

# Classification of Driving Direction in Traffic Surveillance using Magnetometers

Niklas Wahlström, Roland Hostettler, Fredrik Gustafsson, and Wolfgang Birk

This is a post-print of a paper published in *IEEE Transactions on Intelligent Transportation Systems*. When citing this work, you must always cite the original article:

N. Wahlström, R. Hostettler, F. Gustafsson, and W. Birk, “Classification of driving direction in traffic surveillance using magnetometers,” *Intelligent Transportation Systems, IEEE Transactions on*, vol. 15, no. 4, pp. 1405–1418, 2014

**DOI:**

10.1109/TITS.2014.2298199

**Copyright:**

© 2016 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

# Classification of Driving Direction in Traffic Surveillance using Magnetometers

Niklas Wahlström, *Student Member, IEEE*, Roland Hostettler, *Student Member, IEEE*, Fredrik Gustafsson, *Fellow, IEEE*, Wolfgang Birk *Member, IEEE*

**Abstract**—Traffic monitoring using low-cost 2-axis magnetometers is considered. Though detection of metallic vehicles is rather easy, detecting the driving direction is more challenging. We propose a simple algorithm based on a nonlinear transformation of the measurements, which is simple to implement in embedded hardware. A theoretical justification is provided, and the statistical properties of the test statistic are presented in closed form. The method is compared to the standard likelihood ratio test on both simulated data and real data from field tests, where very high detection rates are reported, despite the presence of sensor saturation, measurement noise and near-field effects of the magnetic field.

**Index Terms**—magnetometer, driving direction, classification, traffic surveillance, wireless sensor network

## I. INTRODUCTION

Traffic counting along particular roads is done either manually or electronically for the purpose of road improvement in the long term or re-routing traffic on the shorter time scale. The electronic devices used today include piezo-electric sensors or inductive loops under the road surface or pneumatic tubes on the road surface. Newer developments include radar, infrared light beams or cameras. For collecting road statistics, pneumatic tubes are today the most common solution. Recently the safety aspect of workers deploying the tubes has been in focus [2]. Another drawback of this solution is the short life-length of the tubes, which can be as short as 48 hours [3].

Magnetometers deployed along the road-side or hidden in the lane markers is a promising alternative, since they are cheap and small [4]. Compared to pneumatic tubes, solutions using magnetometers have a longer lifetime since they are not subject to the same amount of wear and do not expose personnel to traffic when mounting and dismounting, hence improving workplace safety.

The magnetometer is preferably part of a wireless sensor network (WSN), where sensor data from several magnetometers are transmitted to nearby nodes for a centralized or decentralized implementation of detection and classification algorithms [5], [6], [7]. However, such sensor nodes also bring certain challenges. Generally, the energy budget is limited as the units are powered by batteries and/or solar panels [8]. Furthermore, computational resources are scarce for reasons such

as power saving (e.g. duty-cycling of the computations or low-power processors) or sharing of the microcontroller between different tasks (measuring, computing, communication, etc.), see, for example, [9] or [10]. Thus, it is very important that the computation time for each task is reduced to a minimum, which emphasizes the need for low-complexity data processing algorithms.

One of the quantities of interest for road administrations, urban planners, or traffic management centers is traffic flow and, associated with that, the driving direction. Consider, for example, a single sensor monitoring a two-way two-lane highway. In order to be able to quantify the traffic volume on the individual lanes, the driving direction is crucial or, if one sensor for each lane is used, one would like to exclude vehicles on the farther lane. Thus, the traffic volume that is normally measured by a simple detector can be analyzed more thoroughly and better conclusions for future measures such as road planning can be drawn. In a similar setting, the driving direction can be used for detecting vehicles driving in the wrong lane, for example while overtaking. This information in turn can be used in order to warn upcoming traffic about a possible hazard in a cooperative collision warning system [11]. A third application where the driving direction is of utmost importance is the detection of wrong-way drivers. Wrong-way drivers are a very hazardous threat to other road users and can cause serious accidents [12]. Particularly on freeways, wrong-way drivers can cause serious head-on collisions and in 2010, wrong-way drivers accounted for 3.1 % of fatal crashes in the USA, causing 1,356 fatalities [13]. Thus, a system for detecting this kind of driving behavior can be of much help for authorities to detect vehicles driving in the wrong direction early on and warning other road users about the threat.

There are a number of different methods using various types of sensors for estimating the driving direction of a vehicle available today. One straight-forward approach is to use imaging sensors such as cameras or infrared laserscanners for tracking vehicles [14], [15]. Such sensors can provide very rich information and the driving direction of vehicles can be determined based on the estimated vehicle trajectory. However, challenging weather or illumination conditions degrade the performance significantly. One way of addressing this problem is to fuse the visual data with another type of sensor as proposed in [16] where a combination of a camera and a microphone array was used. While these approaches are viable and used in practice, they are not well suited for large scale deployment (for example at every freeway ramp) due to their requirements. Solutions more tailored for a system following

This work has been supported by the Swedish Foundation for Strategic Research under the project Cooperative Localization, the Swedish Governmental Agency for Innovation Systems (VINNOVA), and GEVEKO AB.

N. Wahlström and F. Gustafsson are with the Division of Automatic Control, Linköping University, Sweden.

R. Hostettler and W. Birk are with the Division of Systems and Interaction, Luleå University of Technology, Sweden.

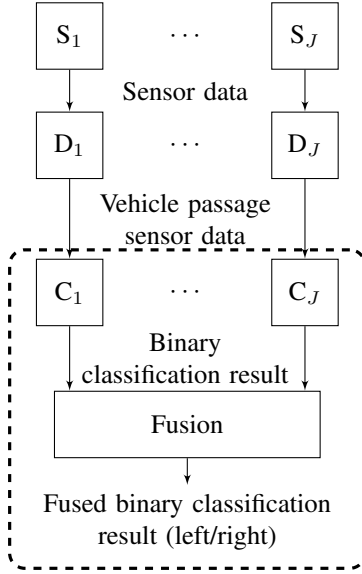


Fig. 1: Illustration of the signal flow. The sensor data from  $J$  sensors ( $S_1$  through  $S_J$ ) is used for detection ( $D_1$  through  $D_J$ ). Each sensor classifies the driving direction individually ( $C_1$  through  $C_J$ ) and a last, optional fusion step can be included to fuse these classification results. The dashed box indicates the parts of the system considered in this paper.

the requirements stated in the beginning are often based on two spatially separated sensors [6], [17]. However, the need for a second sensor can increase the cost considerably (up to twice the cost compared to only one sensor in the worst case) and introduces other challenges such as vehicle re-identification [18], [19] or the requirement of communication between the sensors [20]. Each of these activities will inevitably increase the energy consumption, which is a limited resource. Furthermore, a system based on only one single sensor is presumably more reliable since it does not depend on any second sensor that could break down.

In contrast to these approaches, this paper introduces a method for classifying the driving direction of a vehicle in a fast and efficient way addressing the initially stated requirements of a wireless sensor node. The method is based on one single magnetometer which measures magnetic field distortions induced by vehicles in its vicinity. Intuitively, extracting size and speed from this signal is rather straightforward. The basic principle is that the peak value of the measured signal is related to the size of the car, and the duration of the response is related to the speed of the vehicle.

However, obtaining the driving direction requires more physical insight about the signal and this problem has not been addressed before. In its simplest form, the proposed driving direction classification method only comprises a difference of two inner products of two vectors as it was first shown in [21]. This work is an extension of these findings and provides a more thorough statistical analysis as well as the evaluation of the classifier using more simulation and real measurement data. Specifically, the contributions of this work are:

- Extended version of the driving direction classification

method presented in [21] including an analysis of its statistical properties based on one single sensor.

- Verification of the proposed method using simulations as well as real measurement data.
- Comparison of the proposed method with a standard likelihood classification scheme.
- A sensor fusion strategy for multi-sensor scenarios.

In order to implement a complete system, also the detection of vehicles has to be considered. This can be accomplished by using adaptive thresholds [6] which is proven to be both robust and computationally cheap. However, in this paper, we will only consider the classification as well as the optional fusion step and assume that the detection and association problems are already solved (Figure 1).

The outline of this paper is as follows. The signal model describing the magnetic field distortion caused by a vehicle is presented in Section II. The proposed classifier and its statistical properties are given in Section III and a likelihood classifier is presented in Section IV, which will be used as a comparison to the proposed algorithm. The properties are verified and discussed by using Monte Carlo simulations in Section V and finally, the method is applied to real data in Section VI, followed by conclusions in Section VII.

## II. SIGNAL MODEL

The magnetometer signal induced by a metallic vehicle contains rich information which depends on both, the target trajectory as well as target specific parameters. A typical signal is displayed in Figure 2. In this work, we are only interested in determining the driving direction of the target. Consequently, the method should be insensitive to other quantities such as velocity, distance between the sensor and the trajectory, magnetic signature, and target extension.

