful when developing the complete IFR update, as it reduces the number of GM components created with each update, as well as directly enforces non-negativity upon the posterior pdf. If, alternatively, the "1" and "$-p_D(\boldsymbol{x})$" are treated independently of one another, linearization errors can cause the pdf to go negative, while also producing more GM components within $p^+(\boldsymbol{x})$.

It should be mentioned that this conceptual Gaussian model of $p_D(\boldsymbol{x})$ can be extended into a GMM of $p_D(\boldsymbol{x})$, much like in [13]. This is especially useful in situations where the probability of detection cannot be adequately described by a single Gaussian distribution. However, the proposed model of Eq. (33) is sufficient for the purposes of this work, since the main goal is the investigation of different levels of $p_D(\boldsymbol{x})$ modeling.

3.5 Filtering Equations

This section achieves the Bayesian update of Section 3.2 by way of the assumptions presented in Sections 3.3 and 3.4, resulting in three distinct filters: a GMEKF with residual editing, an IFR filter with zeroth-order approximated $p_D(\boldsymbol{x})$, and an IFR filter with a Gaussian model of $p_D(\boldsymbol{x})$.

3.5.1 GMEKF Prediction

Often credited to the work of [29, 30], the Gaussian mixture extended Kalman filter (GMEKF)—sometimes referred to as the Gaussian sum filter—is one of the first closed-form methods for nonlinear filtering, and is named after the EKF for its algorithmic similarity only; it is not derived from the EKF, as it is constructed via a direct application of GMM pdfs onto the Chapman-Kolmogorov equation and Bayes' rule [31]. Nevertheless, the nomenclature GMEKF is the most prevalent title currently given to the filter, and is thus used here for the sake of clarity.

As the main focus of this work is the development and evaluation of the proposed update, all filters herein are equipped with identical dynamic propagation equations, such that the predictive stage of each filter is essentially equivalent. This helps ensure a fair comparison between the various update stages. Pursuant to the discussion of Section 3.3.5, where the previous posterior is taken as the GM of Eq. (29), the prior distribution that results from the GMEKF propagation of Eq. (29) is also a GM of the form

$$p^{-}(\boldsymbol{x}_k) = \sum_{\ell=1}^{L_k^-} w_{\ell,k}^- p_g\big(\boldsymbol{x}_k \big| \boldsymbol{m}_{\ell,k}^-,\, \boldsymbol{P}_{\ell,k}^-\big)\ , \tag{35}$$

where $w_{\ell,k}^-$, $\boldsymbol{m}_{\ell,k}^-$, and $\boldsymbol{P}_{\ell,k}^-$ are the weight, mean, and covariance of the ℓth GM component, respectively. Recalling the dynamical system defined by Eqs. (23) and (24), the GM components of Eq. (35) are computed as [30]

$$L_k^- = L_{k-1}^+ \tag{36a}$$

$$w_{\ell,k}^- = w_{\ell,k-1}^+ \tag{36b}$$

$$\boldsymbol{m}_{\ell,k}^- = \boldsymbol{f}_k(\boldsymbol{m}_{\ell,k-1}^+) \tag{36c}$$

$$\boldsymbol{P}_{\ell,k}^- = \boldsymbol{F}_k(\boldsymbol{m}_{\ell,k-1}^+)\boldsymbol{P}_{\ell,k-1}^+\boldsymbol{F}_k^T(\boldsymbol{m}_{\ell,k-1}^+) + \boldsymbol{Q}_k\ , \tag{36d}$$

where $\boldsymbol{F}_k(\boldsymbol{m})$ is the Jacobian of the nonlinear dynamics function $\boldsymbol{f}_k(\cdot)$ evaluated at $\boldsymbol{x} = \boldsymbol{m}$. Note that the time index k will frequently be omitted for brevity in future equations.

3.5.2 GMEKF Update with Residual Editing

To provide a baseline reference for a filter where absolutely no modeling of the probability of detection is performed, a GMEKF outfitted with residual editing is presented, the derivation of which is detailed in [12]. Based on the discussion of Section 3.4.1 and assuming that $p^-(\boldsymbol{x})$ is available as Eq. (35), the partial Bayesian update of Eq. (30) when outfitted with residual editing becomes

$$p^{[i]}(\boldsymbol{x}) = \begin{cases} \sum\limits_{\ell=1}^{L^-} w_\ell^- p_g(\boldsymbol{x}|\boldsymbol{m}_\ell^-, \boldsymbol{P}_\ell^-)\ , & \text{if } \kappa_{\text{RE}} < \sum\limits_{\ell=1}^{L^-} w_\ell^- \left|2\pi\boldsymbol{W}_\ell\right|^{-\frac{1}{2}} \exp\big\{-\tfrac{1}{2}\Upsilon\big\} \\ \sum\limits_{\ell=1}^{L^+} w_\ell^+ p_g(\boldsymbol{x}|\boldsymbol{m}_\ell^+, \boldsymbol{P}_\ell^+)\ , & \text{if } \kappa_{\text{RE}} \geq \sum\limits_{\ell=1}^{L^-} w_\ell^- \left|2\pi\boldsymbol{W}_\ell\right|^{-\frac{1}{2}} \exp\big\{-\tfrac{1}{2}\Upsilon\big\} \end{cases}\ , \tag{37}$$

where

$$L^+ = L^- \qquad \boldsymbol{K}_\ell = \boldsymbol{P}_\ell^-\boldsymbol{H}^T(\boldsymbol{m}_\ell^-)\boldsymbol{W}_\ell^{-1}$$

$$w_\ell^+ = \frac{w_\ell^-\kappa_\ell}{\sum_{j=1}^{L^-} w_j^-\kappa_j} \qquad \kappa_{\text{RE}} = \sum_{\ell=1}^{L^-} w_\ell^-\kappa_\ell$$

$$\boldsymbol{m}_\ell^+ = \boldsymbol{m}_\ell^- + \boldsymbol{K}_\ell[\boldsymbol{z}_i - \boldsymbol{h}(\boldsymbol{m}_\ell^-)] \qquad \kappa_\ell = \left|2\pi\boldsymbol{W}_\ell\right|^{-\frac{1}{2}} \exp\Big\{-\frac{1}{2}[\boldsymbol{z}_i - \boldsymbol{h}(\boldsymbol{m}_\ell^-)]^T\boldsymbol{W}_\ell^{-1}[\boldsymbol{z}_i - \boldsymbol{h}(\boldsymbol{m}_\ell^-)]\Big\}$$

$$\boldsymbol{P}_\ell^+ = \boldsymbol{P}_\ell^- - \boldsymbol{K}_\ell\boldsymbol{H}(\boldsymbol{m}_\ell^-)\boldsymbol{P}_\ell^- \qquad \boldsymbol{W}_\ell = \boldsymbol{H}(\boldsymbol{m}_\ell^-)\boldsymbol{P}_\ell^-\boldsymbol{H}^T(\boldsymbol{m}_\ell^-) + \boldsymbol{R}\ ,$$

and where the components L^-, w_ℓ^-, $\boldsymbol{m}_\ell^-$, $\boldsymbol{P}_\ell^-$ are recursively defined as

$$p^{[i-1]}(\boldsymbol{x}) = \sum_{\ell=1}^{L^-} w_\ell^- p_g(\boldsymbol{x}|\boldsymbol{m}_\ell^-, \boldsymbol{P}_\ell^-)\ ,$$

which is initialized as $p^{[0]}(\boldsymbol{x}) = p^-(\boldsymbol{x})$. Setting the threshold value Υ in accordance to a χ^2 probability gate is considered a best practice, and at the very least provides an initial value that can expedite filter tuning [5].

3.5.3 IFR Update

Before accounting for the precise treatments of $p_D(\boldsymbol{x})$ listed in Section 3.4, a more general IFR update is derived by asserting the system assumptions of Section 3.3 onto the Bayesian update of Eq. (22). First, it is noted that the fault-cognizant likelihood equation of Eq. (18), when realized via the false alarm model of Eqs. (25) and (26), becomes

$$\begin{aligned}
\ell(\boldsymbol{Z}|\boldsymbol{x}) &= p_D(\boldsymbol{x})\frac{\lambda^{m-1}}{m(m-1)!\mathrm{e}^{\lambda}}\sum_{i=1}^{m}\left\{\ell_v(\boldsymbol{z}_i|\boldsymbol{x})\prod_{j=1,j\neq i}^{m}\frac{1}{V}\right\}+[1-p_D(\boldsymbol{x})]\frac{\lambda^m}{m!\mathrm{e}^{\lambda}}\prod_{j=1}^{m}\frac{1}{V} \\
&= p_D(\boldsymbol{x})\frac{\lambda^{m-1}}{m!\mathrm{e}^{\lambda}}\sum_{i=1}^{m}\left\{\ell_v(\boldsymbol{z}_i|\boldsymbol{x})\frac{1}{V^{m-1}}\right\}+[1-p_D(\boldsymbol{x})]\frac{\lambda^m}{m!\mathrm{e}^{\lambda}}\frac{1}{V^m} \\
&= \frac{\lambda^m}{m!\mathrm{e}^{\lambda}}\frac{1}{V^m}\left\{p_D(\boldsymbol{x})\frac{V}{\lambda}\sum_{i=1}^{m}\left[\ell_v(\boldsymbol{z}_i|\boldsymbol{x})\right]+[1-p_D(\boldsymbol{x})]\right\} \\
&= \frac{p_F(m)}{V^m}\left\{\frac{p_D(\boldsymbol{x})}{\kappa_c}\sum_{i=1}^{m}\left[\ell_v(\boldsymbol{z}_i|\boldsymbol{x})\right]+[1-p_D(\boldsymbol{x})]\right\}, \qquad (38)
\end{aligned}$$

