Formalization of automation risks for dependability-based safeguarding of the nominal function

Arne Geffert
Institute for Traffic Safety and Automation Engineering
Technische Universität Braunschweig
Braunschweig, Germany
a.geffert@tu-braunschweig.de

Tianxiang Lan
Institute for Traffic Safety and Automation Engineering
Technische Universität Braunschweig
Braunschweig, Germany
t.lan@tu-braunschweig.de

Andreas Dodinoiu
Institute for Traffic Safety and Automation Engineering
Technische Universität Braunschweig
Braunschweig, Germany
a.dodinoiu@tu-braunschweig.de

Rasmus Rüdiger
Institute for Traffic Safety and Automation Engineering
Technische Universität Braunschweig
Braunschweig, Germany
r.ruediger@tu-braunschweig.de

Uwe Becker
Institute for Traffic Safety and Automation Engineering
Technische Universität Braunschweig
Braunschweig, Germany
u.becker@tu-braunschweig.de

Abstract*—To safeguard vehicles of higher automation levels, there is currently no generally accepted approach to handle the potentially deficient nominal function of sensor systems for automated driving. In this paper, corresponding risks depending on sensor quality are covered using the example of a GNSS-based multi-sensor localization system as an absolute position sensor for land vehicles. Therefore, basic measurement quality characteristics like integrity are transferred to dependability measures. Based on the PROFUND approach standardized in IEC 62551 with Petri nets, the dependability of that sensor system (including nominal function) is formalized and related to the Stanford diagram. After exemplifying important cases by means of the net model and considering safety-related design aspects, extended nominal functions of multiple channels are contemplated. For that purpose, the concept of risk genesis and various risk mitigation strategies are discussed. Moreover, a comparison between filtering and voting with Kalman filters is conducted based on numerical simulations. The approach in this paper connects GNSS, dependability (PROFUND), FuSa, SOTIF, data fusion, and fault tolerance for safe automation of road vehicles. Further Petri net models for multi-channel measurement systems and higher levels of the PROFUND net need to be developed to formalize the behavior of degradation strategies and to predict the associated risk of the nominal function by simulation.

Keywords—nominal function, dependability, RAMS, localization, GNSS, Stanford diagram, voter, Kalman, EKF, SOTIF, PEGASUS

I. Introduction

Within the automotive community, vehicles of higher automation levels are often called *open systems* [1]. But what does that mean? The term *open system* must not be confused with open systems known from thermodynamics. In fact, it is more reasonable to use the methodology common to control theory. Therefore, it is advisable to consider system boundaries (cf. Fig. 1) and the system states they comprise. For traditional (manually driven) road vehicle systems, one only needs to consider all the states in the vehicle subsystem (cf. left part of Fig. 1). Neither driver nor traffic states are essential for safeguarding the technical system, as it is the responsibility of the driver to estimate all the states necessary to accomplish the driving task. As soon as manufacturers start to automate vehicles so that engineered systems begin to replace the driver, the boundaries of the technical system broaden, now comprising the states of the (artificial) driver and of traffic (cf. right part of Fig. 1). For ensuring safety, an emergent system property, it is of utmost importance to estimate all relevant states, although the vehicles operate in an apparently uncontrolled (i.e., complex) environment. Currently, technological insufficiencies still limit the observability of some crucial states, thus rendering safeguarding a challenge.

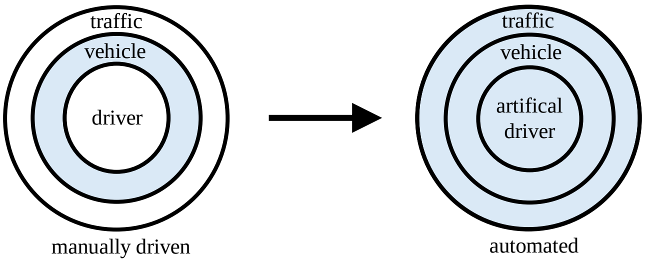

Fig. 1. System boundaries and states for manually driven (left) and for automated (right) road vehicles. Blue shading: system states under consideration.

Before concrete technical safety concepts can be derived, functional safety (FuSa) must be taken into consideration in order to minimize harm and to reduce liability risk when placing a new system on the market. Unlike for traditional automotive systems, not only malfunctioning behavior according to ISO 26262, but also the nominal function (also referred to as *intended* function) is safety-relevant for automated systems, leading to considerations made in SOTIF [2]. Currently, there is no generally accepted approach [3], especially for higher automation levels. According to PEGASUS, automation risks are subdivided into three classes, the first of which covers the interaction of environment and automation, containing sensor deficiencies that may propagate to higher system levels and so evoke unintended and risky behavior of the automation function [4]. In the paper at hand, automation risks mainly address the risks associated with insufficient sensor quality (i.e., the nominal function is deficient, but there is no fault in the sense of ISO 26262). As dependable localization is a basic prerequisite for automated driving, the following discussions refer to the example of a GNSS-based multi-sensor localization system for land vehicles (cf. Fig. 2). Although GNSS currently plays a minor role in safety-relevant car automation, it has a huge potential as soon as there is a solution for overcoming current insufficiencies. A major advantage of GNSS is that it will serve as second (i.e., redundant) absolute position sensor, complementing the digital map (transforming relative information to absolute), which might behave as a single point of failure if no additional absolute information is provided.

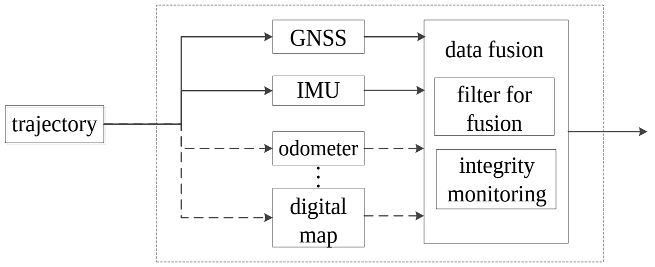

Fig. 2. Exemplary architecture of a GNSS-based multi-sensor localization system for land vehicles [cf. 5, 6]. GNSS: global navigation satellite system, IMU: inertial measurement unit.

Risk, mostly defined as the combination of the severity of an undesired event and its probability of occurrence, may arise due to either systematic or random errors (cf. Fig. 3). According to traditional FuSa theory, only random errors can be quantified according to reliability metrics. In contrast, reliability metrics are not able to describe systematic influences, which can only be treated with qualitative measures.

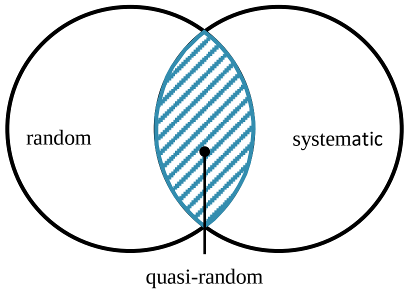

Fig. 3. Classification of errors. Blue shading: measurement errors of GNSS-based land vehicle localization system (nominal function).

At present, most measurement errors of a GNSS-based land vehicle localization system are introduced by complex environment conditions. On the one hand, a larger part of measurement deviations could be reproduced if the measurement was repeated under the same conditions (e.g., same satellite geometry, driving scenario, environment). So from a metrological view, these influences are systematic. On the other hand, it is nearly impractical to rebuild all those complex conditions. That is why the dependability-related behavior of the nominal function of the localization system can be allocated to the intersecting set of Fig. 3, namely in the following sense: although there are systematic influences (e.g., theoretically repeatable multipath due to urban canyons), the behavior of the localization system can be treated as quasi-random. That paradigm is especially valid because – in practice – driving trajectories are repeated in different ways. Because of that, nominal sensor quality can be quantified by reliability metrics and subsequently be used for dependability analysis. In the paper at hand, the term *dependability* is based on railway RAMS (reliability, availability, maintainability and safety) [7]. Dependability includes far more than safety aspects and reveals several trade-offs, e.g., between the orthogonal characteristics *safety* and *availability* [8].

II. Measurement quality and dependability

Impairments of the nominal function can lead to unintended behavior of the system in the SOTIF sense (e.g., if developers do not consider certain disturbances as GNSS multipath, the system may behave in an unsafe manner). Thus, it is necessary for a safe specification (i.e., a safe nominal function) to properly address measurement quality.

A. *GNSS measurement errors*

In the aviation domain, GNSS is in use for a long time. Therefore, the major challenges (concerning measurement quality) have been solved. Contrary to that, land vehicles operate in a different environment containing trees, mountains or high-rise buildings. The local environment still impairs measurement quality (cf. Fig. 4), since several physical effects like reflection, refraction, scattering and diffraction lead to multipath, non-line-of-sight (NLOS) reception or obscuration of signals (cf. Fig. 5). Those effects reduce accuracy and other measurement quality characteristics like integrity, continuity and availability [9]. In order to treat those effects properly, SOTIF must address them by demanding additional requirements augmenting the nominal function (e.g., integrity monitoring), which will subsequently be implemented with a certain ASIL according to ISO 26262 if they are considered safety-relevant.