One way of solving the problem is to approximately model the target as a magnetic dipole. This approximation holds if the distance between the target and the sensor is large in comparison to the characteristic magnetic length of the target [22] which is typically in the range of 1 m to 2 m for passenger cars and 4 m and more for larger vehicles such as busses and trucks. This gives rise to a magnetic dipole field  $\mathbf{h}(t)$  expressed as

$$\mathbf{h}(t) = \frac{3(\mathbf{r}(t) \cdot \mathbf{m})\mathbf{r}(t) - \|\mathbf{r}(t)\|^2\mathbf{m}}{\|\mathbf{r}(t)\|^5}, \quad (1a)$$

where  $\mathbf{r}(t) = [r^x(t) \ r^y(t) \ r^z(t)]^T$  is the position of the target relative to the sensor and  $\mathbf{m} = [m^x \ m^y \ m^z]^T$  is the magnetic dipole moment, which can be considered as a target specific parameter [23]. Two components of the magnetic field (1a) can then be measured with a 2-axis magnetometer

$$\mathbf{y}_k = \begin{bmatrix} y^x(kT) \\ y^y(kT) \end{bmatrix} = \begin{bmatrix} h^x(kT) \\ h^y(kT) \end{bmatrix} + \begin{bmatrix} e^x(kT) \\ e^y(kT) \end{bmatrix} = \bar{\mathbf{h}}_k + \mathbf{e}_k, \quad (1b)$$

where  $T$  is the sampling time,  $k$  denotes the sampling instant,  $\bar{\mathbf{h}}_k$  is a  $2 \times 1$  vector containing the x- and y-components (the first two components) of the  $3 \times 1$  vector  $\mathbf{h}_k = \mathbf{h}(kT)$ , and  $\mathbf{e}_k$

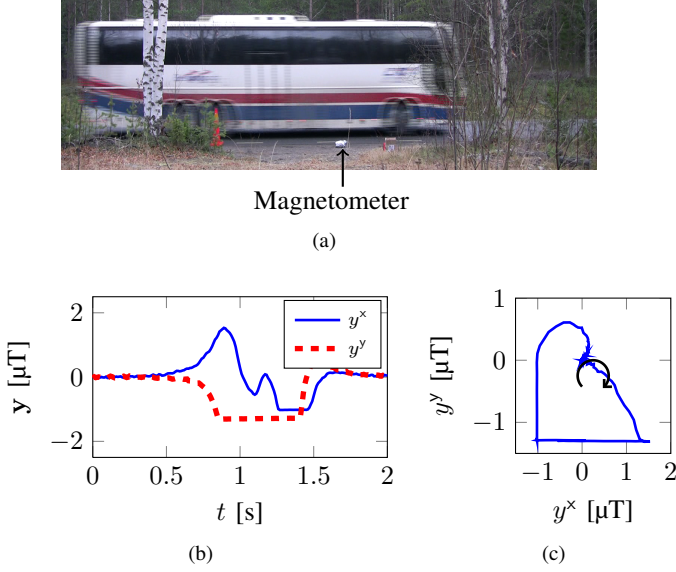


Fig. 2: A metallic vehicle (a) gives rise to a magnetic field distortion (b). The driving direction can be revealed by estimating the rotation direction of the magnetic field components (c).

is measurement noise assumed to be independent, identically distributed, zero mean white Gaussian noise of the form

$$\mathbf{e}_k \sim \mathcal{N}(0, \sigma^2 \mathbf{I}_2), \quad (1c)$$

where  $\mathbf{I}_2$  is the  $2 \times 2$  identity matrix. Furthermore, the following vector notation will be used

$$\mathbf{Y}_{m:n}^\alpha = [y_m^\alpha \ y_{m+1}^\alpha \ \dots \ y_n^\alpha]^\top, \quad (2)$$

where  $\alpha \in \{x, y\}$ .

The model in (1) can now be used to classify the driving direction of the vehicle. The two hypotheses of this binary classification problem will be denoted  $\mathcal{H}_L$  and  $\mathcal{H}_R$  representing that the vehicle passes the sensor from the left or from the right, respectively. One way of doing this is by estimating the unknowns  $\mathbf{r}(t)$  and  $\mathbf{m}$  from the measurement of  $\mathbf{y}_k$  and extract the direction information from the estimated trajectory  $\hat{\mathbf{r}}(t)$ . This can either be done in a batch approach where a whole data batch is used at once or through object tracking using, for example, a Kalman or particle filter as it has been done in [24], [25] and [26]. However, this is a non-linear problem and convergence to a global optimum is not guaranteed. Furthermore, if the target is close to the sensor, a higher order model including more parameters is needed to describe the signal accurately, for example by including higher order moments of the magnetic field or by modeling the target as a grid of dipoles [27]. Unfortunately, the computational cost of the corresponding estimation problem would in the worst case grow exponentially with the number of parameters [28].

Instead, an alternative method based on computing the cross-correlation between the different channels of the measurement is suggested and evaluated in this work. Furthermore, the proposed method will be compared to a likelihood ratio test based on the dipole model (1). This test can be seen as

a common practice procedure and is often used in detection and classification problems in all kinds of disciplines and is thus used as a benchmark [29], [30].

### III. CORRELATION-BASED CLASSIFIER

#### A. Method and Algorithm

It has been shown in [21] that the driving direction information can be obtained by computing the rotation direction of the magnetic field vector components, which is illustrated in Figure 2c. Specifically, the sign of the area

$$f^* = \int \begin{vmatrix} h^x & dh^x \\ h^y & dh^y \end{vmatrix} = \int \begin{vmatrix} h^x(t) & dh^x(t)/dt \\ h^y(t) & dh^y(t)/dt \end{vmatrix} dt \quad (3)$$

is the same as the sign of the area spanned by the position vector  $\mathbf{r}(t)$ . Using discrete time measurements, (3) can be approximated by

$$\begin{aligned} f_1 &= \sum_{k=1}^{N-1} \begin{vmatrix} h_k^x & (h_{k+1}^x - h_k^x)/T \\ h_k^y & (h_{k+1}^y - h_k^y)/T \end{vmatrix} T \\ &= \sum_{k=1}^{N-1} (h_k^x h_{k+1}^y - h_k^y h_{k+1}^x), \end{aligned} \quad (4)$$

which corresponds to the sum of the triangles spanned by two adjacent samples of the trajectory, see Figure 3a. Note that since all the triangles are completely enclosed in the true trajectory,  $f_1$  systematically underestimates the true area  $f^*$ . Using (4) and the measurement model (1b) the proposed correlation-based driving direction classifier is summarized in Algorithm 1, which is an extension of the earlier results in [21]. Note that using the vector notation in (2), subtractions and inner products of vectors are used in order to calculate (5) which is very beneficial for efficient implementation. Also, the algorithm is only parametrized by one single averaging parameter  $p$  introduced in (5a) which can be chosen as described in Section III-C below. The properties of the proposed algorithm are derived in the following section.

#### B. Properties

Given the algorithm as introduced in Algorithm 1, its properties are shown and derived in this section. Note that the assumption

$$\bar{\mathbf{h}}_k = \mathbf{0} \quad \text{for } k \leq 0 \vee k > N, \quad (6)$$

that is, that the magnetic signal has decayed to zero within the given window of  $N$  samples (however,  $\mathbf{e}_k$  is non-zero) is made throughout the remainder of the paper.

1) *Correlation with lag  $p = 1$* : The natural choice for the averaging parameter  $p$  introduced in (5a) is to use  $p = 1$  since it is essentially directly replacing the true magnetic field vector  $\bar{\mathbf{h}}_k$  in (4) with its noisy counterpart  $\mathbf{y}_k$  as

$$\hat{f}_1 = \sum_{k=1}^N (y_k^x y_{k+1}^y - y_k^y y_{k+1}^x), \quad (7)$$

which can be interpreted as being the difference of the cross-correlations between  $y^x$  and  $y^y$  with lag 1 and  $-1$  respectively.

---

**Algorithm 1** Correlation-based driving direction classification
 

---

1) Calculate the test statistic

$$\hat{f}_p = \frac{1}{p} \left( (\mathbf{Y}_{1:N}^x)^\top \mathbf{Y}_{(1+p):N+p}^y - (\mathbf{Y}_{1:N}^y)^\top \mathbf{Y}_{(1+p):N+p}^x \right) \quad (5a)$$

with  $\mathbf{Y}_{m:n}^\alpha$  as defined in (2) and  $p$  the correlation lag (parameter).