where κ_c is a commonly used parameter refereed to as the clutter intensity, which relates the temporal and spatial distribution of false alarms as [8]

$$\kappa_c = \lambda\ell_f(\boldsymbol{z}) ,$$

which, in this case, is

$$\kappa_c = \frac{\lambda}{V} .$$

If the valid likelihood from Eq. (28) is then introduced to Eq. (38), the fault-cognizant likelihood further reduces to

$$\ell(\boldsymbol{Z}|\boldsymbol{x}) = \frac{p_F(m)}{V^m}\left\{\frac{1}{\kappa_c}\sum_{i=1}^{m}\left[p_D(\boldsymbol{x})p_g\big(\boldsymbol{z}_i\big|\boldsymbol{h}(\boldsymbol{x}),\boldsymbol{R}\big)\right]+[1-p_D(\boldsymbol{x})]\right\} . \qquad (39)$$

If the GM assumption of Section 3.3.5 is presumed, such that the prior becomes available as the GM in Eq. (35), then the equivalent Bayesian update from Eq. (22) becomes

$$p^+(\boldsymbol{x}) \propto \frac{1}{\kappa_c}\sum_{i=1}^{m}\sum_{\ell=1}^{L^-}\left\{p_D(\boldsymbol{x})w_\ell^- p_g\big(\boldsymbol{z}_i\big|\boldsymbol{h}(\boldsymbol{x}),\boldsymbol{R}\big)p_g\big(\boldsymbol{x}\big|\boldsymbol{m}_\ell^-,\boldsymbol{P}_\ell^-\big)\right\}+\sum_{\ell=1}^{L^-}\left\{[1-p_D(\boldsymbol{x})]w_\ell^- p_g\big(\boldsymbol{x}\big|\boldsymbol{m}_\ell^-,\boldsymbol{P}_\ell^-\big)\right\} . \qquad (40)$$

Here, the identity of Eq. (58) is applied to the product of Gaussians in Eq. (40) to produce

$$p^+(\boldsymbol{x}) \propto \frac{1}{\kappa_c}\sum_{i=1}^{m}\sum_{\ell=1}^{L^-}\left\{p_D(\boldsymbol{x})\kappa_{\ell,i}^+ w_\ell^- p_g\big(\boldsymbol{x}\big|\boldsymbol{m}_{\ell,i}^+,\boldsymbol{P}_{\ell,i}^+\big)\right\}+\sum_{\ell=1}^{L^-}\left\{[1-p_D(\boldsymbol{x})]w_\ell^- p_g\big(\boldsymbol{x}\big|\boldsymbol{m}_\ell^-,\boldsymbol{P}_\ell^-\big)\right\} . \qquad (41)$$

where

$$\boldsymbol{m}_{\ell,i}^+ = \boldsymbol{m}_\ell^- + \boldsymbol{K}_\ell\big(\boldsymbol{z}_i - \boldsymbol{h}(\boldsymbol{m}_\ell^-)\big) \qquad (42a)$$

$$\boldsymbol{P}_{\ell,i}^+ = \boldsymbol{P}_\ell^- - \boldsymbol{K}_\ell\boldsymbol{H}(\boldsymbol{m}_\ell^-)\boldsymbol{P}_\ell^- \qquad (42b)$$

$$\boldsymbol{W}_\ell = \boldsymbol{H}(\boldsymbol{m}_\ell^-)\boldsymbol{P}_\ell^-\boldsymbol{H}^T(\boldsymbol{m}_\ell^-)+\boldsymbol{R} \qquad (42c)$$

$$\boldsymbol{K}_\ell = \boldsymbol{P}_\ell^-\boldsymbol{H}^T(\boldsymbol{m}_\ell^-)\boldsymbol{W}_\ell^{-1} \qquad (42d)$$

$$\kappa_{\ell,i}^+ = p_g\big(\boldsymbol{z}_i\big|\boldsymbol{h}(\boldsymbol{m}_\ell^-),\boldsymbol{W}_\ell\big) , \qquad (42e)$$

and where $\boldsymbol{H}(\cdot)$ is the Jacobian of $\boldsymbol{h}(\cdot)$. The update of Eq. (41), along with the corresponding components of Eqs. (42), are the basis for each variation of the proposed IFR filter to be presented.

IFR Update with Zeroth-Order Approximated p_D: If a zeroth-order approximation of $p_D(\boldsymbol{x})$ is made, as per Section 3.4.2 and Eq. (32), the IFR update from Eqs. (41) and (42) becomes

$$p^+(\boldsymbol{x}) \propto \frac{1}{\kappa_c}\sum_{i=1}^{m}\sum_{\ell=1}^{L^-}\left\{p_D(\boldsymbol{m}_{\ell,i}^+)\kappa_{\ell,i}^+ w_\ell^- \, p_g\big(\boldsymbol{x}\big|\boldsymbol{m}_{\ell,i}^+,\boldsymbol{P}_{\ell,i}^+\big)\right\} + \sum_{\ell=1}^{L^-}\left\{[1-p_D(\boldsymbol{m}_\ell^-)]w_\ell^- \, p_g\big(\boldsymbol{x}\big|\boldsymbol{m}_\ell^-,\boldsymbol{P}_\ell^-\big)\right\}, \tag{43}$$

which, once proper normalization is accomplished, becomes

$$p^+(\boldsymbol{x}) = \sum_{i=1}^{m}\sum_{\ell=1}^{L^-}\left\{w_{\ell,i}^+ p_g\big(\boldsymbol{x}\big|\boldsymbol{m}_{\ell,i}^+,\boldsymbol{P}_{\ell,i}^+\big)\right\} + \sum_{\ell=1}^{L^-}\left\{w_\ell^+ p_g\big(\boldsymbol{x}\big|\boldsymbol{m}_\ell^-,\boldsymbol{P}_\ell^-\big)\right\}, \tag{44a}$$

where

$$w_{\ell,i}^+ = \frac{p_D(\boldsymbol{m}_{\ell,i}^+)\kappa_{\ell,i}^+ w_\ell^-}{\kappa_c \eta}, \quad w_\ell^+ = \frac{\big[1-p_D(\boldsymbol{m}_\ell^-)\big]w_\ell^-}{\eta}, \quad \text{and} \quad \eta = \sum_{\ell=1}^{L^-} w_\ell^-\left\{\big[1-p_D(\boldsymbol{m}_\ell^-)\big] + \frac{1}{\kappa_c}\sum_{i=1}^{m}\left[p_D(\boldsymbol{m}_{\ell,i}^+)\kappa_{\ell,i}^+\right]\right\}. \tag{44b}$$

Clearly, the posterior of Eq. (44a) is a GM of $L^- \times (m+1)$ components, such that an additional $L^- \times m$ components are generated with each filter iteration. Note that in the trivial case where $p_D(\boldsymbol{x})$ is taken to be a constant, the resulting update can simply be stated as simplified version of Eqs. (44a) and (44b) with $p_D(\boldsymbol{m}) \approx p_D$.