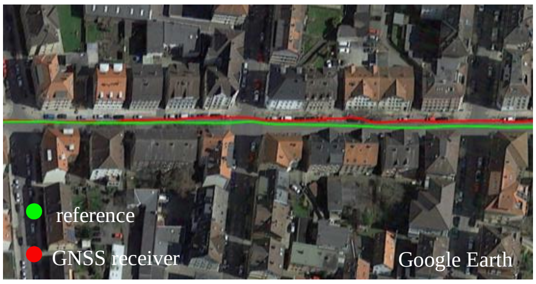

Fig. 4. Low GNSS measurement quality due to local effects [cf. 5, 6].

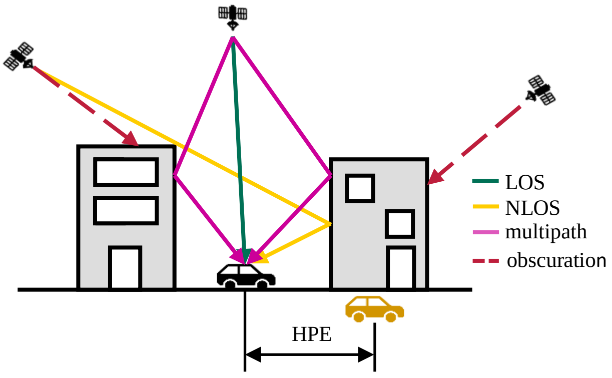

Fig. 5. Typical measurement environment with impairing effects (LOS: line of sight, NLOS: non line of sight, HPE: horizontal position error) [cf. 5, 6].

B. *Measurement quality characteristics*

The four characteristics accuracy, integrity, continuity and availability as shown in Fig. 6 (top left) have a major impact on measurement quality. Special cases neglected, availability is simplified as the function of the other three characteristics and describes the average likelihood that a measurement system is able to provide an output value. In contrast, continuity describes the probability that the system will incessantly maintain its ability to provide an output value for a defined period. Accuracy, as a direct basis for continuity and integrity and so an indirect basis for availability, is the combination of trueness and precision of a measured position [10, 11]. Integrity provides a measure of trust and can limit availability if the integrity function detects unacceptably large position errors (PE). This is done by comparing the protection level (PL) with an alarm limit (AL), cf. Fig. 6 (top right and bottom). The protection level is formed by the estimated variance around the measured position and shall include the ground truth (GT) value with high probability. If the protection level is smaller than the alarm limit, the measured position is considered *trustworthy* and can be used. If the protection level is greater than the alarm limit, the measured position is considered *not trustworthy* (i.e., the system loses its ability to provide an output value). If the GT value lies without the PL, but the integrity function does not detect that outlier (PL < PE), the system maintains its ability to provide an output value, but its state becomes hazardous, since the integrity function erroneously misses to classify the measurement value (MV) as *not trustworthy*. The maximum allowable size of the alarm limit depends on the application and has an impact on the aforementioned trade-off between safety and availability.

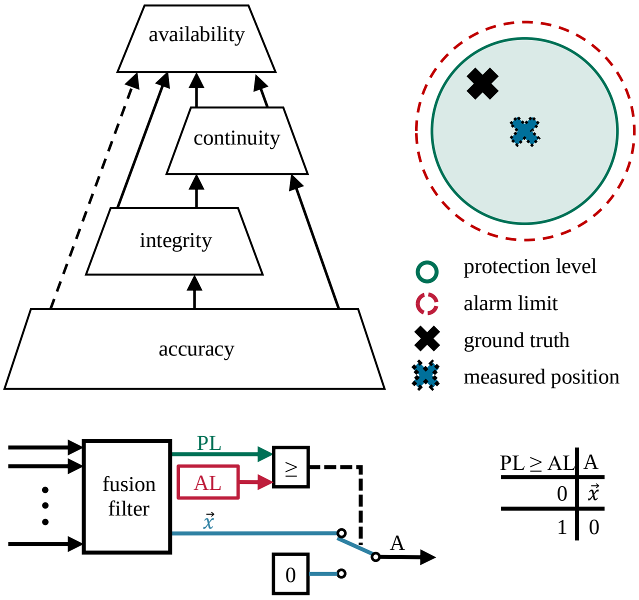

Fig. 6. Measurement quality characteristics (top left), functional principle of the integrity function (top right, bottom left) and associated logic table (bottom right) [cf. 5, 6, 9].

C. *Transferring measurement quality to dependability*

Contrary to measurement quality, dependability is a system characteristic. If the integrity monitoring classifies a measurement value as *not trustworthy* (cf. first falling edge in Fig. 7, top), the system state changes from *operational* to *not operational* (cf. first falling edge in Fig. 7, bottom) if no redundancies or additional channels are provided. That change has the same effect on system dependability like a classic component failure, e.g., a faulty component changing its state from *intact* to *defective*. In both cases, the state changes lead to a system failure. Thus, they have a similar impact on system dependability. However, a rising edge of the measurement state does not necessarily evoke a state change of the system, as can be seen in Fig. 7: the first rising edge in the upper picture does not evoke a rising edge in the lower picture. The reason is that there are other influences (e.g., classic component failure) that prohibit the state change in the lower picture.

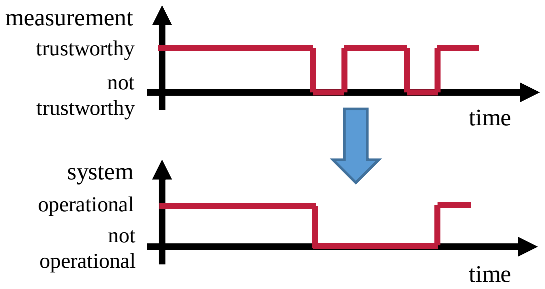

Fig. 7. Temporal behavior of measurement and system state.

The relation between measurement quality and dependability characteristics is depicted in Fig. 8. The UML diagram shows that GNSS measurement quality characteristics (left) have an influence on the dependability characteristics (right) of a system: Measurement integrity influences the safety of a system; the availability of the measurement value has an influence on the system availability etc. Accuracy only has indirect influences on the other characteristics.

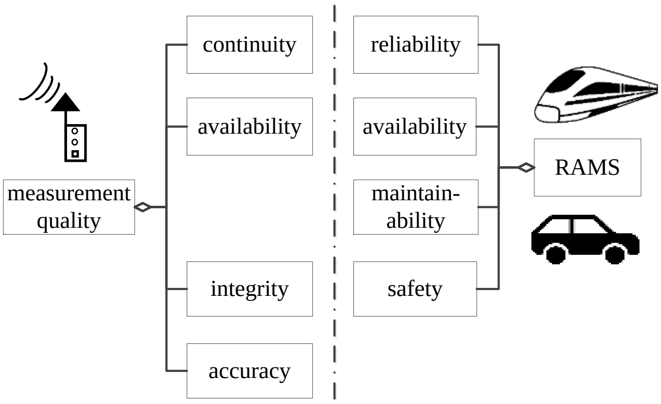

Fig. 8. Relation between measurement quality and dependability (UML) [cf. 5, 12].

III. Formalizing Dependability

For formalizing both functionality and dependability, the PROFUND approach according to Slovák [13] is refined with regard to the *intended* behavior (nominal function). Therefore, Petri nets are used as a means of description.

A. *Petri nets*

Petri nets are a formal means of description for state-based modelling with local states and are standardized, e.g., in IEC 62551 [14]. A Petri net consists of four major elements: place, transition, arc, token and only in some cases of further special elements shown in TABLE I.

A Petri net is a directed graph of places (as passive nodes) and transitions (as active nodes). A place refers to a local state or to an object and is drawn as a circle. If a state becomes active, it will be marked with a token in the shape of a black dot in the circle. Places can only be interconnected via transitions. The link between a place and a transition – or vice versa – is realized by a directed arc showing the path of a token between these two nodes. A transition will be enabled if each place connected by an arc directed to the transition is marked with a token. All these places are called pre-places of the transition (e.g., in Fig. 9 the place *operational* regarding the transition *failure*). An enabled transition fires so that the tokens of all pre-places are absorbed and on each post-place of the transition a new token is created. A post-place is connected with a transition by an arc directed to that place (e.g., in Fig. 9 the place *faulty* regarding the transition *failure*).

There are two special types of arc, test arcs and inhibitor arcs. Such an arc is considered for enabling a transition, but the firing of the transition will not influence the marking of the connected pre-place. The difference between a test arc and an inhibitor arc is that a pre-place linked via a test arc has to be marked for enabling a transition, whereas a pre-place connected via an inhibitor arc does *not* have to be marked. In a place-transition net, an enabled transition can fire immediately and is only controlled by the explained relations. For this paper, it is additionally necessary to model the behavior of a system due to stochastic failures. Therefore, all Petri nets in this paper are Extended Generalized Stochastic Petri nets (EGSPNs) including furthermore stochastic transitions. A stochastic transition is drawn by a shaded rectangle. The enabling and firing follows the aforementioned rules, but additionally the transition will only fire with a certain probability. In consequence, the firing will take place depending on the modelled stochastic distribution [15].