2) Determine the driving direction by testing the sign of (5a)

$$\hat{f}_p \underset{\mathcal{H}_L}{\overset{\mathcal{H}_R}{\geq}} 0 \quad (5b)$$

3) Estimate the variance of the test statistic

$$\begin{aligned} \hat{\sigma}_{\hat{f}_p}^2 = & \frac{\sigma^2}{p^2} \left( (\mathbf{Y}_{1+p:N+p}^x - \mathbf{Y}_{1-p:N-p}^x)^\top \right. \\ & \cdot (\mathbf{Y}_{1+p:N+p}^x - \mathbf{Y}_{1-p:N-p}^x) \\ & + (\mathbf{Y}_{1+p:N+p}^y - \mathbf{Y}_{1-p:N-p}^y)^\top \\ & \cdot (\mathbf{Y}_{1+p:N+p}^y - \mathbf{Y}_{1-p:N-p}^y) \Big) \\ & - \frac{2N}{p^2} \sigma^4 \end{aligned} \quad (5c)$$

4) Estimate the error probability

$$\hat{P}_E = \frac{1}{2} \operatorname{erfc} \left( \frac{|\hat{f}_p|}{\sqrt{2} \hat{\sigma}_{\hat{f}_p}} \right) \quad (5d)$$

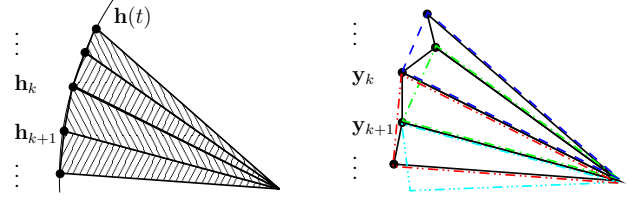

---

Since the measurement noise is assumed to be zero mean and i.i.d. (1c), it can be shown that the estimator (7) is unbiased

$$\begin{aligned} \mathbb{E}[\hat{f}_1] &= \mathbb{E} \left[ \sum_{k=1}^N (y_k^x y_{k+1}^y - y_k^y y_{k+1}^x) \right] \\ &= \sum_{k=1}^N \mathbb{E} [y_k^x y_{k+1}^y - y_k^y y_{k+1}^x] \\ &= \sum_{k=1}^N h_k^x h_{k+1}^y - h_k^y h_{k+1}^x \\ &= f_1. \end{aligned} \quad (8)$$

Further, the variance of (7) is given by

$$\begin{aligned} \sigma_{\hat{f}_1}^2 &\triangleq \operatorname{Var}(\hat{f}_1) = \mathbb{E}[(\hat{f}_1 - \mathbb{E}[\hat{f}_1])^2] \\ &= \mathbb{E} \left[ \left( \sum_{k=1}^N (h_k^x + e_k^x)(h_{k+1}^y + e_{k+1}^y) \right. \right. \\ &\quad \left. \left. - (h_k^y + e_k^y)(h_{k+1}^x + e_{k+1}^x) \right. \right. \\ &\quad \left. \left. - (h_k^x h_{k+1}^y - h_k^y h_{k+1}^x) \right)^2 \right] \\ &= \mathbb{E} \left[ \left( \sum_{k=1}^N h_k^x e_{k+1}^y + h_{k+1}^y e_k^x - h_k^y e_{k+1}^x - h_{k+1}^x e_k^y \right. \right. \\ &\quad \left. \left. + e_{k+1}^y e_k^x - e_{k+1}^x e_k^y \right)^2 \right]. \end{aligned} \quad (9)$$



(a) The quantity  $f_1$  in (4) corresponds to the sum of the triangles spanned by two adjacent samples of the trajectory.

(b) The estimator  $\hat{f}_1$  sums over the black solid line triangles whereas  $\hat{f}_2$  averages over the colored dashed triangles.

Fig. 3: Geometrical interpretation of the estimators.

Analyzing the sum in (9) it can be seen that every  $e_k^\alpha$  appears twice in the whole sum, once scaled by  $h_{k+1}^\beta$  and once by  $-h_{k-1}^\beta$  (where the superscript  $\beta$  denotes the in-plane component perpendicular to  $\alpha$ ). Making use of this and (6) yields

$$\begin{aligned} \sigma_{\hat{f}_1}^2 &= \mathbb{E} \left[ \left( \sum_{k=1}^N (h_{k+1}^y - h_{k-1}^y) e_k^x - (h_{k+1}^x - h_{k-1}^x) e_k^y \right. \right. \\ &\quad \left. \left. + e_{k+1}^y e_k^x - e_{k+1}^x e_k^y \right)^2 \right] \\ &= \sigma^2 \sum_{k=1}^N \left\| (\bar{\mathbf{h}}_{k+1} - \bar{\mathbf{h}}_{k-1}) \right\|^2 + 2N\sigma^4. \end{aligned} \quad (10)$$

From (10) it is seen that the variance is increased by the norm of the (approximate) gradient of the magnetic field vector  $\nabla \mathbf{h}_k \approx \|(\bar{\mathbf{h}}_{k+1} - \bar{\mathbf{h}}_{k-1})\|/2$  as well as the window length. Finally, note that the distribution of  $\hat{f}_1$  is given by  $\hat{f}_1 \sim \mathcal{N}(f, \sigma_{\hat{f}_1}^2)$  as  $N \rightarrow \infty$  as shown in Proposition 1 in Appendix A using  $p = 1$ .

2) *Correlation with lag  $p > 1$ :* As shown above, the variance for  $p = 1$  scales badly if the noise is large since the second term in (10) scales with  $\sigma^4$ . It is thus desirable to reduce this effect. This can be achieved by using an averaging estimator with lag  $p > 1$  in order to reduce the noise sensitivity. Instead of calculating the triangular area of two neighboring measurement points  $k$  and  $k+1$  on the trajectory, larger area segments between the points  $k$  and  $k+p$  are considered (Figure 3b). This yields the cross-correlation estimator with lag  $p$

$$\hat{f}_p = \frac{1}{p} \sum_{k=1}^N (y_k^x y_{k+p}^y - y_k^y y_{k+p}^x). \quad (11)$$

by conducting similar calculations as in (8) it is straightforward to show that for  $p \neq 1$

$$\mathbb{E}[\hat{f}_p] = \frac{1}{p} \underbrace{\sum_{k=1}^N h_k^x h_{k+p}^y - h_k^y h_{k+p}^x}_{\triangleq f_p} \neq f_1 \quad (12)$$

and (11) is thus a biased estimator of (4). However, since only the sign of  $f_1$  is of interest, this is acceptable. The variance

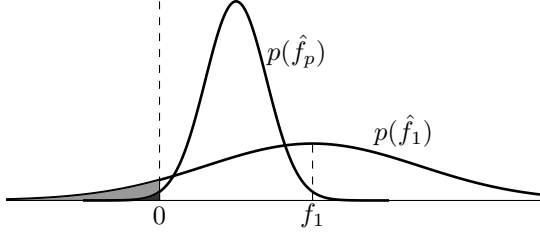


Fig. 4: Comparison of the probability density functions of the estimators  $\hat{f}_1$  and  $\hat{f}_p$ . The expected value of  $\hat{f}_p$  is biased towards zero compared to  $\hat{f}_1$ , however, the error probability (shaded areas) is much smaller.

is given by

$$\begin{aligned}\sigma_{\hat{f}_p}^2 &\triangleq \text{Var}(\hat{f}_p) = \mathbb{E}[(\hat{f}_p - \mathbb{E}[\hat{f}_p])^2] \\ &= \mathbb{E}\left[\left(\frac{1}{p} \sum_{k=1}^N (h_k^x + e_k^x)(h_{k+p}^y + e_{k+p}^y) - (h_k^y + e_k^y)(h_{k+p}^x + e_{k+p}^x) - (h_k^x h_{k+p}^y - h_k^y h_{k+p}^x)\right)^2\right] \\ &= \mathbb{E}\left[\left(\frac{1}{p} \sum_{k=1}^N h_k^x e_{k+p}^y + h_{k+p}^y e_k^x - h_k^y e_{k+p}^x - h_{k+p}^x e_k^y + e_{k+p}^y e_k^x - e_{k+p}^x e_k^y\right)^2\right].\end{aligned}\quad (13)$$

As in (9)-(10), the coefficients for  $e_k^\alpha$  can be grouped which gives

$$\begin{aligned}\sigma_{\hat{f}_p}^2 &= \mathbb{E}\left[\left(\frac{1}{p} \sum_{k=1}^N (h_{k+p}^y - h_{k-p}^y) e_k^x - (h_{k+p}^x - h_{k-p}^x) e_k^y + e_{k+p}^y e_k^x - e_{k+p}^x e_k^y\right)^2\right] \\ &= \frac{\sigma^2}{p^2} \sum_{k=1}^N \|\bar{\mathbf{h}}_{k+p} - \bar{\mathbf{h}}_{k-p}\|^2 + \frac{2N}{p^2} \sigma^4.\end{aligned}\quad (14)$$

From (14) it is seen that the variance is greatly reduced compared to (10). The second term scales with  $1/p^2$  compared to the second term in (10). As for the unbiased estimator  $\hat{f}_1$ , the distribution of  $\hat{f}_p$  converges to the normal distribution  $\hat{f}_p \sim \mathcal{N}(f_p, \sigma_{\hat{f}_p}^2)$  where  $f_p$  is the mean value as illustrated in (12) (see Proposition 1 in Appendix A).

The averaging effect is illustrated in Figure 4. Assuming that the true value  $f_1$  is positive, the expected value of the averaging estimator is moved towards zero due to the bias. However, the averaging reduces the variance and thus the total error probability (the shaded area under the probability density function up to zero) is reduced significantly.

In practice, it is of interest to estimate the error probability coupled to the estimate  $\hat{f}_p$  obtained from (11). For that reason, the variance of (11), which is given in (14), is of interest.