IFR Update with Gaussian p_D: This variant of the IFR update assumes that the probability of detection can be modeled as a Gaussian distribution with a mean that is a function of the state $\boldsymbol{x}$, per the discussion of Section 3.4.3. In this case, the IFR update of Eqs. (41) and (42) becomes

$$p^+(\boldsymbol{x}) \propto \frac{1}{\kappa_c}\sum_{i=1}^{m}\sum_{\ell=1}^{L^-}\left\{\kappa_{\ell,i}^+ w_\ell^- \, p_g\big(p_{D_z}\big|g(\boldsymbol{x}),R_D\big)p_g\big(\boldsymbol{x}\big|\boldsymbol{m}_{\ell,i}^+,\boldsymbol{P}_{\ell,i}^+\big)\right\} + \sum_{\ell=1}^{L^-}\left\{w_\ell^- \, p_g\big(\big[1-p_{D_z}\big]\big|\big[1-g(\boldsymbol{x})\big],R_D\big)p_g\big(\boldsymbol{x}\big|\boldsymbol{m}_\ell^-,\boldsymbol{P}_\ell^-\big)\right\} \tag{45}$$

following the proper substitutions of Eqs. (33) and (34). Next, two simultaneous applications of Eq. (58) are used to evaluate both Gaussian products to yield

$$p^+(\boldsymbol{x}) \propto \frac{1}{\kappa_c}\sum_{\ell=1}^{L^-}\sum_{i=1}^{m}\left\{w_\ell^- \, \kappa_{\ell,i}^+\kappa_{\ell,i}^{D^+} \, p_g(\boldsymbol{x}|\boldsymbol{m}_{\ell,i}^{D^+},\boldsymbol{P}_{\ell,i}^{D^+})\right\} + \sum_{\ell=1}^{L^-}\left\{w_\ell^- \, \kappa_\ell^{D^-} \, p_g(\boldsymbol{x}|\boldsymbol{m}_\ell^{D^-},\boldsymbol{P}_\ell^{D^-})\right\}, \tag{46}$$

where

$$\begin{aligned}
\boldsymbol{m}_{\ell,i}^{D^+} &= \boldsymbol{m}_{\ell,i}^+ + \boldsymbol{K}_{\ell,i}^{D^+}\big(p_{D_z} - g(\boldsymbol{m}_{\ell,i}^+)\big) & \boldsymbol{m}_\ell^{D^-} &= \boldsymbol{m}_\ell^- + \boldsymbol{K}_\ell^{D^-}\big(p_{D_z} - g(\boldsymbol{m}_\ell^-)\big)\\
\boldsymbol{P}_{\ell,i}^{D^+} &= \boldsymbol{P}_{\ell,i}^+ - \boldsymbol{K}_{\ell,i}^{D^+}\boldsymbol{G}(\boldsymbol{m}_{\ell,i}^+)\boldsymbol{P}_{\ell,i}^+ & \boldsymbol{P}_\ell^{D^-} &= \boldsymbol{P}_\ell^- - \boldsymbol{K}_\ell^{D^-}\boldsymbol{G}(\boldsymbol{m}_\ell^-)\boldsymbol{P}_\ell^-\\
W_{\ell,i}^{D^+} &= \boldsymbol{G}(\boldsymbol{m}_{\ell,i}^+)\boldsymbol{P}_{\ell,i}^+\boldsymbol{G}^T(\boldsymbol{m}_{\ell,i}^+) + R_D & W_\ell^{D^-} &= \boldsymbol{G}(\boldsymbol{m}_\ell^-)\boldsymbol{P}_\ell^-\boldsymbol{G}^T(\boldsymbol{m}_\ell^-) + R_D\\
\boldsymbol{K}_{\ell,i}^{D^+} &= \boldsymbol{P}_{\ell,i}^+\boldsymbol{G}^T(\boldsymbol{m}_{\ell,i}^+)\big[W_{\ell,i}^{D^+}\big]^{-1} & \boldsymbol{K}_\ell^{D^-} &= \boldsymbol{P}_\ell^-\boldsymbol{G}^T(\boldsymbol{m}_\ell^-)\big[W_\ell^{D^-}\big]^{-1}\\
\kappa_{\ell,i}^{D^+} &= p_g\Big(p_{D_z}\Big|g(\boldsymbol{m}_{\ell,i}^+),W_{\ell,i}^{D^+}\Big) & \kappa_\ell^{D^-} &= p_g\Big(\big[1-p_{D_z}\big]\Big|\big[1-g(\boldsymbol{m}_\ell^-)\big],W_\ell^{D^-}\Big).
\end{aligned}$$

Note that $\boldsymbol{G}(\cdot)$ is the Jacobian of $g(\cdot)$, and that the Jacobian of $1-g(\boldsymbol{x})$ is simply

$$\frac{\partial}{\partial \boldsymbol{x}}\Big(1-g(\boldsymbol{x})\Big) = -\boldsymbol{G}(\cdot).$$

Therefore, given the update of Eq. (46), the normalized Bayesian posterior is found to be

$$p^+(\boldsymbol{x}) = \sum_{\ell=1}^{L^-}\sum_{i=1}^{m}\left\{w_{\ell,i}^{D^+} \, p_g(\boldsymbol{x}|\boldsymbol{m}_{\ell,i}^{D^+},\boldsymbol{P}_{\ell,i}^{D^+})\right\} + \sum_{\ell=1}^{L^-}\left\{w_\ell^{D^-} \, p_g(\boldsymbol{x}|\boldsymbol{m}_\ell^{D^-},\boldsymbol{P}_\ell^{D^-})\right\}, \tag{47a}$$

where

$$w_{\ell,i}^{D^+} = \frac{w_\ell^- \kappa_{\ell,i}^+ \kappa_{\ell,i}^{D^+}}{\kappa_c \eta} , \qquad w_\ell^{D^-} = \frac{w_\ell^- \kappa_\ell^{D^-}}{\eta} , \qquad \text{and} \quad \eta = \sum_{\ell=1}^{L^-} w_\ell^- \left\{ \kappa_\ell^{D^-} + \frac{1}{\kappa_c} \sum_{i=1}^{m} \left[\kappa_{\ell,i}^+ \kappa_{\ell,i}^{D^+} \right] \right\} . \tag{47b}$$

which corresponds to a Gaussian mixture of $L^- \times (m+1)$ components. In order for this update to be accomplished, some value of p_{D_z} must be reported to the filter by some external source.

4. RESULTS AND DISCUSSION

4.1 Falling Body Simulation

This simulation is specifically designed with low complexity in order to evaluate the behavior of the different update algorithms proposed and is based on the system presented in [32], where the unidimensional position and velocity of a falling object is tracked by a ground-based observer, yielding the state vector $\boldsymbol{x} = [y\ \dot{y}]^T$. As all filters are equipped with identical predictive algorithms, a linear dynamical model of the form $\ddot{y} = -g$, where $g = 32.2\,\text{ft}/\text{s}^2$, is taken to eliminate any linearization errors that might result from a system with nonlinear dynamics. The initial distribution is taken to be a GM that is formed by splitting the multivariate Gaussian distribution

$$p^+(\boldsymbol{x}_0) = p_g\left(\boldsymbol{x}_0 \middle| \boldsymbol{m}_0, \boldsymbol{P}_0\right) , \tag{48}$$

where

$$\boldsymbol{m}_0 = \begin{bmatrix} 1000\text{ ft} \\ 0\ \text{ft}/\text{s} \end{bmatrix} \qquad \text{and} \qquad \boldsymbol{P}_0 = \begin{bmatrix} 2000\text{ft}^2 & 0 \\ 0 & 500(\text{ft}/\text{s})^2 \end{bmatrix} ,$$

into a 9 component GM by application of a three-component splitting library across both position and velocity [33]. To manage the number of components generated by the IFR filter, the pruning technique detailed in [12] is utilized, where $w_{\text{thresh}} = 10^{-8}$. Note that the number of components remains tractable at all times, with the maximum number of components accumulated in each trial, averaged across all MC trials, being 191 for the IFR filter with Gaussian model and 70 for the IFR filter with zeroth-order approximation. Since the GMEKF with residual editing does not naturally generate additional components, the number of components never exceeds 9.

Each filter is subjected to 10,000 Monte Carlo (MC) trials, where each trial consists of a 10 second interval where measurement scans are simulated every second. Per the discussion of Section 3.3.3, faulty measurements are taken to be spatially uniform across a sensor field-of-view on the interval [0 ft, 5000 ft]—such that the sensor volume is $V_{\text{range}} = 5000$ ft—and are temporally Poisson according to Eq. (25) with rate parameter specified in each analysis.

Valid measurement returns conform to the nonlinear model of Eq. (27), where

$$h(\boldsymbol{x}) = \sqrt{r_{\text{obs},1}^2 + (y - r_{\text{obs},2})^2} \tag{49}$$

is the range to the object, and where $r_{\text{obs},1} = 500$ ft and $r_{\text{obs},2} = 250$ ft describe the offset position of the observer and $R = 10\ \text{ft}^2$ is the sensor noise variance corresponding to Eq. (28). The probability of detection is randomly sampled from the Gaussian model of Section 3.4.3, where

$$g(\boldsymbol{x}) = p_{D,\max} - \frac{p_{D,\max} - p_{D,\min}}{V_{\text{range}}} h(\boldsymbol{x}) \tag{50}$$

is the nonlinear function of the mean. Note that values for $p_{D,\max}$, $p_{D,\min}$, and R_D are specified for each analysis.