TABLE I. ELEMENTS OF A PETRI NET [CF. 14, 15]

⦿	▬	▨
place marked with token inside	transition	stochastic transition
⟶	- - ➝	- - ○
arc	test arc	inhibitor arc

Petri nets can grow in complexity quite fast and become hard to understand. To reduce complexity and to enhance transparency, a so-called fusion place can be drawn several times in a Petri net, representing a copy of the original place (e.g., the *operational* place in Fig. 11 is a copy of the original *operational* place in Fig. 12). Furthermore, a part of a Petri net can be aggregated into a subnet. A subnet is represented in a Petri net by an aggregated element such as a super place, a super transition or a super arc as shown in TABLE II. Those elements follow the explained modelling rules of EGSPNs.

TABLE II. AGGREGATED ELEMENTS FOR HIERARCHICAL PETRI NET MODELLING [CF. 14]

◎	▭	⟹	= = ➤
super place	super transition	super arc	super test arc

The approach of using Petri nets to model the system behavior is superior to other modelling methods due to a detailed look at the active states at every simulation step. That often leads to a less complex model and allows to analyze the local states of a system in contrast to the global states of a Markov chain. If the global states are still needed for the analysis of a system, the Petri net can be transformed into a so-called reachability graph. Under certain circumstances, that graph can be interpreted as a Markov chain.

B. PROFUND

Several years ago, Slovák developed the PROFUND approach for modelling the dependability of technical systems [13]. Having led to the standard ISO IEC 62551 [14], PROFUND is a hierarchical approach that strictly separates between the technical **pro**cess, **fun**ctionality and their **d**ependability. As PROFUND originally focusses on railway systems, it must be adapted to automotive systems by explicitly modelling the behavior of the nominal function.

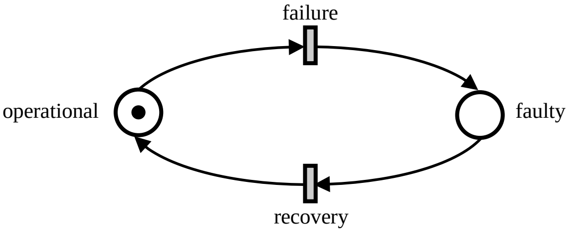

Fig. 9. Failure model: a failure transfers the operational into the faulty state.

In this paper, only the functionality and the dependability of the measurement system are modelled. For a complete PROFUND analysis, at least the control function (comprising the measurement system), the technical process (e.g., automotive traffic) and the dependability of both would have to be modelled in addition.

Failure and *fault*, being crucial terms of dependability analysis, are understood in the way depicted in Fig. 9. Whereas *fault* is a state, *failure* is a dynamic process that transfers the *operational* into the *faulty* state. Thus, a failure can happen, but it cannot persist. Only the fault (or the faulty state) can be persistent.

With the help of aggregated Petri nets, the hierarchy of the regarded system is shown in the summary net depicted in Fig. 10. On the left-hand side, the data flow (transforming input to output data) is modelled. The functionality net decides what kind of output datum is created. Additionally, the output classifier net chooses between a safe and a (potentially) hazardous output datum. To be able to do so, the two latter nets are influenced by two dependability nets. The first one models the dependability of classic resources (e.g., defective devices), whereas the second models the dependability of the nominal function. In this case, nominal function is synonymous with measurement quality.

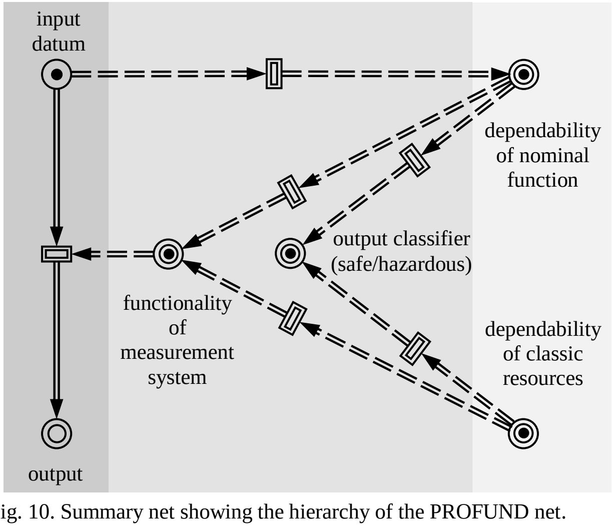
input datum
dependability of nominal function
output classifier (safe/hazardous)
functionality of measurement system
dependability of classic resources
output

Fig. 10. Summary net showing the hierarchy of the PROFUND net.

In the following, the nets are presented in detail. It should be noted that the model presented is only an example showing the method to create probabilistic dependability nets for the nominal function. Certain assumptions are arguable and can be chosen depending on the application. Apart from that, the colors used in the following Petri nets serve as a means of recognition since they indicate the relation of elements to a certain net (e.g., a red fusion place is related to the red quality of integrity net). The use of colors shall not lead to the association with Colored Petri nets, which are another class of Petri net.

C. *Data flow net*

The data flow net (Fig. 11) models the fact that input data must be transformed to output data if the measurement system is operational. If not, the output place *no output datum* will be marked. For the sake of clarity, it must be remarked that the hierarchical structure of the nets is maintained via fusion places (as pre-places of the respective transition). For example, the *input datum* will only be transformed to *output datum* if the fusion place *operational* is marked. But it is decided in another net if that fusion place will be marked (see below).

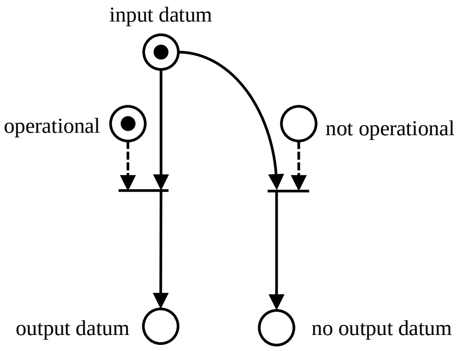

Fig. 11. Data flow net. Fusion places refer to Fig. 12.

D. *Functionality net*

The measurement system fails if its device-related resource is in fail-safe mode or if the integrity function marks the measurement value as *not trustworthy* (PL ≥ AL), see Fig. 12. Functionality is only restored if the measurement value becomes trustworthy and the device-related resource has left the fail-safe mode. It is noteworthy that in line with this modelling, the measurement system is regarded operational even if a device-related resource fails hazardously (but stays functional) or if the integrity function erroneously misses the detection of an outlier (PL < AL, but measurement error is out of tolerance). Therefore, the output classifier (safe/hazardous) is added. It is also remarkable that a great many of dependability engineers would name the *operational* place with the common term *available*. According to the understanding of this paper, availability has a slightly different meaning, since availability is defined as the mean probability of being in the *operational* state.

E. *Output classifier (safe/hazardous)*

The output classifier (Fig. 13) defines whether the output datum is hazardous or (ostensibly) safe. It is (ostensibly) safe, e.g., if the device-related resource is in fail-safe mode or if the condition PL ≥ AL holds (in both cases, the system is both *safe* and *not operational*). The adverb 'ostensibly' is used because the combination of several states, e.g., *not operational* and *safe* can lead to additional risk on higher (not modelled) levels of the PROFUND net (e.g., a not operational measurement system could require a degradation strategy of an automated vehicle, leading to higher risk compared to nominal operation). Furthermore, there is no connection (e.g., test arc) between the output classifier and the data flow, since the risk that arises due to a 'hazardous output datum' must be modelled on a higher level of the PROFUND net (out of scope of this paper).

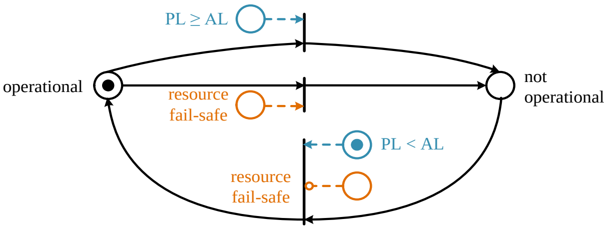

Fig. 12. Functionality net (functionality of measurement system).

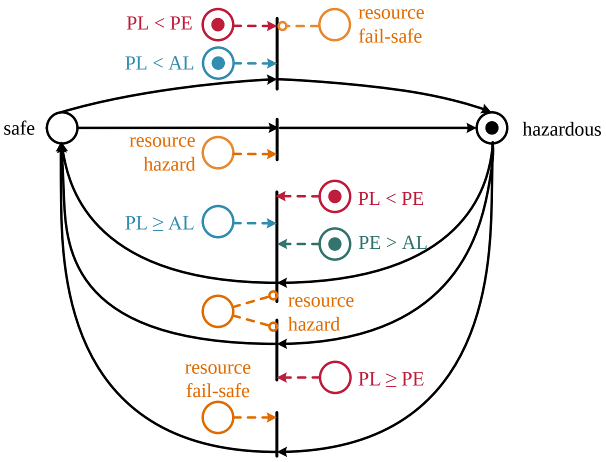

Fig. 13. Output classifier net. Fusion places refer to Fig. 14 and Fig. 15.

F. *Classic dependability net*

A level deeper, there is the dependability net of classic resources (e.g., defective devices), containing the places *intact*, *fail-safe* and *hazard* (see Fig. 14). Based on Slovák [13], the transitions are modelled as super transitions with the possibility to be detailed. As device-related dependability is not the focus of this paper, the net is not furtherly refined.