Noting that

$$\mathbb{E}[\|\mathbf{y}_{k+p} - \mathbf{y}_{k-p}\|^2] = \mathbb{E}\left[(y_{k+p}^x - y_{k-p}^x)^2 + (y_{k+p}^y - y_{k-p}^y)^2\right]$$

and letting  $z_k^\alpha = y_{k+p}^\alpha - y_{k-p}^\alpha$  it follows that  $z_k^\alpha \sim \mathcal{N}(h_{k+p}^\alpha - h_{k-p}^\alpha, 2\sigma^2)$ . Thus,

$$\mathbb{E}[(z_k^x)^2 + (z_k^y)^2] = (h_{k+p}^x - h_{k-p}^x)^2 + 2\sigma^2 + (h_{k+p}^y - h_{k-p}^y)^2 + 2\sigma^2$$

and finally

$$\mathbb{E}[\|\mathbf{y}_{k+p} - \mathbf{y}_{k-p}\|^2] = \|\bar{\mathbf{h}}_{k+p} - \bar{\mathbf{h}}_{k-p}\|^2 + 4\sigma^2. \quad (15)$$

Using (14) and (15) we can then estimate  $\text{Var}(\hat{f}_p)$  as follows

$$\begin{aligned}\hat{\sigma}_{\hat{f}_p}^2 &= \frac{\sigma^2}{p^2} \sum_{k=1}^N (\|\mathbf{y}_{k+p} - \mathbf{y}_{k-p}\|^2 - 4\sigma^2) + \frac{2N}{p^2} \sigma^4 \\ &= \frac{\sigma^2}{p^2} \sum_{k=1}^N \|\mathbf{y}_{k+p} - \mathbf{y}_{k-p}\|^2 - \frac{2N}{p^2} \sigma^4.\end{aligned}\quad (16)$$

The probability that the sign of the estimated  $\hat{f}_p$  is wrong compared to the sign of  $f_p$  is given by

$$\Pr(\text{sgn}(\hat{f}_p) \neq \text{sgn}(f_p)) = \frac{1}{2} \text{erfc}\left(\frac{|f_p|}{\sqrt{2}\sigma_{\hat{f}_p}}\right). \quad (17)$$

Since neither the true  $f_p$  nor  $\sigma_{\hat{f}_p}^2$  are known, the estimated values  $\hat{f}_p$  and  $\hat{\sigma}_{\hat{f}_p}^2$  can be used instead. The estimated error probability (17) then becomes

$$\hat{P}_E = \frac{1}{2} \text{erfc}\left(\frac{|\hat{f}_p|}{\sqrt{2}\hat{\sigma}_{\hat{f}_p}}\right), \quad (18)$$

which can be evaluated numerically. Note, however, that (18) has a slightly different meaning than (17). It indicates the probability of  $f_p$  having a different sign compared to the given  $\hat{f}_p$ .

### C. Parameter Tuning

The lag  $p$  introduced in (11) will improve the classification result as explained in Section III-B. Our objective is to choose a value  $p$  which minimizes the overall estimated error probability (18). In theory this could be done for each detection separately. However, that would require a non-linear search in the parameter  $p$  for each detection, which does not meet the needs for a computationally efficient implementation. Instead we will minimize the mean of the estimated error probabilities from a training set of estimation data  $l = 1, 2, \dots, L$  and use this value  $p$  afterwards. Given a set of estimation data  $(\mathbf{Y}_{1:N})_{1:L}$  we compute  $p$  as

$$\begin{aligned}p &= \underset{p}{\text{argmin}} \frac{1}{L} \sum_{l=1}^L \hat{P}_{E,l} \\ &= \underset{p}{\text{argmin}} \sum_{l=1}^L \text{erfc}\left(\frac{|\hat{f}_{p,l}|}{\sqrt{2}\hat{\sigma}_{\hat{f}_{p,l}}}\right).\end{aligned}\quad (19)$$

After finding this value  $p$ , all future classifications can be performed by using Algorithm 1 and the given value for  $p$ .

#### D. Sensor Fusion

Information from multiple sensors can be fused together in order to arrive at a joint-classification of multiple sensors. The fusion rule (49) for Bernoulli random variables as derived in Appendix B is used in order to reach a joint decision of the driving direction of  $J$  sensors as follows. Let  $p_j$  be the probability that the car is passing the sensor from left to right (hypothesis  $\mathcal{H}_L$  true) which is given by

$$p_j = \Pr(\mathcal{H}_L | \hat{f}_{p,j}, \hat{\sigma}_{\hat{f}_{p,j}}) = \frac{1}{2} \operatorname{erfc} \left( \frac{\hat{f}_{p,j}}{\sqrt{2} \hat{\sigma}_{\hat{f}_{p,j}}} \right). \quad (20)$$

A Bernoulli random variable  $k$  with

$$p(k|p_j) = p_j^k (1 - p_j)^{1-k} \quad \text{for } k \in \{0, 1\} \quad (21)$$

can now be used to represent the probability of each hypothesis where the value  $k = 0$  is assigned to  $\mathcal{H}_R$  and  $k = 1$  to  $\mathcal{H}_L$ . Finally, using (49) (see Appendix B) and (21) yields the sensor fusion decision rule given by

$$\Pr(\mathcal{H}_L | \mathbf{p}) = \frac{\prod_{j=1}^J p_j}{\prod_{j=1}^J (1 - p_j) + \prod_{j=1}^J p_j} \stackrel{\mathcal{H}_L}{\geq} \frac{1}{2}, \quad (22)$$

where  $\mathbf{p} = [p_1 \dots p_J]^T$ . Equation (22) can also be rewritten as

$$\Pr(\mathcal{H}_L | \mathbf{p}) = \frac{\prod_{j=1}^J \Pr(\mathcal{H}_L | \hat{f}_{p,j})}{\prod_{j=1}^J \Pr(\mathcal{H}_R | \hat{f}_{p,j}) + \prod_{j=1}^J \Pr(\mathcal{H}_L | \hat{f}_{p,j})}, \quad (23)$$

which can be interpreted as the joint probability for all sensors indicating  $\mathcal{H}_L$  at the same time, normalized by the sum of the same and its complementary event, that is, all sensors indicating  $\mathcal{H}_R$  at the same time.

Notice that the classification is performed in a distributed manner by first computing the ratios  $p_j$  in each sensor according to (20), and then, these values are fused according to (23).

#### IV. LIKELIHOOD TEST

In order to benchmark the proposed driving direction classification algorithm proposed in Section III, a generalized likelihood ratio test based on the measurement model (1) is derived in this section which is then compared to the proposed method in the following sections.

##### A. Single Sensor

Consider a vehicle passing the sensor with a constant velocity and constant lateral distance. The position  $\mathbf{r}_k$  can be rewritten as

$$\mathbf{r}_k = \mathbf{r}(kT) = \begin{bmatrix} v(kT - t_{\text{CPA}}) \\ r^y \\ 0 \end{bmatrix}, \quad (24)$$

where  $v$  is the vehicle speed,  $t_{\text{CPA}}$  is the closest point of approach time, and  $r^y$  the lateral distance between the target and the sensor.

It can be safely assumed that most of the vehicles will adhere to the known speed limit  $v_{\text{limit}}$  for a given road and thus, vehicles passing the sensor can be classified according to the following two hypotheses:

$$\mathcal{H}_L: \quad \boldsymbol{\theta}_1^* = [v_{\text{limit}} \quad r_1^y]^T, \quad (25a)$$

$$\mathcal{H}_R: \quad \boldsymbol{\theta}_2^* = [-v_{\text{limit}} \quad r_2^y]^T, \quad (25b)$$

where  $\boldsymbol{\theta}_i^*$  is the hypothesis parametrization which is known. The lateral distances  $r_1^y$  and  $r_2^y$  are derived from the road geometry. For example, in a traffic counting scenario, they would correspond to the distances to the closer and farther lane, respectively. On the other hand, when detecting wrong-way drivers on a freeway ramp, they would be equal. The hypothesis  $\mathcal{H}_L$  corresponds to a vehicle passing the sensor from left to right and  $\mathcal{H}_R$  to a vehicle passing the sensor from right to left.

The remaining parameters

$$\boldsymbol{\theta} = [m^x \quad m^y \quad m^z \quad t_{\text{CPA}}]^T \quad (26)$$

in (1a) are all unknown (opposite to  $\boldsymbol{\theta}^*$  which holds the known parameters determined as described above) and the measurement model can be rewritten as

$$\bar{\mathbf{h}}_k(\boldsymbol{\theta}_i^*, \boldsymbol{\theta}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \cdot \frac{(3\mathbf{r}_k \mathbf{r}_k^T + \|\mathbf{r}_k\|^2 \mathbf{I}_3)}{\|\mathbf{r}_k\|^5} \mathbf{m}, \quad (27)$$

where the position  $\mathbf{r}_k$  is a function of the parameters  $r_i^y$ ,  $v_i$  and  $t_{\text{CPA}}$  as given by (24). The measurement model is linear in the unknown vehicle dependent parameters  $\mathbf{m}$  and non-linear in  $t_{\text{CPA}}$ . These have to be estimated before the actual likelihood test can be performed. This can be done by using a maximum likelihood estimator which yields a generalized likelihood ratio test (GLRT) [29].