At each time step t_k, the average filter error standard deviations $\sigma_{\text{filt},\xi}$ and Monte Carlo error standard deviations $\sigma_{\text{MC},\xi}$ of the ξth state of $\boldsymbol{x}$ are calculated as

$$\sigma_{\text{filt},\xi} = \sqrt{\frac{1}{N_{\text{MC}}} \sum_{i=1}^{N_{\text{MC}}} \hat{\boldsymbol{P}}_k^i(\xi,\xi)} \qquad \text{and} \qquad \sigma_{\text{MC},\xi} = \sqrt{\frac{1}{N_{\text{MC}}} \sum_{i=1}^{N_{\text{MC}}} \left(e_\xi^i - \bar{e}_\xi\right)^2} , \tag{51}$$

where N_{MC} is the number of MC trials, and where e_ξ^i is the ξth element of the estimation error vector for the ith MC trial defined as

$$\boldsymbol{e}_k^i = \boldsymbol{x}_k^i - \hat{\boldsymbol{x}}_k^i \ ,$$

with the average filter error at time t_k calculated as

$$\bar{e}_\xi = \frac{1}{N_{\mathrm{MC}}} \sum_{i=1}^{N_{\mathrm{MC}}} e_\xi^i \ . \tag{52}$$

Note that $\boldsymbol{x}_k^i$ is the truth at t_k of the ith MC trial, and that $\hat{\boldsymbol{x}}_k^i$ and $\hat{\boldsymbol{P}}_k^i$ are the extracted state and covariance estimate, respectively, of the ith trial at t_k calculated using algorithms in [12]. Generally, a well-behaved filter yields $\sigma_{\mathrm{filt},\xi} \approx \sigma_{\mathrm{MC},\xi}$, where $\sigma_{\mathrm{filt},\xi} > \sigma_{\mathrm{MC},\xi}$ indicates conservative estimation, while $\sigma_{\mathrm{filt},\xi} < \sigma_{\mathrm{MC},\xi}$ signifies overconfident filter performance. In practice, a conservative filter is customarily more desirable than one that is overconfident, as overconfidence often leads to issues such as filter divergence [34].

4.1.1 Filter Configurations

GMEKF with Residual Editing: This filter, outlined in Section 3.5.2, is considered a framework with the lowest amount of $p_D(\boldsymbol{x})$ modeling, where the only mechanism being applied is a residual check. Not only does it not necessitate a p_{D_z} return, it also does not need any function of $p_D(\boldsymbol{x})$ to be specified. Only a selection residual editing threshold of $\Upsilon = 16.45$ is required, which corresponds to a χ^2 value of 99.95%, a value which was selected after preliminary tuning of the filter.

IFR Filter with Zeroth-Order Approximated p_D: This version of the proposed filter is constructed in Eqs. (44), and is representative of a moderate amount of $p_D(\boldsymbol{x})$ modeling as it also does not require a value of p_{D_z} to be reported. Instead, it calculates a zeroth-order approximation of the probability of detection as

$$p_D(\boldsymbol{m}) = g(\boldsymbol{m}) \ .$$

IFR Filter with State-Dependent p_D: This filter, described in Eqs. (47), possesses the highest level of $p_D(\boldsymbol{x})$ modeling and requires a reported probability of detection return, p_{D_z}, from some external sensing operation. In this simulation, it is randomly sampled according to the distribution of Eq. (33) with mean $g(\boldsymbol{x})$ defined by Eq. (50).

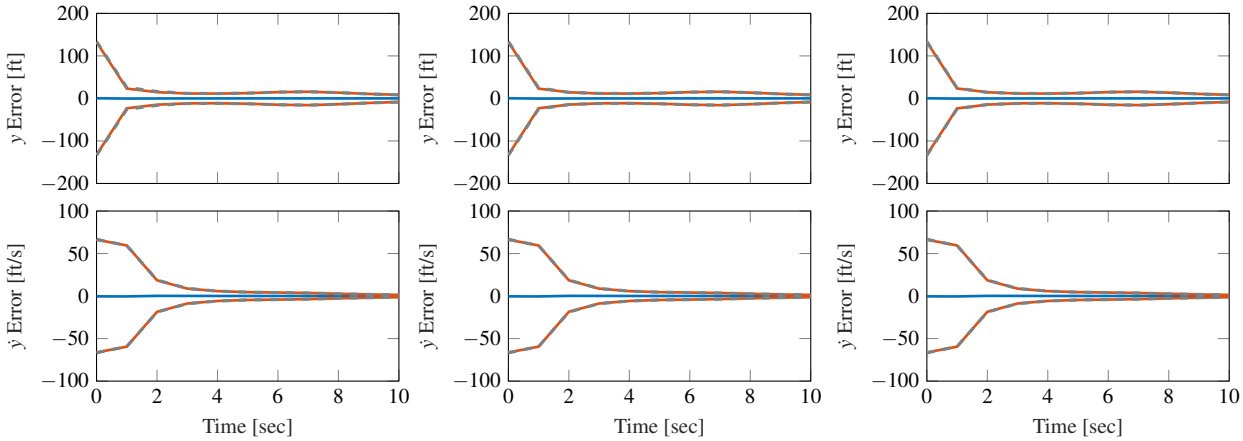
4.1.2 Analysis 1: Ideal Sensing Conditions

Before analyzing the filters under extreme stress, it is useful to first prove that they operate well in prototypical conditions. Therefore, this analysis is intended to mimic more ideal sensing conditions, where the probability of detection remains relatively constant and near 1 during the entirety of the simulation ($p_{D,\min} = 0.95$, $p_{D,\max} = 0.99$, $R_D = 0.0025$), and clutter is generated at a very low rate of $\lambda = 0.001$. The results of all three filters are presented in Fig. 1, where performance of each is nearly identical, and all are observed to be well-behaved as the average filter and MC standard deviations match almost exactly.

As none of the filters show signs of degeneracy, it can be said that any one of the filters is an appropriate selection in similar systems with nearly constant/unity probability of detection and essentially non-existent clutter. If minimal modeling and good computational efficiency is desired, the GMEKF with residual editing is advisable, as it performs as well as the other filters in this instance. This outcome is by design, of course, and this analysis provides a good baseline for the remaining analyses.

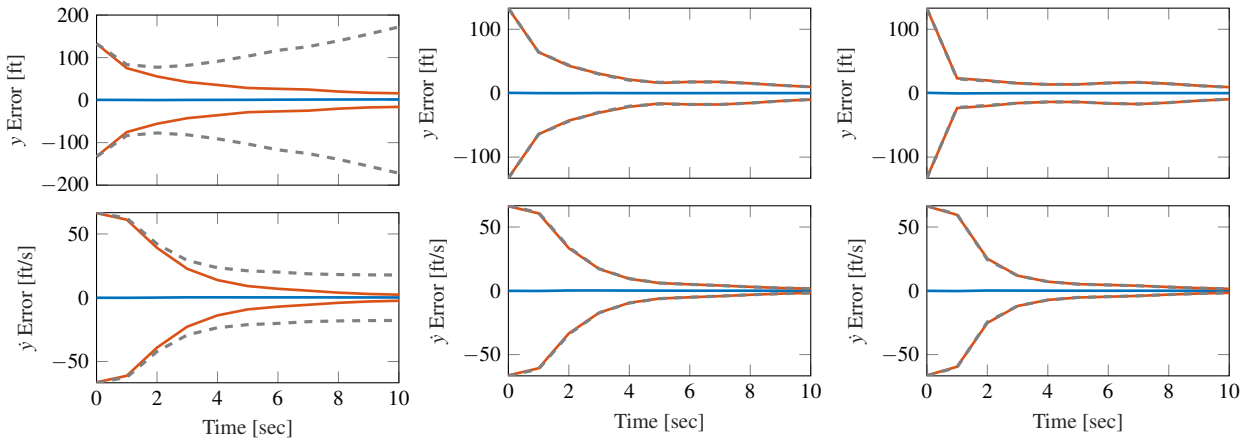
4.1.3 Analysis 2: Cluttered Environment with Highly Variable p_D

To evaluate filtering performance under less idealistic conditions, clutter returns are generated at a rate of $\lambda = 0.25$, and the probability of detection mean of Eq. (50) is given parameters of $p_{D,\min} = 0.10$, $p_{D,\max} = 0.99$, and $R_D = 0.0025$, such that $p_D(\boldsymbol{x})$ varies much throughout the simulation. The corresponding MC results are found in Fig. 2, where it is immediately clear that all filters exhibit higher levels of uncertainty than in Fig. 1, as the quality of incoming



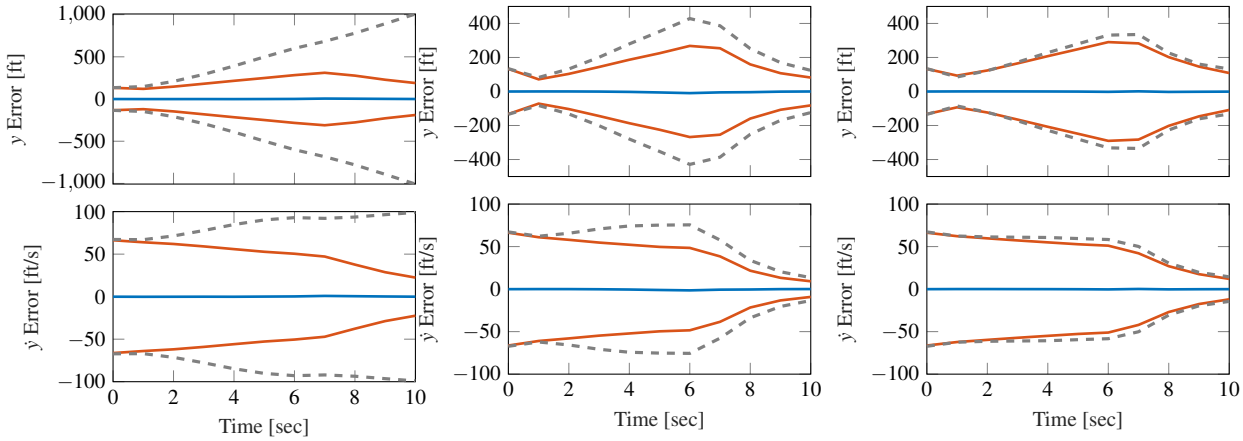
(a) **Residual Editing** (b) **Zeroth-Order Approximation** (c) **Gaussian Model**

Fig. 1: Filter results in ideal sensing environment plotted as $\overline{e}_\xi$ (——), $3\sigma_{\text{filt},\xi}$ (——) and $3\sigma_{\text{MC},\xi}$ (- - -).