G. *Dependability net of nominal function*

For the sake of simplicity, only the measurement quality aspect of the nominal function is modelled (see Fig. 15). The net runs from left to right (in the order green, blue, red) and is interconnected with the other nets via fusion places. To keep the net simple, the order is modelled implicitly by priorities if no explicit control structure (e.g., fusion place with test arc) is provided.

When the place *input datum* in the data flow net (Fig. 11) is marked, the respective fusion place in the green localization function measurement error net of Fig. 15 is also marked. If measurement conditions are disadvantageous (e.g., multipath), the fusion place named '-' will be marked. As all pre-places are marked, the stochastic transition can fire with a certain probability and the measurement error (ME) place will change from *low* to *high*. The probability is defined by a distribution belonging to the transition. It must be modelled so that the error behavior of the receiver is correctly approximated.

Whereas the green net only models the behavior of the pure measurement process, the integrity monitoring function (e.g., receiver autonomous integrity monitoring (RAIM)), commonly used for ensuring safety in the GNSS domain, is modelled in addition so that the whole net in Fig. 15 characterizes the behavior of a *monitored localization function*. The blue net, called *safeguarding function*, is the first part of the integrity monitoring and consists of observing the measurement error and raising an alarm. The upper place *trust in MV* (MV: measurement value) is equivalent to the *trustworthy* state in Fig. 7, whereas the lower place *alarm* leads to the *not trustworthy* state in Fig. 7. As the measurement error (green net) is normally not observable to the user, statistical procedures are used to estimate the PL instead. If PL ≥ AL, the probability of the true error being out of tolerance is high, so that an alarm is raised not to use the current measurement value. In more detail, that modelled behavior is only *a part* of a real safeguarding function since a system reaction to that alarm would have to be complemented on a higher hierarchical level (e.g., degradation strategy).

The red net, called *quality of integrity*, is needed to model the cases in which the integrity routine is wrong, e.g., missed detection (false negative). Most of the control structure (esp. test and inhibitor arcs) is needed for taking those cases into account. The '& place' together with the adjacent super transition and the inhibitor arc represent a logic NAND constraint, e.g., if PE > AL and PL < AL, the red net must stay in (or switch into) the faulty state.

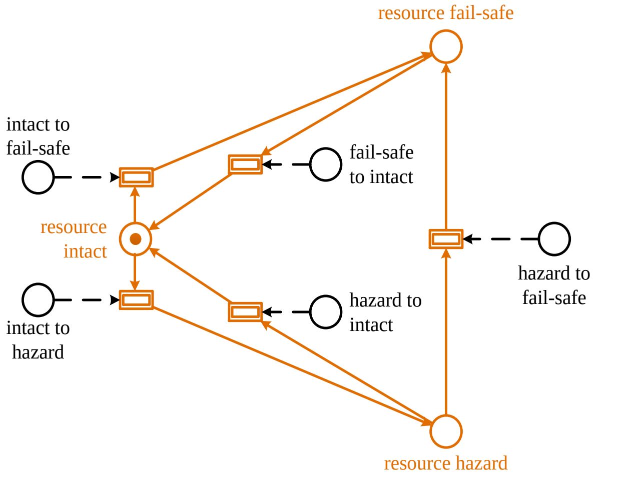

Fig. 14. Dependability of classic resources (transition from fail-safe to hazard intentionally neglected) [cf. 13].

H. Synthesis

All in all, the net runs as follows: When the receiver takes a new measurement, the *input datum* place (Fig. 15) is marked, enabling the green net (Fig. 15) to work. After that, the blue net simulates the observation process possibly followed by an alarm. Then the NAND constraint nets are executed, enabling the red net to decide whether the decision of the blue net was true or false. Thereafter, the classic dependability net (Fig. 14) runs. Subsequently, state changes take place in the functionality and output classifier net (Fig. 12 and Fig. 13) according to the changes made in the dependability nets. Finally, one of the two transitions in the data flow net (Fig. 11) fires in order to transform the input datum to an output state.

IV. Exemplifying Automation Risk

Initially, the Petri nets developed in chapter III can be used to illustrate the risks associated with a monitored localization function. In addition, they can be parameterized in order to simulate the behavior of the localization system (out of scope of this paper). Such a simulation can be useful in an early state of the development process, e.g., for supporting architectural decisions.

A. Relation of Petri net and Stanford diagram

In the GNSS domain, the Stanford diagram (Fig. 16) is a well-known illustration of the quality of integrity, thus being an indicator of the system state. Whereas the ordinate displays the size of PL in relation to AL, the abscissa reveals the true position error, which is unobservable in most applications. The dependability net of Fig. 15 is designed in such a way that each measurement value is assigned to one area (i.e., system state) of the Stanford diagram [5].

B. Missed detection rate

Missed detection is the case that corresponds to the marking depicted in Fig. 15: the measurement error is high (PE > AL), but the safeguarding function erroneously raises no alarm (PL < AL), which automatically enforces the red net to be in the *faulty* state (PL < PE). In practice, that global state can occur if the measurement environment contains at least one condition that is not included in the specification (i.e., out of scope of the nominal function), e.g., complex multipath evoked by trees or high-rise buildings. In unfavorable cases, the integrity monitoring understimates the measurement error, although it is hazardously high (e.g., automated vehicle assumes to drive on the right lane, but actually is on the left with no alarm being raised). Depending on the domain, the *missed detection* is also named false negative, dangerous undetected [12], hazardously misleading information (HMI, cf. Stanford diagram in Fig. 16) or hazard. It is noteworthy that the measurement system is regarded operational, although it has failed hazardously.

C. False alarm rate

Another important case is the false alarm, which occurs if the measurement error is low (PE ≤ AL), but the safegurading function erroneously raises an alarm (PL ≥ AL), which automatically enforces the red net in Fig. 15 to be in the *intact* state (PL ≥ PE). In practice, that state can occur if the integrity function is too conservative, thus overestimating measurement errors. Depending on the domain, the *false alarm* is also named false positive or safe detected [12]. It is noteworthy that the measurement system is regarded not operational, although it is ostensibly safe. Additional risk could result from the traffic process, e.g., if the vehicle performs a degradation strategy.

D. Safety-related probabilistic design

According to Junietz et al. [17], safety requirements may not be directly derived from tolerable accident rates. But SOTIF provides neither quantitative metrics for failure rates of the nominal function nor acceptance criteria (i.e., tolerable failure rates). If a design was based on tolerable accident rates, several practical issues would arise, e.g., that the rates are dependent on market penetration [17].

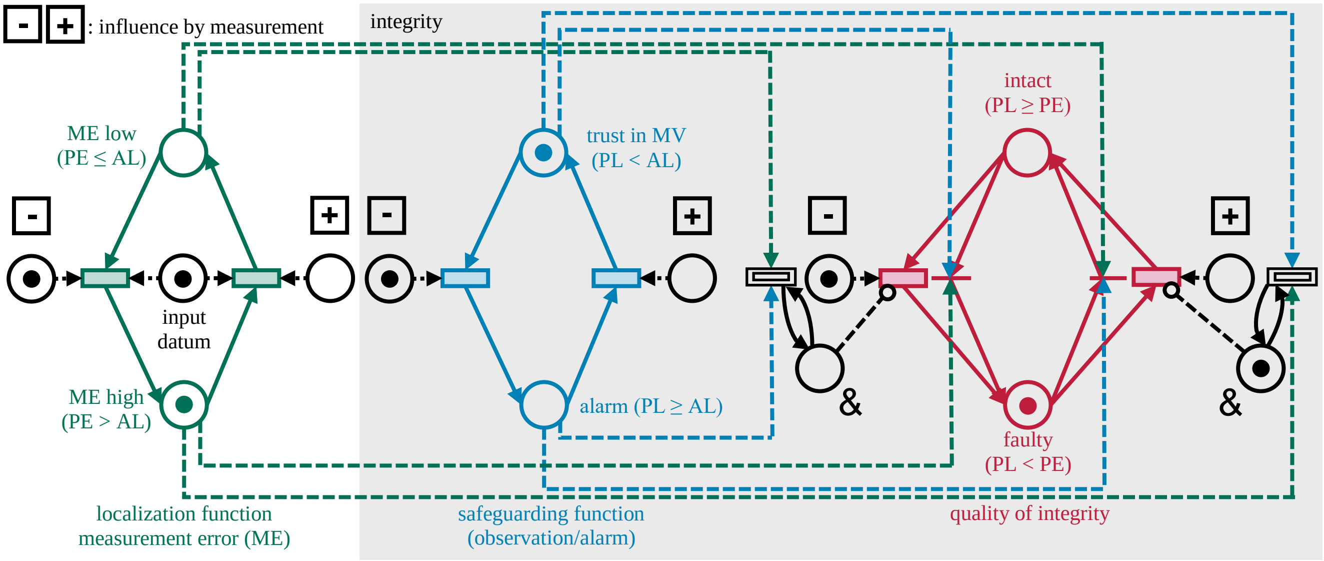

Fig. 15. Dependability of nominal function: error behavior (green) and integrity monitoring (blue and red). +/-: advantageous/disadvantageous influence by measurement conditions; ME: measurement error; MV: measurement value [cf. 5].