The joint probability density function for all  $N$  vector samples is given by

$$\begin{aligned} p(\mathbf{Y}_{1:N}; \boldsymbol{\theta}_i^*, \boldsymbol{\theta}) &= \prod_{k=1}^N p(\mathbf{y}_k; \boldsymbol{\theta}_i^*, \boldsymbol{\theta}) \\ &= \frac{1}{(2\pi\sigma^2)^{2N/2}} \exp \left( -\frac{1}{2\sigma^2} \sum_{k=1}^N \|\mathbf{y}_k - \bar{\mathbf{h}}_k(\boldsymbol{\theta}_i^*, \boldsymbol{\theta})\|_2^2 \right) \\ &= \frac{1}{(2\pi\sigma^2)^{2N/2}} \exp \left( -\frac{1}{2\sigma^2} \|\mathbf{Y}_{1:N} - \mathbf{H}_{1:N}(\boldsymbol{\theta}_i^*, \boldsymbol{\theta})\|_2^2 \right), \end{aligned} \quad (28)$$

where the measurement samples are stacked as

$$\mathbf{Y}_{m:n} = \begin{bmatrix} \mathbf{Y}_{m:n}^x & \mathbf{Y}_{m:n}^y \end{bmatrix}^T, \quad (29)$$

$\mathbf{Y}_{m:n}^\alpha$  is as defined in (2), and equivalently for  $\mathbf{H}_{1:N}$ . The maximum likelihood estimator [31] for the parameters  $\boldsymbol{\theta}$  is then simply

$$\hat{\boldsymbol{\theta}}_i = \underset{\boldsymbol{\theta}}{\operatorname{argmax}} p(\mathbf{Y}_{1:N}; \boldsymbol{\theta}_i^*, \boldsymbol{\theta}) \quad \text{for } i = 1, 2. \quad (30)$$

where the estimate  $\hat{\boldsymbol{\theta}}_i$  depends on the hypothesis  $\boldsymbol{\theta}_i^*$ .

Once the estimation  $\hat{\boldsymbol{\theta}}_i$  is obtained, the likelihood ratio can be calculated as

$$l = \frac{p(\mathbf{Y}_{1:N}; \boldsymbol{\theta}_1^*, \hat{\boldsymbol{\theta}}_1)}{p(\mathbf{Y}_{1:N}; \boldsymbol{\theta}_2^*, \hat{\boldsymbol{\theta}}_2)} \stackrel{\mathcal{H}_L}{\geq} \frac{\mathcal{H}_L}{\mathcal{H}_R}. \quad (31)$$

If  $l > 1$ , the hypothesis  $\mathcal{H}_L$  is more likely to be true and  $\mathcal{H}_R$  otherwise. Using (28) and (31), the log-likelihood ratio is given by

$$\begin{aligned} \lambda &= \log(l) \\ &= -\frac{1}{2\sigma^2} \left\| \mathbf{Y}_{1:N} - \mathbf{H}_{1:N}(\boldsymbol{\theta}_1^*, \hat{\boldsymbol{\theta}}_1) \right\|_2^2 \\ &\quad + \frac{1}{2\sigma^2} \left\| \mathbf{Y}_{1:N} - \mathbf{H}_{1:N}(\boldsymbol{\theta}_2^*, \hat{\boldsymbol{\theta}}_2) \right\|_2^2 \\ &\stackrel{\mathcal{H}_L}{\underset{\mathcal{H}_R}{\geq}} 0 \end{aligned} \quad (32)$$

and the decision rule becomes

$$\tilde{\lambda}_1 \stackrel{\mathcal{H}_L}{\underset{\mathcal{H}_R}{\geq}} \tilde{\lambda}_2 \quad (33)$$

with

$$\tilde{\lambda}_i = - \left\| \mathbf{Y}_{1:N} - \mathbf{H}_{1:N}(\boldsymbol{\theta}_i^*, \hat{\boldsymbol{\theta}}_i) \right\|_2^2 \quad (34)$$

Note that the two test statistics  $\tilde{\lambda}_1$  and  $\tilde{\lambda}_2$  are easily calculated. However, the parameter estimation step to be executed still requires solving a (separable) non-linear problem. Hence, this method is not well tailored for implementation in systems with limited computational power since it requires an iterative solver which might not converge to the global optimum. Finally, the likelihood algorithm is summarized in Algorithm 2.

---

**Algorithm 2** Generalized likelihood ratio test driving direction classification

---

- 1) Estimate the model parameters  $\boldsymbol{\theta}$  as defined in (26) for the hypotheses  $\{\mathcal{H}_L, \mathcal{H}_R\}$  using the maximum likelihood estimator

$$\hat{\boldsymbol{\theta}}_i = \underset{\boldsymbol{\theta}}{\operatorname{argmax}} p(\mathbf{Y}_{1:N}; \boldsymbol{\theta}_i^*, \boldsymbol{\theta}) \quad \text{for } i = 1, 2. \quad (35a)$$

with  $\boldsymbol{\theta}_i^*$  as defined in (25).

- 2) Calculate the test statistic

$$\begin{aligned} \lambda &= \log(l) \\ &= -\frac{1}{2\sigma^2} \left\| \mathbf{Y}_{1:N} - \mathbf{H}_{1:N}(\boldsymbol{\theta}_1^*, \hat{\boldsymbol{\theta}}_1) \right\|_2^2 \\ &\quad + \frac{1}{2\sigma^2} \left\| \mathbf{Y}_{1:N} - \mathbf{H}_{1:N}(\boldsymbol{\theta}_2^*, \hat{\boldsymbol{\theta}}_2) \right\|_2^2 \end{aligned} \quad (35b)$$

with  $\mathbf{Y}_{1:N}$  and  $\mathbf{H}_{1:N}(\boldsymbol{\theta}_i^*, \hat{\boldsymbol{\theta}}_i)$  as defined in (29).

- 3) Determine the driving direction by testing the sign of (35b)

$$\lambda \stackrel{\mathcal{H}_L}{\underset{\mathcal{H}_R}{\geq}} 0 \quad (35c)$$


---

### B. Sensor Fusion

If there is data from more than one sensor available, the likelihood test can take advantage of all this information in order to make a well-balanced decision. Let  $\boldsymbol{\theta}_j$  be the

parameters and  $\mathbf{Y}_{1:N,j}$  be the measurements for the  $j$ th sensor. Then the joint PDF for all measurement is given by

$$p(\boldsymbol{\Upsilon}; \boldsymbol{\theta}_i^*, \boldsymbol{\Theta}) = \prod_{j=1}^J p(\mathbf{Y}_{1:N,j}; \boldsymbol{\theta}_i^*, \boldsymbol{\theta}_j), \quad (36)$$

where  $\boldsymbol{\Upsilon} = [\mathbf{Y}_{1:N,1}^\top \quad \mathbf{Y}_{1:N,2}^\top \quad \dots \quad \mathbf{Y}_{1:N,J}^\top]^\top$  and  $\boldsymbol{\Theta} = [\boldsymbol{\theta}_1^\top \quad \boldsymbol{\theta}_2^\top \quad \dots \quad \boldsymbol{\theta}_J^\top]^\top$ . The overall likelihood ratio and decision rule is then given by

$$l_J = \frac{p(\boldsymbol{\Upsilon}; \boldsymbol{\theta}_1^*, \hat{\boldsymbol{\Theta}}_1)}{p(\boldsymbol{\Upsilon}; \boldsymbol{\theta}_2^*, \hat{\boldsymbol{\Theta}}_2)} = \frac{\prod_{j=1}^J p(\mathbf{Y}_{1:N,j}; \boldsymbol{\theta}_1^*, \hat{\boldsymbol{\theta}}_{1,j})}{\prod_{j=1}^J p(\mathbf{Y}_{1:N,j}; \boldsymbol{\theta}_2^*, \hat{\boldsymbol{\theta}}_{2,j})} \stackrel{\mathcal{H}_L}{\underset{\mathcal{H}_R}{\geq}} 1, \quad (37)$$

which results in the log likelihood

$$\begin{aligned} \lambda_J &= \log(l_J) \\ &= \sum_{j=1}^J \frac{1}{2\sigma_j^2} \tilde{\lambda}_{1,j} - \sum_{j=1}^J \frac{1}{2\sigma_j^2} \tilde{\lambda}_{2,j} \\ &= \sum_{j=1}^J \frac{1}{2\sigma_j^2} \left( \tilde{\lambda}_{1,j} - \tilde{\lambda}_{2,j} \right) \stackrel{\mathcal{H}_L}{\underset{\mathcal{H}_R}{\geq}} 0. \end{aligned} \quad (38)$$

Finally, (38) can be rewritten as

$$\sum_{j=1}^J \frac{1}{\sigma_j^2} \tilde{\lambda}_{1,j} \stackrel{\mathcal{H}_L}{\underset{\mathcal{H}_R}{\geq}} \sum_{j=1}^J \frac{1}{\sigma_j^2} \tilde{\lambda}_{2,j}. \quad (39)$$

Notice that the classification is performed in a distributed manner by first computing the ratios  $\tilde{\lambda}_{i,j}$  in each sensor according to (34), and then, these values are fused according to (39).

## V. SIMULATION

Before applying the classifier derived in the previous section on real data, a simulation will be used to visualize and validate the properties of the proposed estimator. In the end of this section, the proposed classifier is also compared to the generalized likelihood test presented in Section IV.