(a) **Residual Editing** (b) **Zeroth-Order Approximation** (c) **Gaussian Model**

Fig. 2: Filter results in cluttered sensing environment plotted as $\overline{e}_\xi$ (——), $3\sigma_{\text{filt},\xi}$ (——) and $3\sigma_{\text{MC},\xi}$ (- - -).



(a) **Residual Editing** (b) **Zeroth-Order Approximation** (c) **Gaussian Model**

Fig. 3: Filter results for failing sensor plotted as $\overline{e}_\xi$ (——), $3\sigma_{\text{filt},\xi}$ (——) and $3\sigma_{\text{MC},\xi}$ (- - -).

measurements has significantly decreased. By the uncertainty profile of Fig. 2a, the GMEKF outfitted with residual editing is no longer capable of producing a valid filtering solution and shows highly divergent behavior. Meanwhile, the other filters show good consistency between MC and average filtering standard deviations, indicating healthy estimation solutions. However, the filter with a Gaussian model p_D of Fig. 2c produces estimates with lower uncertainty than that of the filter with zeroth-order approximation of Fig. 2b. This is attributed to two factors. First, as p_D varies more, there is a greater amount of information gain from the Gaussian model. Secondly, near the beginning of the simulation, the probability of detection is generally lower such that fewer valid measurements are being generated. Therefore, information gained from p_{D_z} returns is more influential at this time. As the $p_D(\boldsymbol{x})$ nears 1 towards the end of the simulation, valid measurements become the dominant contributor of incoming information, and both IFR filters achieve similar performance.

As expected, residual editing is not robust enough for cluttered sensing conditions with low detection probability, and the GMEKF of Section 3.5.2 is not capable of acceptable performance. Both IFR filters, on the other hand, are considered well-behaved, although modeling the p_D explicitly shows a significant advantage over the simple zero-order approximation here.

4.1.4 Analysis 3: Sensor Failure/Model Mismatch

As the proposed filters of this work are intended to be fault resistant, this analysis mimics a possible scenario with a failed sensor, where the filter is informed of an incorrect model for $p_D(\boldsymbol{x})$ in the time interval between two and six seconds. While the IFR filters expect a relatively high probability of detection with $p_{D,\text{min,filt}} = 0.95$, $p_{D,\text{max,filt}} = 0.99$, and $R_{D,\text{filt}} = 1$, starting at second two, the actual probability of detection is generated according to $p_{D,\text{min,true}} = 0.05$, $p_{D,\text{max,true}} = 0.15$, and $R_{D,\text{true}} = 0.0025$. Note that the covariance $R_{D,\text{filt}}$ is inflated here to reflect a lack of trust in the model, and it is ill-advised to ever set $R_{D,\text{filt}} < R_{D,\text{true}}$. To further stress the system, the rate of faulty measurements is increased tenfold to $\lambda = 2.5$. The results of this analysis are found in Fig. 3.

Upon inspection, it is clear that the GMEKF with residual editing of Fig. 3a still performs the worst by far, and that the IFR filters operate well, by comparison. However, the zeroth-order approximate IFR filter of Fig. 3b becomes significantly more overconfident than the Gaussian model IFR filter of Fig. 3c, while also producing higher overall errors. This implies that the IFR filter with a Gaussian model for $p_D(\boldsymbol{x})$ is the most robust to extreme model mismatch, but it is noted that it can also be the most sensitive if the filter covariance is not increased enough; i.e., if $R_{D,\text{filt}}$ is too small in situations with extreme model mismatch, then the filter of Fig. 3c performs much worse. It is found that in cases where the model of $p_D(\boldsymbol{x})$ is very poorly known and model mismatch of the probability of detection occurs, that setting $R_{D,\text{filt}} \geq 1$ is sufficient, as this not only expands the conventional 3σ confidence interval to include any return $0 \leq p_{D_z} \leq 1$, but also almost completely removes the information gain from the probability of detection.

4.2 Orbit Determination Simulation

This orbit determination (OD) simulation involves the tracking of the satellite O3B FM7 (Catalog ID 40080), maintained in Earth centered inertial (ECI) coordinates, and initialized on August 15, 2021 at 8:30AM UTC. To sufficiently address Eq. (23), satellite motion is governed by simple Keplerian two body dynamics, and no process noise is imparted to the system. Tracking is done by a ground-based observer located at the Haleakalā observatories on Maui, where measurement scans of right-ascension and declination are simulated every minute for 2.5 days, a duration that allows for multiple satellite passes and three "night" periods for the observer. It is assumed that the optical sensor has valid noise standard deviations of 3 arcseconds with a field-of-view spanning 3° by 3°, and clutter is uniformly generated across it at a rate of $\lambda = 50$.

Following procedures in [4], the admissible region of Fig. 4 is formed and subsequently approximated by a GM of 165 components, which provides a distribution to initialize the filter. To limit the increasing number of components, GM pruning and merging are applied [12], where $w_{\text{thresh}} = 10^{-10}$ and $d_{\text{merge}} = 0.05$. Additionally, splitting procedures are used to split each existing GM component into 18 smaller components (applied across 6 dimensions using a 3 component splitting library), which is performed anytime the number of components falls below 3, as a higher component number generally diminishes detrimental effects caused by nonlinearities [33].

Results are generated across 350 MC trials according to Eqs. (51) and (52), where both the zeroth-order approximation and Gaussian model versions of the IFR filter from Eqs. (44)–(47) are tested. In this simulation, the GMEKF with residual editing is found to fail so often that it is infeasible to generate any meaningful MC results, and is thus left

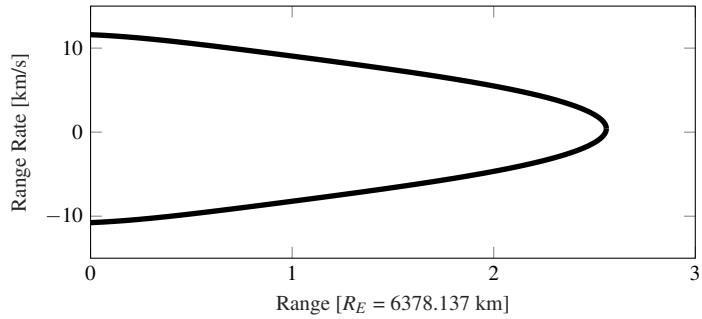


Fig. 4: Admissible region generated for O3B FM7.

out of this comparison entirely. Clearly, this indicates that a more nuanced approach to fault resistance than simply screening measurement residuals is needed for a simulation of this nature.

4.2.1 Probability of Satellite Detection

A four-part model of the probability of detection is built as

$$g(\boldsymbol{x}) = p_{D,1}\ p_{D,2}(\boldsymbol{x})\ p_{D,3}(\boldsymbol{x})\ p_{D,4}\ , \tag{53}$$

where each term $p_{D,i}$ corresponds to a different factor affecting detectability. While this model may not cover all aspects of satellite observability, it is intended to be more complete than a simple visibility check, and provide a higher level of complexity with which to test the filters. Again, the simulated probability of detection is randomly sampled according to the model of Section 3.4.3, where Eq. (53) is the state-dependent mean and $R_D = 0.03^2$ is the variance. An example of the simulated probability of detection over the duration of the simulation is presented in Fig. 5, the profile of which illustrates nine distinct periods of higher detectability.