According to the authors' view, a probabilistic design including mortality rates (as discussed by Stiller [3] and Junietz et al. [17]) is necessary, but not sufficient. The rates will be used for ensuring dependability, which goes far beyond safety. Such a dependability-related design will ensure that undesired events like missed detection or false alarm will happen with an acceptably low frequency, thus satisfying the German Ethic Commission's claim for a positive balance of risks [18] as well as requirements for user acceptance and trust. Therefore, rough estimations for the design of the nominal sensor function have to be made in order to overcome practical problems like the dependence on market penetration. Those estimations deliver an initial balance of risks built on simulation-based evidence, which must be replaced by empirical evidence through lifecycle-based field observation [19]. If empirical evidence shows that the initial risk balance is not met, manufacturers must change their design, even during the lifecycle. Apart from that, manufacturers must additionally comply with SOTIF criteria (transcending the consideration of mortality rates) for the purposes of admission and justiciability.

V. Nominal Function Dependability

The Petri net models described in chapter III represent the behavior of a single-channel model. Depending on both system architecture and safeguarding strategy, the nominal function possibly consists of multiple channels, i.e., several of the abovementioned Petri net models in parallel. The following subchapters discuss the possibilities for enhancing nominal function dependability in connection with the Petri net models.

A. *Risk genesis*

According to Schnieder's model of risk genesis (Fig. 17), damage can only occur if exposition and hazard coincide. In other words, damage can only occur if there is at least one object that is exposed to a hazard. In addition, hazard is not yet damage. As the other chapters of this paper concentrate on modelling the probability of occurrence of a hazard, one has to keep in mind that calculated missed detection rates (i.e., hazard rates) are not equal to damage rates or mortality rates. But according to the authors' view, it is advisable to treat them equally as a conservative approach, as long as there is too little experience with localization systems for safety-relevant road traffic applications. Another approach for a simulation-based prediction considering both exposition and hazard comprises the modelling of higher levels of the PROFUND net (i.e., the traffic process and its dependability), which is out of scope of this paper.

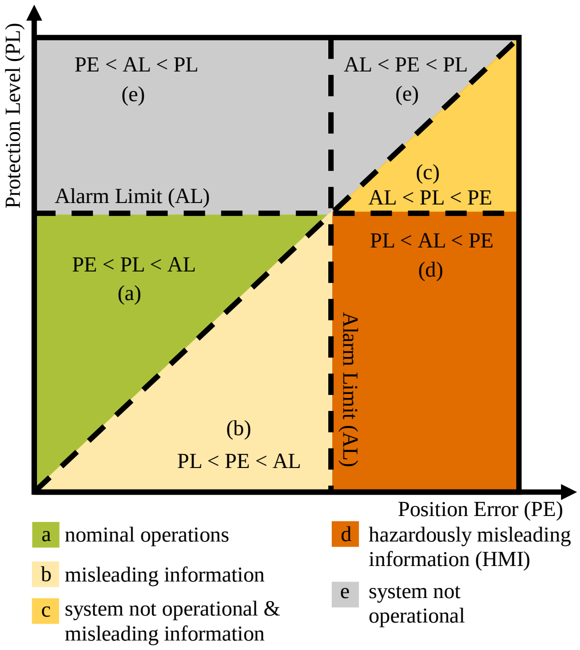

Fig. 16. Stanford diagram [cf. 16], adapted to the terminology of this paper (system-oriented terminology, classic dependability neglected). Relation with PROFUND nets: b and d are hazardous (Fig. 13), c and e are not operational (Fig. 12).

B. *Risk mitigation strategies*

On the technical level, a wide range of strategies is known in the GNSS domain. In most cases, there is a well-known trade-off between safety and availability. To find a risk mitigation strategy that goes beyond the trade-off, several possibilities exist: better sensors (e.g., GNSS receiver, antenna

[21]), better correction algorithms (e.g., ray tracing, probabilistic modelling [22]) and better data fusion (e.g., map matching [23], complex filters [24], Monte Carlo methods [25]) are often discussed. In addition, integrity monitoring (esp. RAIM) has its origin in the aviation domain and thus must be adapted to the needs of ground traffic, since it often fails to correctly handle effects coming from the local environment [9] (cf. subchapter II.A). In theory, integrity is a safety concept consisting of several layers (cf. Fig. 18). At best, the innermost layer, the integrity of the GNSS itself, has a defined integrity risk (e.g., when using Galileo Public Regulated Service). The next layer is SBAS (satellite-based augmentation system, e.g., EGNOS in Europe), trying to reduce the integrity risk further. As it is designed for aviation, it lacks to comply with ground traffic requirements, but still enhances localization quality (cf. experiences made in railway research [26, 27]). The outermost safety layer – (R)AIM – is an extra risk reduction factor and roughly corresponds to the integrity monitoring depicted in Fig. 15.

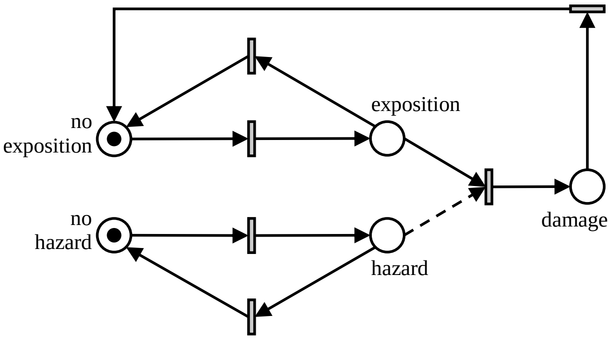

Fig. 17. Generic Petri net model of risk genesis [20].

On the functional level (cf. Fig. 17), the integrity function is tantamount to a monitoring of the measurement value, based on statistical means. If the monitoring suspects an incorrect measurement value, an alarm is instantaneously raised – according to common GNSS terminology. But in an automated vehicle system, the alarm has to trigger a safety-related reaction on higher levels of the PROFUND net, e.g., an emergency braking. That is why the combination of observation, alarm and automated safety-related reaction must be considered as safeguarding. Referring to Fig. 17, the safeguarding process shall ensure that the system stays in the *no hazard* state. A missed detection could violate that mechanism.

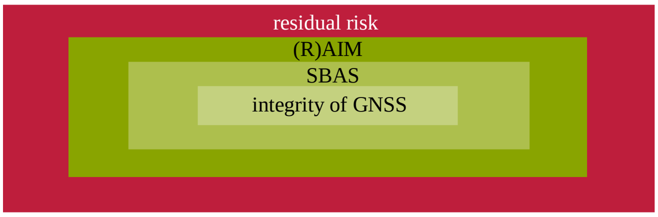

Fig. 18. Integrity layers of GNSS-based localization [cf. 5].

Types of safety-related strategies are various. Most commonly, fail-safe or fail-silent strategies suggest themselves when GNSS integrity monitoring is regarded. The reason is that for a measurement system (as shown in Fig. 10), monitoring and raising an alarm (not to use the measurement value) are equivalent to a fail-safe/silent strategy. But for an automated vehicle, the fail-safe mode of a localization sensor (system) would only necessarily propagate towards higher system levels if the vehicle architecture consisted of a single channel. Such a single-channel architecture (as it is modelled in the PROFUND nets of chapter III) would fail to comply with dependability requirements.

As fail-safe (and even hazardous) modes of sensor channels are inevitable, multi-channel systems have to be considered (i.e., several of the nets depicted in chapter III in parallel), enabling the design of fail-operational systems, degraded modes, etc. Therefore, the design principles of fault-tolerance must be transferred to the design of the nominal function. To put it clearly, multiple channels have to be implemented for ensuring the dependability of the nominal function. If – in a second step – malfunctioning behavior (according to ISO 26262) is addressed, even more channels might be added.

For the realization of fail-operational resp. fault-tolerant systems, further strategies have to be implemented that exceed the aforementioned possibilities on the technical level, namely with respect to the system architecture: voting is a suitable means of enhancing dependability with multiple (i.e., redundant) channels. Voters dedicated to safety-critical applications can have an output only if there is consistency among a specific number of channels. Different voting algorithms have various requirements on the number of consistent inputs. Besides the consistency checking, other information (e.g., PL) can also be utilized by the voter to promote consensus on correct information [28]. In classic fault-tolerant design theory (dedicated for reducing malfunctioning behavior), homogeneous redundancy aims at reducing random errors, whereas heterogeneous redundancy makes use of diverse technology, design principles, etc. and thus also minimizes the risk originating from systematic errors. Transferred to the dependability of the nominal function, diverse redundancy is helpful to reduce the systematic influences of the environment (e.g., GNSS multipath, which is treated differently by the receivers). The principle of such a design is depicted in Fig. 19: The original measurement system (one channel, cf. Fig. 2) is multiplied and the output signals of both systems are subsequently fed into a voter. The different colors of the shadings indicate that different components, algorithms and design principles are used. It would also be possible to replace certain sensors by others (e.g., replace the IMU by a camera in one channel and leave the IMU in the other).

If one channel (of a multi-channel architecture) loses its trustworthiness, the measurement system may be designed to stay operational. Therefore, degradation strategies need to be defined, e.g., speed restrictions. If merely one channel is not trustworthy, speed might be unlimited, but the associated risk is higher. The system with n channels may be designed in a way that the risk associated with using n-1 channels is tolerable. If subsequently more channels lose their trustworthiness, the vehicle must drive slower. Only when the number of trustworthy channels falls below a critical number, it has to perform a minimum risk maneuver.