Consider a simulation setup with a vehicle heading in positive x-direction, starting at  $\mathbf{r}_1 = [-5 \quad 1 \quad 0]^\top$  and ending at  $\mathbf{r}_N = [5 \quad 1 \quad 0]^\top$  divided into  $N = 100$  data points in between. Furthermore, consider a magnetic dipole moment of  $\mathbf{m} = [1 \quad 1 \quad 1]^\top$ . We will simulate this example with different levels of the signal to noise ratio (SNR), which is defined as

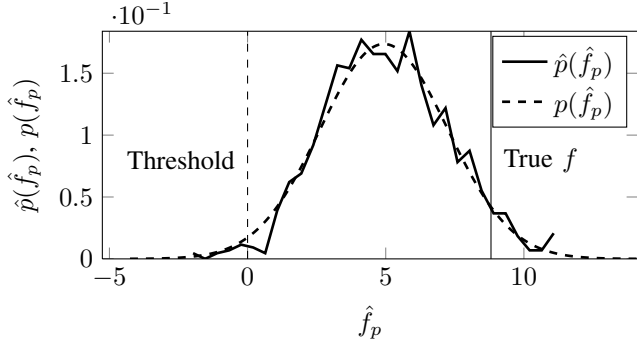
$$\text{SNR} = 10 \log_{10} \left( \frac{\frac{1}{N} \sum_{k=1}^N \|\bar{\mathbf{h}}_k\|^2}{\sigma^2} \right) \text{dB}, \quad (40)$$

where  $\sigma^2$  is the variance of the measurement noise.

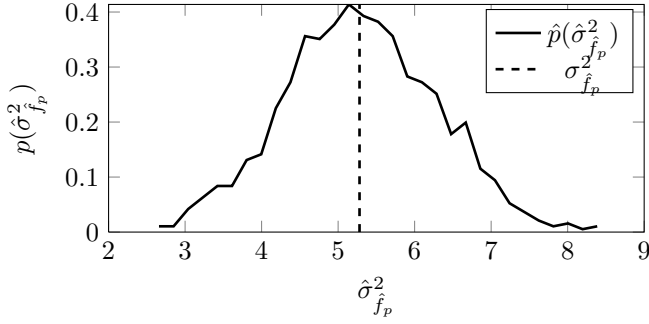
### A. Estimate and Variance Estimate

In (12) and (14) expressions for the mean  $f_p = \mathbb{E}[\hat{f}_p]$  and variance  $\sigma_{f_p}^2 = \mathbb{E}[(\hat{f}_p - f_p)^2]$  of the estimators are given. These expressions are verified by performing 1,000 Monte Carlo simulations for the presented example with different noise realization for each run. The result is presented in





(a) Empirical distribution  $\hat{p}(\hat{f}_p)$  together with the theoretical distribution  $p(\hat{f}_p)$ .



(b) Empirical distribution  $\hat{p}(\hat{\sigma}_{f_p}^2)$  together with the theoretical variance.

Fig. 5: Simulation results of the estimator (a) and its estimated variance (b) for a scenario with SNR = -10 dB and  $p = 15$ .

Figure 5a with SNR = -10 dB and  $p = 15$  together with the theoretical distribution of the estimator  $\hat{f}_p \sim \mathcal{N}(f_p, \sigma_{f_p}^2)$ .

According to the result, the theoretical distribution corresponds well to the empirical one. Furthermore, note that the estimate is biased  $E[\hat{f}_p] \neq f_1$ , as already stated in (12). We can also conclude that the Gaussian assumption of the estimator distribution is indeed valid.

Each sequence of data does not only provide us with the estimate (11), but also with an estimate of its variance (16). In Figure 5b, this estimated variance is compared with the true variance, using the same 1,000 Monte Carlo simulations as previously. According to this result, the variance estimate seems to be unbiased as expected from the derivation.

### B. Dependency of $P_E$ on SNR and $p$

In (17) a scheme for computing the error probability is proposed by knowing the true mean and variance of the estimator. This value has been compared with the actual error classification rate by performing 1,000 Monte Carlo simulations for different values of  $p$  and SNR. The result is provided in Figure 6 and the theoretical values display a good agreement with the simulations. Also the classification performance increases with higher SNR which is natural.

Furthermore, this result also shows the classification improvement of the cross-correlation method by choosing a lag  $p > 1$ . Also note, that for the chosen simulation scenario, there is an optimal  $p \approx 15$  which is fairly independent of the

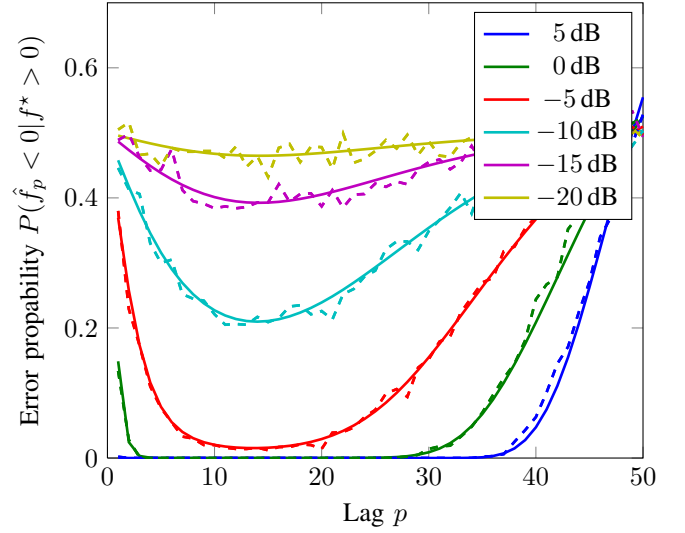


Fig. 6: Classification performance as a function of the lag  $p$  for different SNRs. The solid line is the theoretical performance  $P_E$  according to (17) and the dashed line the average error probability of the 1,000 Monte Carlo simulations.

SNR. However, this will depend on the magnetic moment of the vehicle  $\mathbf{m}$ , the trajectory  $\mathbf{r}_k$ , and the data length  $N$ .

### C. Comparison with Likelihood Test

As a reference, the proposed classifier can be compared to the likelihood test presented in Section IV. Again, 1,000 Monte Carlo simulations with  $p = 15$  for the correlation classifier and the true values for  $v_1, v_2, r_1^y$  and  $r_2^y$  for the likelihood test were run. Figure 7 shows the error rates for the two classifiers as functions of the SNR. As can be expected, both classifiers perform well for high SNRs, down to about -5 dB where the error rates start to increase until the point of “tossing a coin” somewhere below -20 dB is reached. It should be noted, however, that the correlation classifier requires an SNR of about 5 dB higher than the likelihood classifier in order to achieve the same classification rate. This is not very surprising since the likelihood test is expected to be the optimal test for this scenario since the likelihood test is performed under the same model and model parameter as have been used in the simulation. However, it will be shown that the likelihood test is more sensitive to violations of the model assumptions such as dipole model, speed, or trajectory, which were used to derive the GLRT. These results are presented in the next Section.

## VI. EXPERIMENTAL RESULTS AND DISCUSSION

The simulations in the preceding section indicate that the proposed classifier works according to the expectations from its derivation. This section will now show how the classifier performs on real data where a bigger amount of uncertainty and challenges are to be expected.

### A. Experiment Setup

In order to verify the proposed algorithm, real world experiments have been conducted on a two-way country road

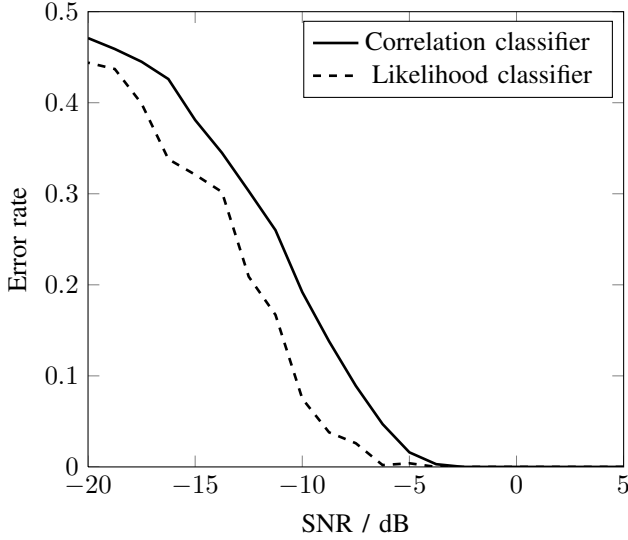


Fig. 7: Classification performance compared with the likelihood test as a function of the SNR.

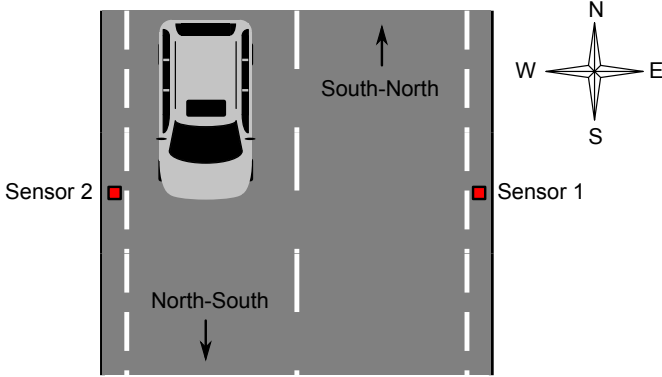


Fig. 8: Illustration of the experiment setup showing the two sensors on each side of the road as well as the driving directions.

with moderate traffic density and a speed limit of 90 km/h. Two commercially available 2-axis magnetometers (Honeywell HMC6042, [32]) sampling at 100 Hz were deployed on both sides of the road as illustrated in Figure 8. The traffic was measured during three separate periods for a total of 158 minutes. In total 362 vehicles traveling south-north (close to sensor 1) and 305 vehicles traveling north-south (close to sensor 2) were measured.