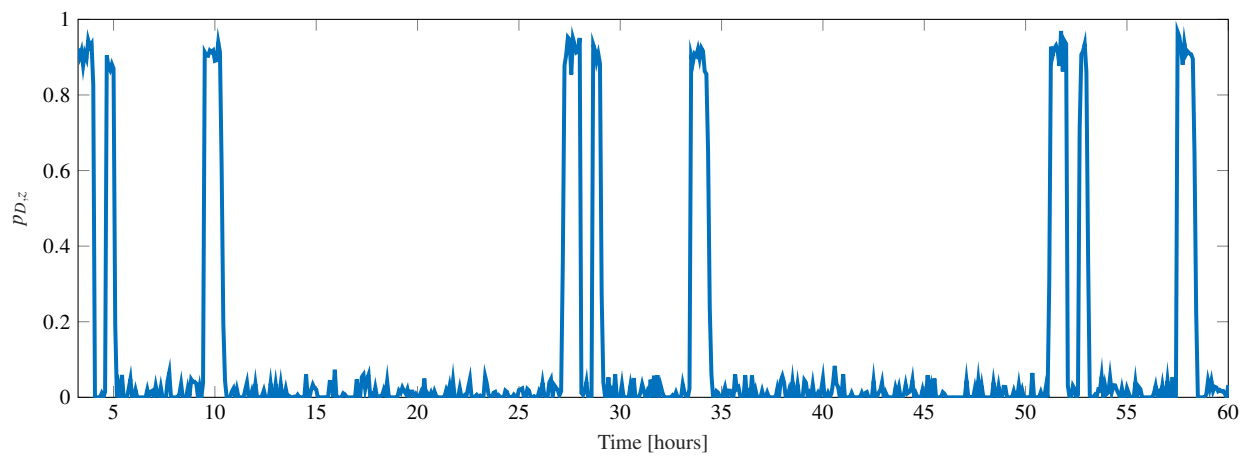


Fig. 5: Profile of randomly sampled probabilities of detection of orbiting satellite.

Observer Sky Brightness: It is assumed that, for valid sensor returns to be possible, the observer's sky must be sufficiently dark, a factor that is dictated by the elevation of the Sun α_{sun} with respect to observer horizon [35]. This model assumes that the sky begins to darken at nautical twilight ($\alpha_{\text{NT}} = -6°$), and reaches maximal darkness at astronomical twilight ($\alpha_{\text{AT}} = -18°$), the transition between which is modeled via a sigmoid function as

$$p_{D,1} = \frac{1}{1 + \exp\{c_1(c_2 + \alpha_{\text{sun}})\}}\ , \tag{54}$$

where choosing $c_1 = 2\frac{1}{\text{deg}}$ and $c_2 = 9^\circ$ in Eq. (54) yields a smooth function of $p_{D,1}$ that behaves as desired, which is plotted in Fig. 6a. Note that this specific model assumes that the observer cannot detect satellites under daylight conditions, which neglects the recent advancements in day-time observing [36].

Satellite Elevation: For valid satellite observations to be possible, the satellite must, of course, be above the observer horizon. This is simulated via $p_{D,2}(\boldsymbol{x})$, which is also modeled as a sigmoid function as

$$p_{D,2}(\boldsymbol{x}) = \frac{1}{1+\exp\{c_1(c_2-\alpha_{\text{sat}})\}} \ , \tag{55}$$

where α_{sat} is the satellite's elevation and, in this case, $c_1 = 5\frac{1}{\text{deg}}$ and $c_2 = -0.5^\circ$ produce simplistic behavior desirable for Eq. (55), the profile of which is shown in Fig. 6b.

Satellite Illumination: In order for the satellite to be visible to the observer, it must be sufficiently illuminated by the Sun. Put simply, a near-Earth satellite may be 1) not illuminated at all during an umbral eclipse, 2) partially illuminated during a penumbral eclipse, or 3) fully illuminated if no eclipse occurs [35]. Again, a sigmoid function is used to model a smooth transition between $p_{D,3} = 0$ and $p_{D,3} = 1$, such that

$$p_{D,3}(\boldsymbol{x}) = \frac{1}{1+\exp\{c_1(c_2-\theta)\}} \ , \tag{56}$$

where θ is the apparent angle between the centers of the Earth and Sun, calculated as

$$\begin{aligned}
\theta &= \cos^{-1}\left(\frac{\boldsymbol{r}_E \cdot \boldsymbol{r}_S}{|\boldsymbol{r}_E|\,|\boldsymbol{r}_S|}\right) \\
c_1 &= \frac{7.5}{\theta_1-\theta_0} \\
c_2 &= \theta_0 + \frac{\theta_1-\theta_0}{2} \\
\theta_0 &= |\theta_E - \theta_S| \\
\theta_1 &= \theta_E + \theta_S \ ,
\end{aligned}$$

and where θ_E and θ_S are the semi-diameters of the earth and sun, respectively, given by

$$\theta_E = \sin^{-1}\left(\frac{R_E}{|\boldsymbol{r}_E|}\right) \quad \text{and} \quad \theta_S = \sin^{-1}\left(\frac{R_S}{|\boldsymbol{r}_S|}\right).$$

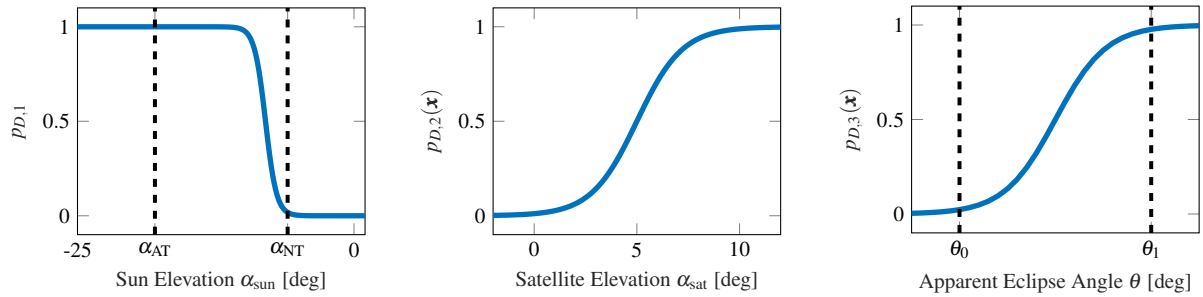
Note that $\boldsymbol{r}_S$ is the position vector from the satellite to Sun, and $\boldsymbol{r}_E$ is the position vector from the satellite to Earth, whereas R_E and R_S are the radii of Earth and Sun, respectively. More importantly, $\theta < \theta_0$ corresponds to an umbral eclipse, $\theta_0 \leq \theta \leq \theta_1$ corresponds to a penumbral eclipse, and $\theta > \theta_1$ corresponds to no eclipse at all. Fig. 6c contains a representative profile of the detectability function in Eq. (56).

Pointing Direction: Given the small size of the observer's FOV relative to the horizon, it is unrealistic to assume that the observer is able to always capture the satellite within the image frame. Instead, this model proposes that the observer is equipped with a fairly successful sensor tasking scheme, such that the satellite remains in view of the sensor approximately 91% of the time. If the vector to the satellite is considered the nominal pointing direction, then perturbations of right ascension δ_{ra} and declination δ_{d} are randomly sampled from a normal distribution as

$$\delta_{\text{ra}}, \delta_{\text{d}} \sim p_g\left(\delta \middle| 0, \left(\frac{3}{4}\right)^2 \text{deg}^2\right) \ ,$$

which are then used to offset the pointing direction from its nominal state. Thus, if $\max[|\delta_{\text{ra}}|, |\delta_{\text{d}}|] > 1.5^\circ$, then the satellite is outside the FOV, and a detection does not occur. Accordingly, $p_{D,4}$ is given by

$$\begin{aligned}
p_{D,4} &= \Pr\{|\delta_{\text{ra}}| \leq 1.5^\circ\}\Pr\{|\delta_{\text{d}}| \leq 1.5^\circ\} \\
&= 0.91107 \ .
\end{aligned} \tag{57}$$



(a) **Plot of $p_{D,1}$** (b) **Plot of $p_{D,2}$** (c) **Plot of $p_{D,3}$**

Fig. 6: Profiles of three different detectability factors contributing to the overall probability of detection.