There are many degrees of freedom for the system designer to develop a suitable architecture. The designer must figure out if dependability requirements are fulfilled in a better way by means of data fusion (often a minimum variance estimator, e.g., Kalman filter) or by means of voting. In most cases, the optimum will be found by identifying the best combination of both strategies. An exemplary comparison, showing strengths and weaknesses of both approaches, is conducted in the following subchapter. More comprehensive strategies (out of scope of this paper) can be deducted from classic fault-tolerant design principles, e.g., using redundancy for building logic cascades that identify and exclude erroneous measurements.

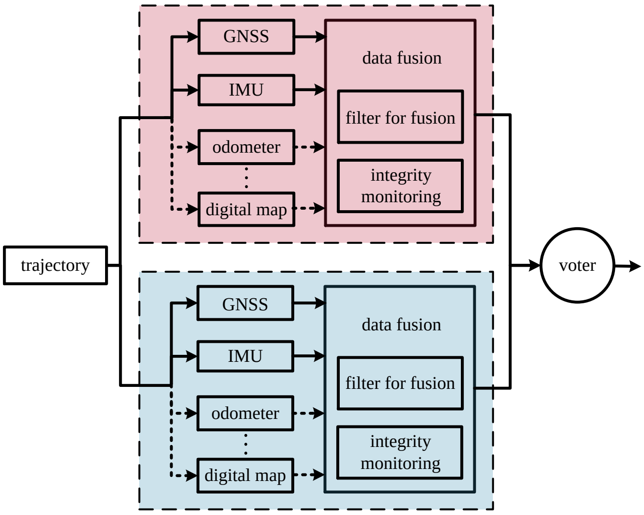

Fig. 19. Diversely redundant multi-channel measurement system [cf. 5].

C. Filtering vs. voting

A comparison between filtering and voting is conducted by simulation in MATLAB. The Kalman filter is chosen in the simulation among the filtering techniques. For the purpose of comparison, a double-channel system is constructed and the system architecture is depicted in Fig. 20. Each channel contains an integrated localization system. One takes information from an IMU and a GNSS receiver and the other one from an odometer and another GNSS receiver. IMU and odometer share no common error sources. Furthermore, GNSS measurements are impaired by modelled environmental effects such as NLOS. To provide an approximate indepence of measurements, they come from two independent receivers which have an installation distance of one meter in along-track direction. Due to the nonlinear characteristics of GNSS measurements in the pseudorange domain, the extended Kalman filter (EKF) is applied in both channels. The EKFs estimate the sensor errors and feed the estimation back to correct the errors in the sensors. The two channels provide the estimated positions r as well as dedicated horizontal protection levels (HPL, integrity monitoring).

The outputs of the two channels are then fed to a central Kalman filter (central KF) and a voter. The purpose of the central KF is to further improve the measurement quality. The estimated position error should be reduced compared to the outputs of the two channels. As a consequence, the values of the measurement quality characteristics should improve.

On the contrary, the voter determines the output of the overall system based on the state of the channels (depending on the trustworthiness of the input signals, cf. relation between PL and AL). In addition, the estimated positions from two channels should be in an acceptable range, so that the integrated systems in the two channels can be regarded as consistent and therefore operational.

For a voter implementing a 1oo2 logic, the overall system is operational when (classic resource dependability neglected):

i) at least one of the HPLs is smaller than the horizontal alarm limit (HAL) and
ii) the distance between the estimated positions is within a defined threshold.

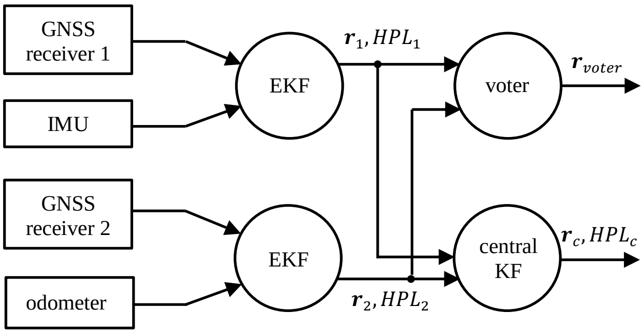

Fig. 20. Architecture of simulated system (**r**: position vector).

When condition ii) is fulfilled, the channel that has a protection level smaller than AL will pass its estimated position to the voter. When both channels have acceptable protection levels, the estimated position coming from the main channel will be passed to the output. In the simulation, the channel with IMU and GNSS is considered the main channel since it provides more localization information such as the pose of the vehicle.

The determination of HAL is dependent on the application. In the simulation, the value of HAL is assumed to be 12 m. The value of the threshold for condition ii) is usually determined by experience and it is a compromise between high false alarm rate and high missed detection rate [28]. In the simulation, it is assumed to be 5 m.

From the simulation results depicted in Fig. 21, it can be concluded that GNSS is able to correct the drift of IMU and odometer, but the positions estimated by both channels are still noisy due to environmental effects. As shown in Fig. 22, the HPL of the first channel is smaller than the HAL most time. By contrast, the HPL of channel 2 is less conservative, leading to higher availability but theoretically also to more (hazardously) misleading information.

With the aid of the voter, availability as well as MTTF (mean time to failure, a measure for continuity) can be raised in comparison to channel 1. However, since the two channels provide estimated positions of similar accuracies, especially in along track direction, the voter cannot improve the accuracy further but maintain an accuracy similar to the main channel. Contrary to that, the central KF is designed to reduce the noise in the estimated position, therefore, the accuracy of its output data is correspondingly better (Fig. 23). The measurement quality characteristics for the position estimated by the central KF are also calculated through the HPL. Since HPL_c is calculated with a more conservative method, the availability and MTTF are worse than those of the voter. A detailed comparison is shown in TABLE III.

In summary, both strategies have their individual strengths and drawbacks. The voter can increase availability as well as continuity, but it also inherits the missed detection errors from the main channel, requiring further countermeasures such as more complex voting architectures.

The Kalman filter can raise accuracy, which should indirectly contribute to higher dependability. To achieve an availabilty level comparable with the voter, a proper integrity monitoring algorithm should be developed for the central KF to avoid overconservative estimation. All in all, a combined architecture utilizing both voting and filtering seems beneficial.

TABLE III. QUALITY CHARACTERISTICS OF CHANNELS 1 AND 2, VOTER AND CENTRAL KF

	radial accuracy in m (95%)	availability in %	MTTF in s	missed detection rate in h^{-1}
IMU + GNSS	7.67	72.44	11.03	9.13
odometer + GNSS	6.05	98.73	77.89	0
voter	7.29	99.07	192.39	9.13
central KF	5.69	55.63	21.94	0

VI. Benefit for Dependability of Automated Driving

The formal approach developed in this paper depicts an additional facet for rendering automated driving more dependable (and thus safer). As experiences made in other domains (e.g., railway, aviation) prove, it is necessary for a safety-oriented system design to include safety principles right from the beginning, for example by means of formal methods (and not only by means of testing in a late phase of the development process). Thus, the PROFUND approach can serve as an *additional* method to enhance dependabililty in an early stage of the development process, when there is only a functional (and not yet a technical) model of the localization system. The formal approach based on Petri nets is advantageous over classic approaches (e.g., state machines), since it can cope with the complexity of automotive automation systems, e.g., because Petri nets are capable of modelling multi-state systems more efficiently. But as road vehicle systems are considered *open*, applying the PROFUND approach will be time-consuming. In contrast to other domains, it will most likely not be possible to ensure completeness. That is why the scenario-based approach – common to the automotive domain – proves useful: crucial scenarios can be modelled and simulated according to the PROFUND approach. The results of those simulations will shed light on the questions whether the system satisfies its dependability requirements or what kind of further measures (e.g., additional sensors) are required. It is worth pointing out that those valuable results will already be available in an early design phase.

Apart from the PROFUND approach, the redundancy and diversity principles discussed in chapter V might also prove benefical for safeguarding the nominal function of automated driving, though expensive. As longstanding experience made in the rail and in the aviation domain shows, both principles in combination are an effective means to mitigate random and systematic errors in the event that single-channel systems fail to meet dependability requirements.

VII. Conclusion and Outlook

Quantifying the automation risks of the nominal function involves a lot of scientific and technical domains. Therefore, terminology is still inconsistent and needs to be unified in order to avoid misunderstandings. The paper at hand suggests an approach for bringing together GNSS, data fusion, FuSa, SOTIF, dependability (PROFUND) and fault tolerance for road vehicle automation. Whereas the content of the first part of this paper is already modelled in Petri nets, the second part (e.g., multi-channel architectures) is still qualitative in most sections. As Petri nets can be simulated with the aim to calculate dependability metrics, multi-channel measurement systems as well as higher levels of the PROFUND net need to be modelled in order to formalize the behavior of degradation strategies and to predict the associated risk of the nominal function. In this way, the authors contribute to rendering vehicles of higher automation levels as so-called open systems controllable, with the aim to make automated vehicles behave in a more dependable manner.

Acknowledgment

We express our sincere thanks to Anna-Lena Müller for exhaustively correcting, editing and formatting this paper. In addition, we thank Christian Frohn for his support.