In addition, a video recording was conducted along with the magnetometer measurements. From the video recording the time when the vehicles were in front of the sensor (passing time) as well as their driving directions were manually determined in order to establish a ground truth. From this ground truth, a 1.5 s sequence of the magnetometer signal centered around the true passing time for each vehicle was extracted from the raw measurement data. These data sequences were then used in order to evaluate the direction classification algorithms as presented in the previous sections. In this way, the detection problem does not affect the comparison of the two classification algorithms.

Note that the two sensor setup is only used for increasing the amount of data and for evaluating the presented fusion framework. The presented classifier in its simplest form still only needs one sensor for classifying the driving direction.

### B. Results

As indicated above, the ground truth data was used in order to measure the performance of the two classifiers. For each passage, a 1.5 s long time window from the magnetometer signal was extracted and the driving direction was determined. For the likelihood classifier introduced in Section IV, the chosen parametrization was as follows:

$$\begin{aligned} \mathcal{H}_L : \quad \theta_1^* &= [25 \text{ m/s} \quad 3.5 \text{ m}] \\ \mathcal{H}_R : \quad \theta_2^* &= [-25 \text{ m/s} \quad 6.5 \text{ m}] \end{aligned}$$

which corresponds to the actual road geometry and speed limit at the place where the measurements were performed. The correlation classifier was tuned using the tuning algorithm described in Section III-C using the first measurement set (71 vehicles close to sensor 1 and 85 vehicles close to sensor 2) resulting in  $p = 11$ .

Finally, the two algorithms were run on the remaining two datasets and the results are shown in Table I. As it can be seen in the table, the correlation classifier performs very well. For the vehicles passing close to the sensor, only one out of the 511 vehicles is misclassified. As expected due to the lower SNR compared to vehicles passing close to the sensor, the performance is worse for the vehicles passing on the lane farther from the sensor. Here, in total 57 vehicles are wrongly classified.

Also the likelihood classifier performs well for vehicles close to the sensor but not quite as well as the cross-correlation classifier. 23 vehicles are wrongly classified for the case where the vehicles pass close to the sensor and 132 – around twice as many as for the correlation classifier – are wrongly classified for vehicles passing far from the sensor.

The incorrect classifications can be divided into the following three main sources of error which affect the two classifiers differently:

- 1) If multiple vehicles are present at the same time, the single target assumption made in the dipole model (1a) is violated. This situation occurs if two vehicles heading in different directions are passing each other in front of the sensors or if a train of vehicles is passing the sensors with short distance between the vehicles.
- 2) For large vehicles, the dipole model (1a) is also violated. This is because the dipole model does not assume any geometrical extent of the vehicle. Furthermore, for very big vehicles this will give rise to a saturated signal as displayed in Figure 2, which clearly violates the assumption in (1a).
- 3) Finally, if the SNR is poor the classification result will become worse as also concluded in the simulations in Section V.

In order to quantify the impact of the first source of error on the two classifiers, the available data was first split into two groups – one group where the vehicles are more than 2

TABLE I: Results of applying the driving direction classification to the measurement data. For example, 290 out of the 291 vehicles traveling south-north were classified correctly by the correlation classifier using the measurements of sensor 1.

| Direction                       | # Vehicles | Sensor 1    |            | Sensor 2    |            | Fusion      |            |
|---------------------------------|------------|-------------|------------|-------------|------------|-------------|------------|
|                                 |            | Correlation | Likelihood | Correlation | Likelihood | Correlation | Likelihood |
| South-North (Close to sensor 1) | 291        | 290         | 278        | 265         | 215        | 283         | 273        |
| North-South (Close to sensor 2) | 220        | 189         | 164        | 220         | 210        | 211         | 207        |
| Total                           | 511        | 479         | 442        | 485         | 425        | 494         | 480        |

TABLE II: Results of applying the driving direction classification to the measurement data. For example, 710 out of the 722 vehicles which are more than 2 seconds separated from the closest vehicle, were classified correctly by the cross-correlation based classifier.

| Distance to closest vehicle | # Vehicles | Correlation | Likelihood |
|-----------------------------|------------|-------------|------------|
| More than 2 s               | 722        | 710         | 642        |
| Less than 2 s               | 300        | 254         | 225        |

seconds separated from each other and one group where the vehicles are less than 2 seconds separated from each other. Each vehicle gives rise to two data sets since we have two sensors. Therefore, in total  $511 \times 2 = 1,022$  data sets are considered here. The classification results for these two groups are presented in Table II. According to these results both classifiers are degraded when the single target assumption is violated, however the likelihood test suffers more from this violation than the correlation classifier does.

Furthermore, when the distance between the vehicles is larger than 2 seconds and the single target assumption applies, the correlation classifier also performs better than the likelihood test. In order to further investigate these differences, this group of data has been grouped into 8 classes of different SNR levels, each class having an interval of 5 dB. In Figure 9 the classification result for these groups is presented as a function of the SNR, which was computed as

$$\begin{aligned} \text{SNR} &= 10 \log_{10} \left( \frac{\frac{1}{N} \sum_{k=1}^N \|\bar{\mathbf{h}}_k\|^2}{\sigma^2} \right) \text{ dB} \\ &\approx 10 \log_{10} \left( \frac{\frac{1}{N} \sum_{k=1}^N \|\mathbf{y}_k\|^2 - \sigma^2}{\sigma^2} \right) \text{ dB}, \end{aligned}$$

where  $\sigma^2$  is the variance of the measurement noise.

According to Figure 9 the correlation classifier performs better than the likelihood test for all present SNR levels. Also notice that the correlation classifier has a zero error rate where  $\text{SNR} \gtrsim 10$  dB, which the likelihood does not. As a matter of fact, the likelihood test performs even worse for  $\text{SNR} \gtrsim 25$  dB corresponding to large vehicles close to the sensor. As explained above, this is because large vehicles violate the point target assumption, which the dipole model (1a) is based upon. However, it is important to note that the performance of the correlation classifier is not affected by these model violations since it does not use that model explicitly, but only one property of it and thus, the second source of error only affects the likelihood classifier. According to the experimental results, this property is still valid even in

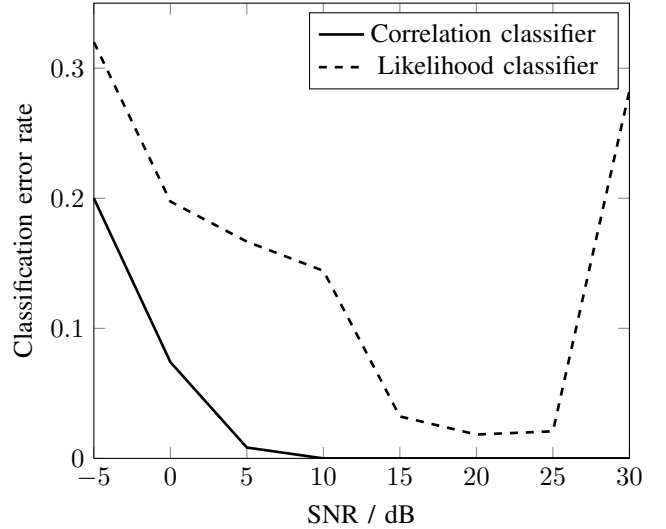


Fig. 9: Classification performance compared with the likelihood test as a function of SNR.

a near field scenario since we still have excellent classification results, even for high SNRs.

It is instructive to compare the classification performance evaluated on real data in Figure 9 with the performance on simulated data in Figure 7. When evaluating the classification on simulated data, the likelihood classifier performs better than the correlation classifier does. However, for real data the likelihood classifier is heavily disadvantaged and it performs worse than the correlation classifier due to the reasons discussed above.

Finally, the last two columns in Table I show the results for sensor fusion where the classification of both sensors were taken into account. Clearly, the overall performance is improved; the total number of correct classifications including vehicles on both, the closer and farther lane, is larger than for the case when only one sensor was used. Note that it appears as if the fusion results were worse when comparing it to, for example, the results of sensor 1 and vehicles travelling in the south-north direction. However, this is a biased comparison since one preselects the cases favorable for sensor 1 by looking at the isolated results for the south-north direction. The important result is the overall performance where the fused result is better than that of the individual sensors.

### C. Discussion

As the results in the preceding sections indicate, the proposed driving direction classifier performs very well and generally outperforms the likelihood classifier. The challenges

encountered in the previous section are discussed in more detail in this section.

The first problem was related to scenarios with multiple vehicles in the scene. This violates the single target assumption, which both classifiers are based upon. The likelihood test could be extended to handle this case by modifying the model (27) to include multiple dipoles, which requires even more parameters to estimate. For the proposed correlation classifier such an extension is not straightforward since the feature has been extracted under the single target assumption and is otherwise not valid. This might be considered as a limitation of this classification method.