4.2.2 Orbit Determination Simulation Discussion

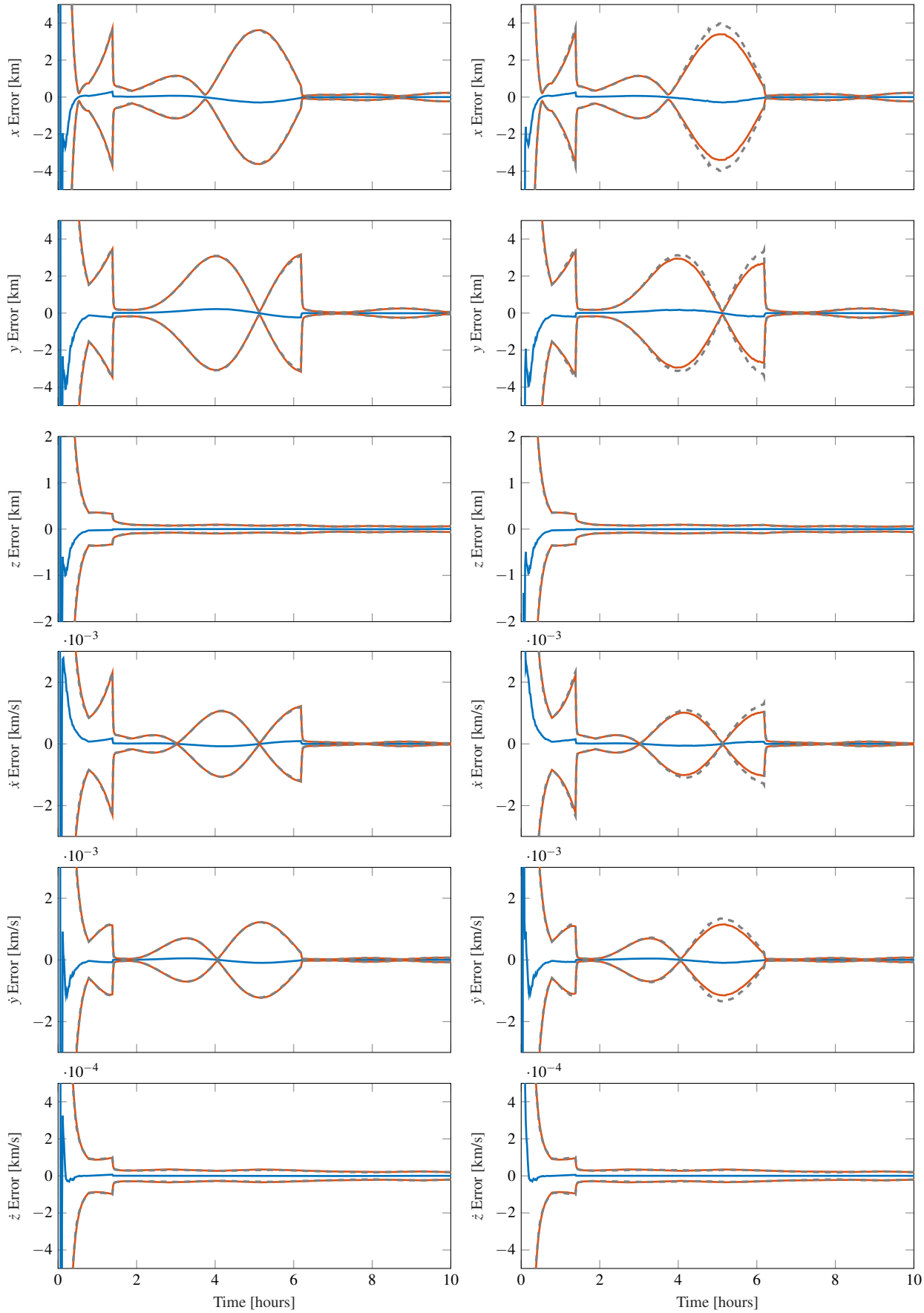
The MC results for both IFR filters are found in Figs. 7 and 8, which, at a glance, contain remarkably similar results. Neither of the filters fail during the simulation and, towards the end of the simulation, both filters acheive nearly identical levels of uncertainty. However, earlier in the simulation, especially between hours three and six, the zeroth-order approximated filter of Fig. 7b shows a slight overconfidence that the IFR filter with Gaussian model of Fig. 7a does not have. As this simulation is relatively complex, it is difficult to attribute this overconfidence to one specific factor, yet it can be stated that the proposed filter with Gaussian modeled $p_D(\boldsymbol{x})$ does exhibit better behavior than the zeroth-order approximation variant. Additionally, near the end of the simulation, the average filter uncertainty $3\sigma_{\text{filt},\xi}$ (——) of Fig. 8a is slightly lower than that of Fig. 8b. This indicates that the high fidelity IFR filter with Gaussian model produces a more certain estimate than that of the medium fidelity IFR filter with zeroth-order approximation, which is most likely due to the additional information gain from explicitly modeling the state-dependent probability of detection. Note that this difference in uncertainty is slight enough to the point that it is imperceptible in Fig. 8, which is simply because the optical measurements taken by the observer are much more accurate than the probability of detection returns supplied to the filter.

5. CONCLUSION

By way of a proposed intrinsic fault resistance (IFR) update, this work investigates probability of detection p_D modeling, specifically the selection of a level of fidelity. Three distinct models of detection probability are considered: 1) a low fidelity model where p_D is neglected completely, 2) a medium fidelity model where p_D is approximated as state-independent, and 3) a high fidelity model where p_D is taken as a Gaussian pdf. Results show that in ideal sensing conditions, any choice of fidelity produces well-behaved filters, whereas cluttered sensing environments with varying p_D require medium-to-high fidelity models to ensure robust operations. As such, fidelity selection is, as expected, highly dependent upon the system at hand and resources available. In this work, the medium fidelity p_D with zeroth-order approximation is robust enough to prevent filter failure in most analyses, but does not always achieve the accuracy or consistency of the high fidelity model. In general, the more variable the probability of detection is, the greater the benefits of the high fidelity model.

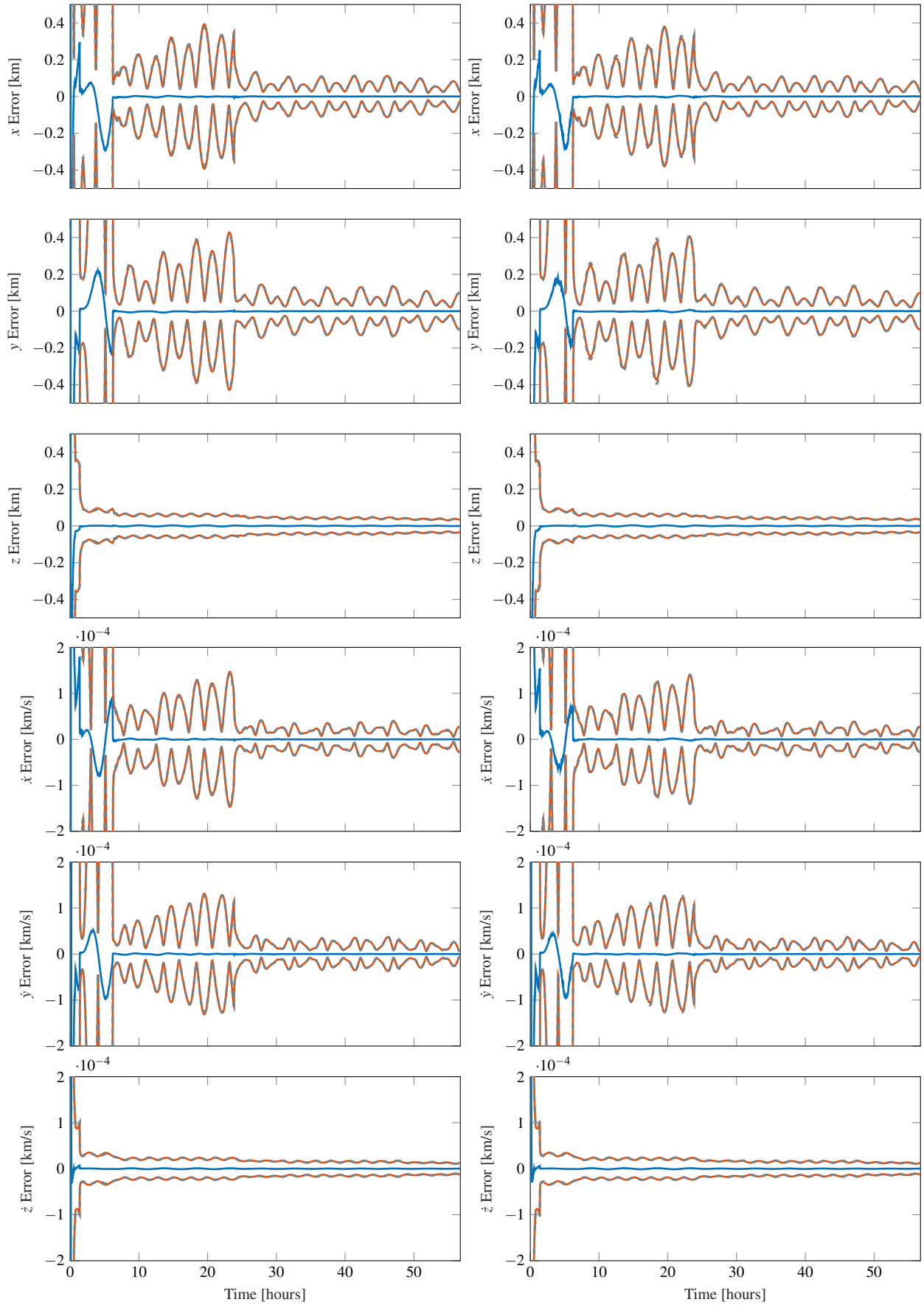
ACKNOWLEDGMENTS

This research is supported by the Department of Defense (DoD) through the National Defense Science & Engineering Graduate (NDSEG) Fellowship Program under fellowship number F-8445592924.



(a) **IFR filter with Gaussian model** (b) **IFR filter with zeroth-order approximation**

Fig. 7: Zoomed MC results of first 10 hours of OD simulation plotted as $\overline{e}_\xi$ (——), $3\sigma_{\text{filt},\xi}$ (——) and $3\sigma_{\text{MC},\xi}$ (- - -).



(a) **IFR filter with Gaussian model** (b) **IFR filter with zeroth-order approximation**

Fig. 8: Zoomed MC results of entire OD simulation plotted as $\overline{e}_\xi$ (——), $3\sigma_{\text{filt},\xi}$ (——) and $3\sigma_{\text{MC},\xi}$ (- - -).