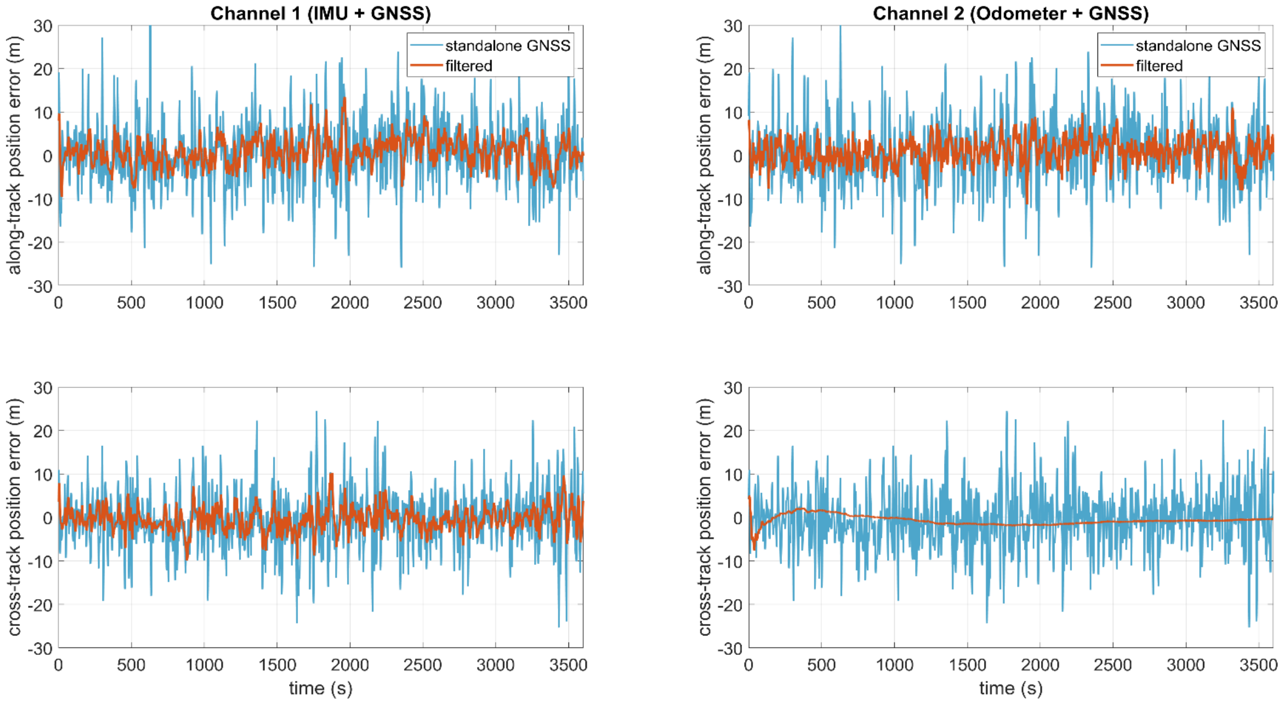

Fig. 21. Simulation results of standalone GNSS vs. filtered: position errors in along-track (top) and cross-track (bottom) direction for both channels. The error in the cross-track position of channel 2 is overly optimistic due to a simple model.

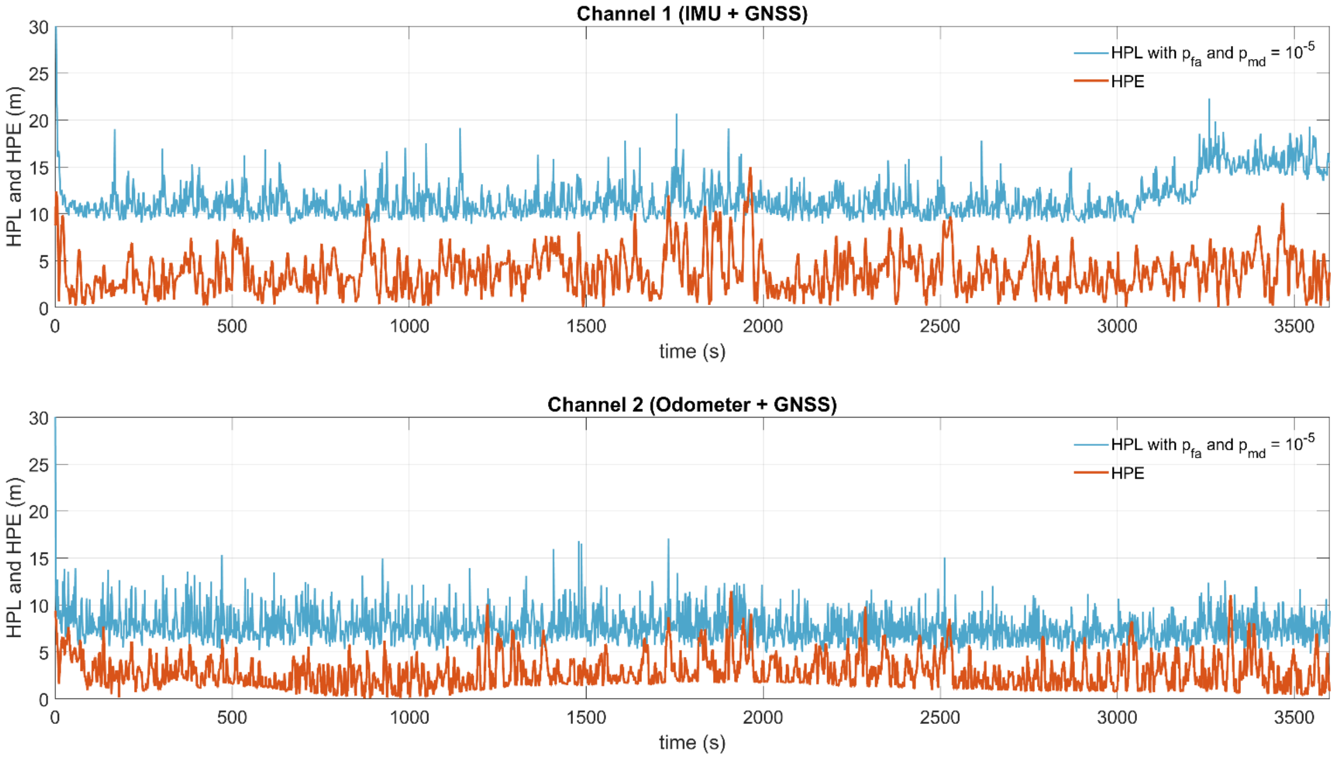

Fig. 22. Simulation results of channel 1 (top) vs. channel 2 (bottom): horizontal position error (HPE) vs. horizontal protection level (HPL). HPL calculation based on parameters p_{fa} (probability of false alarm) and p_{md} (probability of missed detection). HAL = 12 m.