REFERENCES

[1] DART Mishap Investigation Board. Overview of the DART mishap investigation results. Tech report, NASA. Available at http://www.nasa.gov/pdf/148072main_DART_mishap_overview.pdf, 2006.

[2] Cornelius J. Dennehy and J. Russell Carpenter. A summary of the rendezvous, proximity operations, docking, and undocking (RPODU) lessons learned from the defense advanced research project agency (DARPA) orbital express (OE) demonstration system mission. Technical Report TM–2011–217088, NASA Langley Research Center, April 2011.

[3] Ba-Ngu Vo, Mahendra Mallick, Yaakov Bar-Shalom, Stefano Coraluppi, Richard Osborne III, Ronald P. S. Mahler, and Ba-Tuong Vo. Multitarget tracking. *Wiley Encyclopedia of Electrical and Electronics Engineering*, pages 1–15, 2015.

[4] Kyle J. DeMars and Moriba K. Jah. Probabilistic initial orbit determination using Gaussian mixture models. *Journal of Guidance, Control, and Dynamics*, 36(5):1324–1335, September 2013.

[5] J. Russell Carpenter and Christopher N. D'Souza. Navigation filter best practices. Technical Report TP–2018–219822, NASA Goddard Space Flight Center, April 2018.

[6] Raphael Some, Richard Doyle, Larry Bergman, William Whitaker, Wesley Powell, Michael Johnson, Montgomery Goforth, and Michael Lowry. Human and robotic space mission use cases for high-performance spaceflight computing. *AIAA Infotech*, August 2013.

[7] Wesley Powell, editor. *High-Performance Spaceflight Computing (HPSC) Program Overview*, Bedford, MA, June 2018. Space Computing & Connected Enterprise Resiliency Conference, NASA Goddard Space Flight Center Electrical Engineering Division.

[8] Ronald P. S. Mahler. *Advances in Statistical Multisource-Multitarget Information Fusion*. Artech House, 2014.

[9] Ronald P. S. Mahler. Multitarget Bayes filtering via first-order multitarget moments. *IEEE Transactions on Aerospace and Electronic Systems*, 39(4):1152–1178, October 2003.

[10] Ba-Ngu Vo and W-K Ma. The Gaussian mixture probability hypothesis density filter. *IEEE Transactions on Signal Processing*, 54(11):4091–4104, November 2006.

[11] Gustaf Hendeby and Rickard Karlsson. Gaussian mixture PHD filtering with variable probability of detection. In *17th International Conference on Information Fusion*, pages 1–7. IEEE, 2014.

[12] Gunner S. Fritsch and Kyle J. DeMars. Nonlinear Gaussian mixture filtering with intrinsic fault resistance. *Journal of Guidance, Control, and Dynamics*, Accepted by JGCD.

[13] Taek Lyul Song, Darko Mušicki, and Kim Da Sol. Target tracking with target state dependent detection. *IEEE Transactions on Signal Processing*, 59(3):1063–1074, March 2011.

[14] Wolfgang Koch. On exploiting 'negative' sensor evidence for target tracking and sensor data fusion. In *10th International Conference on Information Fusion*, volume 8, pages 28–39. Elsevier, 2007.

[15] Rudolph Emil Kalman. A new approach to linear filtering and prediction problems. *Transactions of the ASME–Journal of Basic Engineering*, 82(Series D):35–45, March 1960.

[16] Simo Särkkä. *Bayesian Filtering and Smoothing*. Cambridge University Press, 2013.

[17] Ronald P. S. Mahler. *Statistical Multisource-Multitarget Information Fusion*. Artech House, 2007.

[18] Yaakov Bar-Shalom, X. Rong Li, and Thiagalingam Kirubarajan. *Estimation with Applications to Tracking and Navigation: Theory Algorithms and Software*. John Wiley & Sons, 2004.

[19] James S. McCabe. *Multitarget Tracking and Terrain-Aided Navigation Using Square-Root Consider Filters*. PhD thesis, Missouri University of Science and Technology, 2018.

[20] Wolfgang Koch. *Tracking and Sensor Data Fusion*. Springer, 2014.

[21] Arthur Gelb. *Applied Optimal Estimation*. MIT press, 1974.

[22] Branko Ristic. *Particle Filters for Random Set Models*. Springer, 2013.

[23] David M. Blei, Alp Kucukelbir, and Jon D. McAuliffe. Variational inference: A review for statisticians. *Journal of the American Statistical Association*, 112(518):859–877, 2017.

[24] Christopher M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.

[25] Stanley F. Schmidt. Applications of state space methods to navigation problems. *Advances in Control Systems*, 3:293–340, 1966.

[26] Bernard A. Kriegsman and Yee-Chee Tao. Shuttle navigation system for entry and landing mission phases. *Journal of Spacecraft and Rockets*, 12(4):213–219, April 1975.

[27] Renato Zanetti, Greg Holt, Robert Gay, Christopher D'Souza, Jastesh Sud, Harvey Mamich, Michael Begley,

Ellis King, and Fred D. Clark. Absolute navigation performance of the Orion exploration flight test 1. *Journal of Guidance, Control, and Dynamics*, 40(5):1106–1116, May 2017.

[28] James S. McCabe and Kyle J. DeMars. Terrain relative navigation with anonymous features. In *AIAA SciTech 2019 Forum*, San Diego, CA, January 2019.

[29] Harold W. Sorenson and Daniel L. Alspach. Recursive Bayesian estimation using Gaussian sums. *Automatica*, 7(4):465–479, July 1971.

[30] Daniel L. Alspach and Harold W. Sorenson. Nonlinear Bayesian estimation using Gaussian sum approximations. *IEEE Transactions on Automatic Control*, AC-17(4):439–448, August 1972.

[31] Sheldon M. Ross. *Introduction to Probability Models*, volume 11. Academic Press, 2014.

[32] Richard P. Wishner, John A. Tabaczynski, and Michael Athans. A comparison of three non-linear filters. *Automatica*, 5:487–496, 1969.

[33] Kyle J. DeMars, Robert H. Bishop, and Moriba K. Jah. Entropy-based approach for uncertainty propagation of nonlinear dynamical systems. *Journal of Guidance, Control, and Dynamics*, 36(4):1047–1057, July-August 2013.

[34] Harold W. Sorenson and J. E. Sacks. Recursive fading memory filtering. *Information Sciences*, 3(2):101–119, April 1971.

[35] T. S. Kelso. Visually observing earth satellites. *Satellite Times*, 3(1), September 1996.

[36] Jeffrey Shaddix, Jacob Brannum, Alex Ferris, Austin Hariri, Ari Larson, Tyler Mancini, and Jeff Aristoff. Daytime GEO tracking with "Aquila": Approach and results from a new ground-based SWIR small telescope system. In *Advanced Maui Optical and Space Surveillance Technologies Conference*, page 82, Wailea, HI, September 2019.

[37] Yu-Chi Ho and R. C. K. Lee. A Bayesian approach to problems in stochastic estimation and control. *IEEE Transactions on Automatic Control*, 9(4):333–339, 1964.

APPENDIX

Given $\boldsymbol{h}(\cdot)$, $\boldsymbol{R}$, $\boldsymbol{m}$, and $\boldsymbol{P}$ are of appropriate dimensions and $\boldsymbol{R}$ and $\boldsymbol{P}$ are symmetric, positive definite [37, 30]

$$p_g(\boldsymbol{z}|\boldsymbol{h}(\boldsymbol{x}),\boldsymbol{R})p_g(\boldsymbol{x}|\boldsymbol{m},\boldsymbol{P}) = p_g(\boldsymbol{z}|\boldsymbol{h}(\boldsymbol{m}),\boldsymbol{H}(\boldsymbol{m})\boldsymbol{P}\boldsymbol{H}^T(\boldsymbol{m})+\boldsymbol{R})p_g(\boldsymbol{x}|\boldsymbol{\mu},\boldsymbol{\Pi}) \ , \tag{58a}$$

where

$$\boldsymbol{\mu} = \boldsymbol{m}+\boldsymbol{K}[\boldsymbol{z}-\boldsymbol{h}(\boldsymbol{m})] \tag{58b}$$

$$\boldsymbol{\Pi} = \boldsymbol{P}-\boldsymbol{K}\boldsymbol{H}(\boldsymbol{m})\boldsymbol{P} \tag{58c}$$

$$\boldsymbol{K} = \boldsymbol{P}\boldsymbol{H}^T(\boldsymbol{m})[\boldsymbol{H}(\boldsymbol{m})\boldsymbol{P}\boldsymbol{H}^T(\boldsymbol{m})+\boldsymbol{R}]^{-1} \ . \tag{58d}$$

Note that $\boldsymbol{H}(\boldsymbol{m})$ is the Jacobian of $\boldsymbol{h}(\boldsymbol{x})$ evaluated at $\boldsymbol{x}=\boldsymbol{m}$.