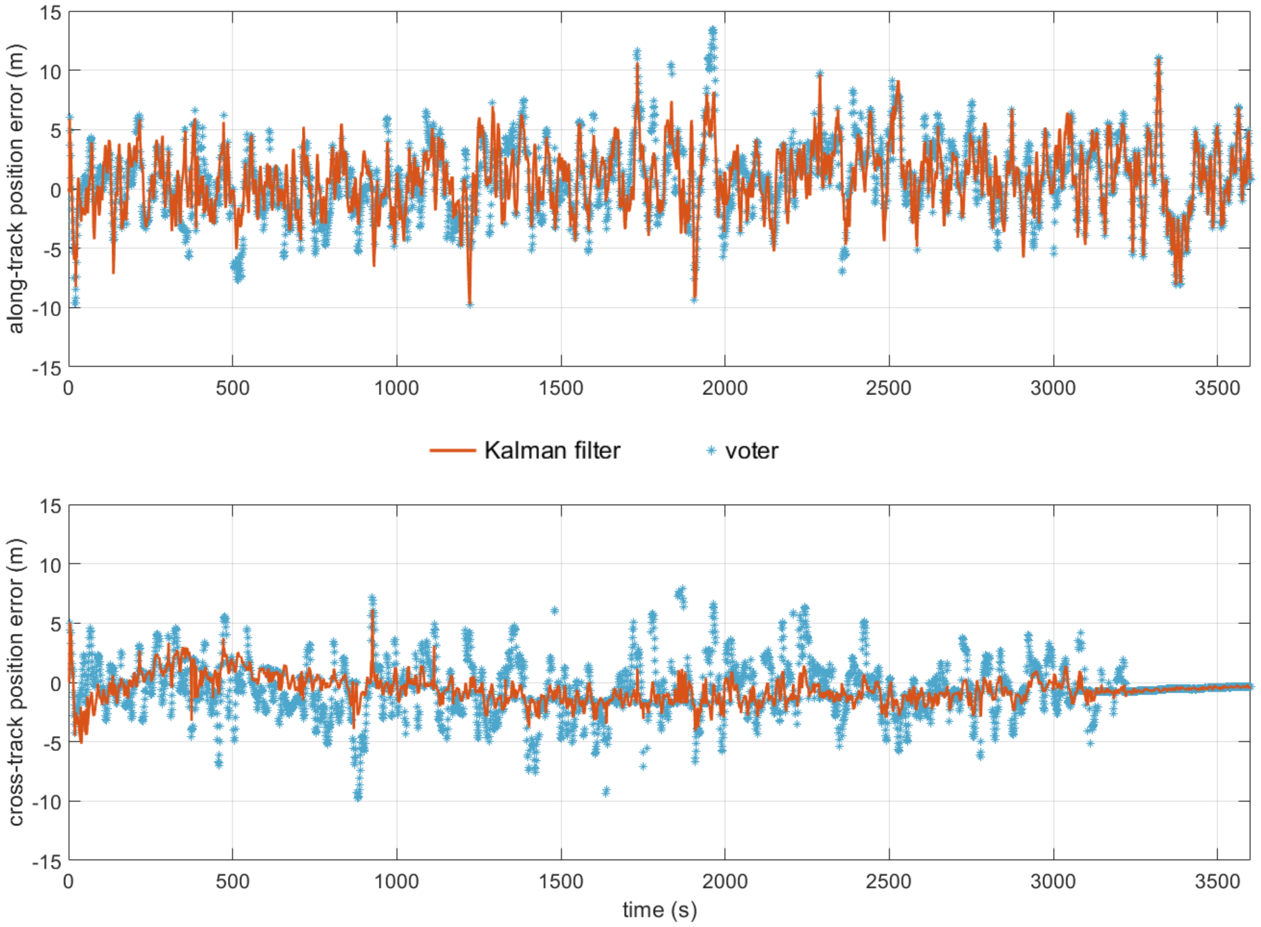

Fig. 23. Simulation results of along-track (top) and cross-track (bottom) position error: voter ($\boldsymbol{r}_{voter}$) vs. Kalman filter ($\boldsymbol{r}_c$), cf. Fig. 20.

References

[1] A. Burkert, “Roboterautos – Tabubruch einer Branche,” *ATZ*, vol. 121, no. 6, pp. 8–13, 2019.

[2] *Road vehicles -- Safety of the intended functionality; ISO/PAS 21448:2019*, 2019.

[3] M. Stiller, *Absicherung von Systemen für das (hoch)automatisierte Fahren.* [Online] Available: https://www.pegasusprojekt.de/files/tmpl/pdf/AutoTest%202016%20Abstract.pdf. Accessed on: Jul. 30 2019.

[4] M. Büker, B. Kramer, E. Böde, S. Vander Maelen, and M. Fränzle, “Identifikation von Automationsrisiken hochautomatisierter Fahrfunktionen in PEGASUS,” in *AAET: Automatisiertes und vernetztes Fahren*, ITS mobility e. V., Ed., 2019, pp. 315–329.

[5] A. Geffert, A. Dodinoiu, and U. Becker, “Multiperspektivischer Ansatz zur domänenübergreifenden Formalisierung von Verlässlichkeit am Beispiel der fahrzeugautonomen Ortung,” in *safe.tech 2019*, München, 2019.

[6] A. Geffert, T. Lan, A. Dodinoiu, and U. Becker, “Sicherheitsgerichteter, simulationsbasierter Entwurf fusionierter Fahrzeugortungssysteme in frühen Phasen des Entwicklungsprozesses,” in *DGON-Symposium: Positionierung und Navigation für Intelligente Verkehrssysteme*, Berlin, 2018.

[7] *Bahnanwendungen - Spezifikation und Nachweis von Zuverlässigkeit, Verfügbarkeit, Instandhaltbarkeit und Sicherheit (RAMS) - Teil 1: Generischer RAMS-Prozess; DIN EN 50126-1:2018-10*, Deutsche Fassung EN 50126-1:2018-10, 2018.

[8] E. Schnieder, “Verlässlichkeit von Verkehrssystemen im Verfügbarkeits-Sicherheits-Diagramm,” *Signal + Draht*, vol. 95, no. 10, pp. 6–9, 2003.

[9] N. Zhu, J. Marais, D. Bétaille, and M. Berbineau, “GNSS Position Integrity in Urban Environments: A Review of Literature,” *IEEE Transactions on Intelligent Transportation Systems*, https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8248655.

[10] M. Wegener, “Über die metrologische Qualität der Fahrzeugortung,” Ph.D. dissertation, Dept. Mech. Eng., Tech. Univ. Braunschweig, Braunschweig, 2013.

[11] Bureau International des Poids et Mesures, Ed., “International vocabulary of metrology: Basic and general concepts and associated terms (VIM),” 2012. [Online] Available: https://www.bipm.org/utils/common/documents/jcgm/JCGM_200_2012.pdf. Accessed on: Aug. 27 2019.

[12] D. Lu, “GNSS for Train Localisation Performance Evaluation and Verification,” Ph.D. dissertation, Dept. Mech. Eng., Tech. Univ. Braunschweig, Braunschweig, 2014.

[13] R. Slovák, “Methodische Modellierung und Analyse von Sicherungssystemen des Eisenbahnverkehrs,” Ph. D. dissertation, Dept. Mech. Eng., Tech. Univ. Braunschweig, Braunschweig, 2006.

[14] *Analysemethoden für Zuverlässigkeit - Petrinetze (IEC 62551:2012); Deutsche Fassung EN 62551:2012*, DIN EN 62551:2013-08; VDE 0050-4:2013-08 (IEC 62551:2012), 2013.

[15] E. Schnieder, *Methoden der Automatisierung: Beschreibungsmittel, Modellkonzepte und Werkzeige für Automatisierungssysteme*. Braunschweig/Wiesbaden: Friedr. Vieweg & Sohn Verlagsgesellschaft mbH, 1999.

[16] T. Walter, A. Hansen, and P. Enge, *Validation of the WAAS MOPS Integrity Equation.* [Online] Available: http://web.stanford.edu/group/scpnt/gpslab/pubs/papers/Walter_IONAM_1999_WAAS_MOPS_Integrity_Validation.pdf. Accessed on: Aug. 27 2019.

[17] P. Junietz, U. Steininger, and H. Winner, “Macroscopic Safety Requirements for Highly Automated Driving,” *Transportation Research Record Journal of the Transportation Research Board*, vol. 2673, no. 3, pp. 1–10, 2019.

[18] Federal Ministry of Transport and Digital Infrastructure, “Ethics Commission: Automated and Connected Driving,” Jun. 2017. [Online] Available: https://www.bmvi.de/SharedDocs/EN/publications/report-ethics-commission.pdf?__blob=publicationFile. Accessed on: Jul. 30 2019.

[19] L. Schnieder, P. Krumbach, “Positive Risikobilanzierung als ein Zulassungskriterium des hochautomatisierten Fahrens,” in *safe.tech 2019*, München, 2019.

[20] E. Schnieder, “(Verkehrs)sicherheit als regelungstechnische Aufgabe,” *at*, vol. 62, no. 12, pp. 829–841, https://www.degruyter.com/downloadpdf/j/auto.2014.62.issue-12/auto-2014-1133/auto-2014-1133.pdf, 2014.

[21] P. D. Groves, Z. Jiang, M. Rudi, and P. Strode, “A Portfolio Approach to NLOS and Multipath Mitigation in Dense Urban Areas,” in *Proceedings of the 26th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2013)*, Nashville, TN, 2013, pp. 3231–3247.

[22] M. Obst, S. Bauer, and G. Wanielik, “Urban multipath detection and mitigation with dynamic 3D maps for reliable land vehicle localization,” in *Proceedings of the 2012 IEEE/ION Position, Location and Navigation Symposium*, Myrtle Beach, SC, USA, Apr. 2012 - Apr. 2012, pp. 685–691.

[23] R. Toledo-Moreo, D. Betaille, and F. Peyret, “Lane-Level Integrity Provision for Navigation and Map Matching With GNSS, Dead Reckoning, and Enhanced Maps,” *IEEE Trans. Intell. Transport. Syst.*, vol. 11, no. 1, pp. 100–112, 2010.

[24] J. Liu, B.-g. Cai, T. Tang, and J. Wang, “A CKF based GNSS/INS train integrated positioning method,” in *2010 International Conference on Mechatronics and Automation: (ICMA); 4-7 August 2010*, Xi'an, China, 2010, pp. 1686–1689.

[25] O. Heirich, “Bayesian Train Localization with Particle Filter, Loosely Coupled GNSS, IMU, and a Track Map,” *Journal of Sensors*, vol. 2016, pp. 1–15, 2016.

[26] Satellite technology for advanced railway signalling, “D5.1 State of the art of EGNSS projects for the rail application,” 2017. [Online] Available: http://www.stars-rail.eu/wp-content/uploads/2017/05/D5.1-State-of-the-art-of-EGNSS-system-for-the-rail-application-updated…pdf. Accessed on: Aug. 23 2019.

[27] Satellite technology for advanced railway signalling, “D5.3 EGNSS Target Performances to meet railway safety requirements,” 2017. [Online] Available: http://www.stars-rail.eu/wp-content/uploads/2018/07/STR-WP5-D-ANS-034-07_-_D5.3_-_EGNSS_Target_Performances_to_meet_railway_safety_requirements_.pdf. Accessed on: Aug. 23 2019.

[28] G. Latif-Shabgahi, J. M. Bass, and S. Bennett, “A Taxonomy for Software Voting Algorithms Used in Safety-Critical Systems,” *IEEE Trans. Rel.*, vol. 53, no. 3, pp. 319–328, 2004.