However, in case of multiple targets both classifiers most likely will classify according to the vehicle with the highest SNR. Since the SNR decays cubically with the distance to the magnetic source, see (1), the vehicle with the highest SNR is most likely the one closest to the sensor. However, this assumption might not always be valid. In Figure 10 a scenario with two vehicles of different sizes is presented where the classification result for the sensor closest to the smallest vehicles can not be resolved with the present algorithm, since both vehicles affect the sensor signal in the same order of magnitude. This also reveals the robustness of the proposed algorithm in cluttered environments. If the magnitude of the clutter is lower than the magnitude of signal induced by the vehicle, the driving direction will most likely be classified correctly. Due to the cubically decay in SNR as a function of distance, only clutter in the immediate vicinity of the sensor will be a potential problem. The same reasoning would be valid for the detection performance as well. However, as explained previously, the detection problem is not analyzed further in this work.

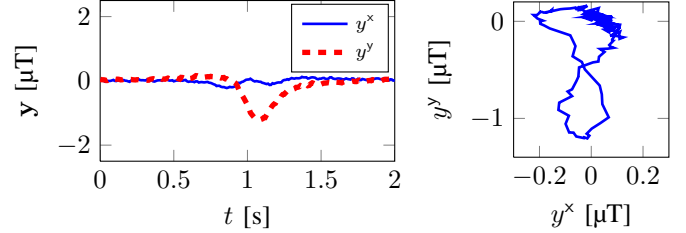
A possibility of handling multiple vehicles in the scene is to design a more elaborate post-processing strategy in a multisensor scenario. For example, instead of just fusing the individual classifications from two sensors on opposite sides of the road as proposed in Section III-D, one could inspect the uncertainty of the two classifications more thoroughly. If both sensors classify an event in the opposite driving directions with a high certainty, this could be an indication of two vehicles passing each other in front of the sensors. Further, in a more extensive multi-sensor scenario, the detection and classification from each sensor could be associated and processed to perform tracking in a multitarget-multisensor scenario [33]. In such scenarios the possibility to fuse information from other sensor types can also be analyzed.

The second problem for the correlation classifier was related to poor signal-to-noise ratios. Obviously, this is a problem that affects any algorithm to some extent; the important aspect is at which level the degradation becomes critical. Taking the experimental results in Figure 9 into account, one sees that even at an SNR as low as 0 dB, the classification error rate is below 10 %. This indicates that the correlation classifier performs relatively well even in these conditions. Nevertheless, it is apparent that the SNR can not be arbitrarily low.

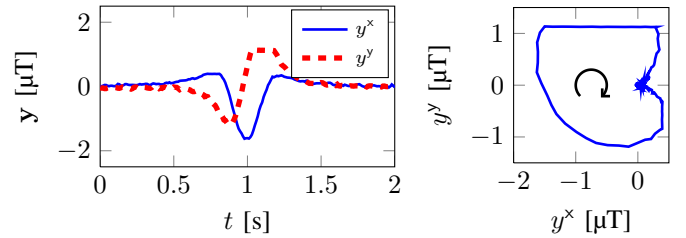
Furthermore, a disadvantage of the proposed method is the fact that large averaging windows (large  $p$ ) will cause problems at high speeds (compared to the sampling rate). If



(a) Two meeting vehicles



(b) Signal from sensor 1 (closer to the camera)



(c) Signal from sensor 2 (farther from the camera)

Fig. 10: A scenario is presented with vehicles of different sizes meeting in front of the sensors. Sensor 2 will classify according to the driving direction of the larger close-by white vehicle (c), whereas the classification result for sensor 1 is ambiguous since the the smaller close-by white vehicle and the farther larger vehicle affect the signal with the same magnitude.

a vehicle passes the sensor very fast, only few points of the loop trajectory will be measured. Averaging over these few points will have unwanted effects and it might happen that the estimation becomes wrong. However, note that the correlation classifier behaves better in general since no assumption on the vehicle trajectory and/or speed was made, compared to the likelihood test in Section IV. Clearly, the likelihood test could also be extended to take variations in these parameters into account, however, at a cost of higher complexity.

Finally, it is important to obtain an indication about whether the classification was successful or not for the two problematic cases for the correlation classifier. For the case of signals with poor SNR, this is again reflected in the classification's uncertainty (that is, in the variance of the feature  $\hat{f}_p$ ). For multiple vehicles, the only remedy is to rely on the data of multiple sensors as indicated above.

## VII. CONCLUSIONS

By using measurements from a 2-axis magnetometer, a fast and efficient method for classifying the driving direction of a vehicle has been proposed. Its properties were first analyzed theoretically and then verified by using Monte Carlo simulations before it was applied to real measurement data

from commercially available sensor. The method was also compared to a generalized likelihood ratio test and it was shown how the method can be extended to incorporate data from multiple independent sensor nodes in a sensor network.

The results show that the performance of the proposed method is very good for vehicles passing on the lane close to the sensor where more than 99 % of all vehicles were classified correctly (sample size: 511 vehicles). As it can be expected, performance degrades as the signal-to-noise ratio decreases which reflects in the fact that the classification rate for vehicles on the lane farther from the sensor dropped to approximately 89 %. Comparing this method to the generalized likelihood ratio test, it is seen that the latter is outperformed in every aspect. It is very likely that one particular reason for this behavior is the fact that one or more of the assumptions (for example, dipole model, speed, or trajectory) made to derive the GLRT are violated. The biggest difficulties arise in cases where two vehicles meet right in front of the sensor. Apparently, the vehicle generating the stronger field distortion will dominate the signal and the second vehicle is shadowed.

The proposed algorithm as presented in this work is restricted to work on single target scenarios, where possible extensions to handle multiple targets have been discussed. Future work should focus on developing these ideas further and evaluate them on real data. Furthermore, the proposed method will be implemented in a real sensor platform and the performance in a real time system will be analyzed.

#### APPENDIX A DISTRIBUTIONS

**Proposition 1.** Consider the estimator  $\hat{f}_p$  given by (11) and the measurement signal  $\mathbf{y}_k$  as in (1b) and (1c). Then, the estimator  $\hat{f}_p$  is distributed according to

$$\hat{f}_p \sim \mathcal{N}(f_p, \sigma_{\hat{f}_p}^2) \quad \text{when } N \rightarrow \infty \quad (41)$$

where  $f_p$  and  $\sigma_{\hat{f}_p}^2$  are given by (12) and (14), respectively.

*Proof.* As stated above, the mean and variance of  $\hat{f}_p$  have been derived in (12) and (14). What remains to show is that  $\hat{f}_p$  is normal distributed. Starting by expanding the original expression for  $\hat{f}_p$  yields

$$\begin{aligned} \hat{f}_p &= \frac{1}{p} \sum_{k=1}^N (y_k^x y_{k+p}^y - y_k^y y_{k+p}^x) \\ &= \frac{1}{p} \sum_{k=1}^N (h_k^x + e_k^x)(h_{k+p}^y + e_{k+p}^y) \\ &\quad - (h_k^y + e_k^y)(h_{k+p}^x + e_{k+p}^x) \\ &= \frac{1}{p} \sum_{k=1}^N h_k^x h_{k+p}^y + h_k^x e_{k+p}^y + e_k^x h_{k+p}^y + e_k^x e_{k+p}^y \\ &\quad - (h_k^y h_{k+p}^x + h_k^y e_{k+p}^x + e_k^y h_{k+p}^x + e_k^y e_{k+p}^x). \end{aligned} \quad (42)$$

Similar to (13)–(14), the sum in (42) can now be rearranged so that all the deterministic coefficients of  $e_k^\alpha$  are gathered

together:

$$\begin{aligned} \hat{f}_p &= \frac{1}{p} \sum_{k=1}^N h_k^x h_{k+p}^y - h_k^y h_{k+p}^x + (h_{k+p}^y - h_{k-p}^y) e_k^x \\ &\quad - (h_{k+p}^x - h_{k-p}^x) e_k^y + e_{k+p}^y e_k^x - e_{k+p}^x e_k^y. \end{aligned} \quad (43)$$

Distributing the sum to the individual terms finally gives

$$\begin{aligned} \hat{f}_p &= \frac{1}{p} \sum_{k=1}^N (h_k^x h_{k+p}^y - h_k^y h_{k+p}^x) \\ &\quad + \frac{1}{p} \sum_{k=1}^N (h_{k+p}^y - h_{k-p}^y) e_k^x \\ &\quad - \frac{1}{p} \sum_{k=1}^N (h_{k+p}^x - h_{k-p}^x) e_k^y \\ &\quad + \frac{1}{p} \sum_{k=1}^N e_{k+p}^y e_k^x - \frac{1}{p} \sum_{k=1}^N e_{k+p}^x e_k^y. \end{aligned} \quad (44)$$

Since  $\mathbf{e}_k \sim \mathcal{N}(0, \sigma^2 \mathbf{I})$ , the second and third term in (44) will also be normal distributed. Furthermore, since all  $e_1^x, \dots, e_N^x, e_1^y, \dots, e_N^y$  are independent, the overall variance is simply the sum of the individual variances for each variable and the distributions for these two terms are

$$\mathcal{N}\left(0, \frac{\sigma^2}{p^2} \sum_{k=1}^N (h_{k+p}^y - h_{k-p}^y)^2\right) \quad (45a)$$

and

$$\mathcal{N}\left(0, \frac{\sigma^2}{p^2} \sum_{k=1}^N (h_{k+p}^x - h_{k-p}^x)^2\right). \quad (45b)$$

The last two terms in (44) are sums of normal product distributed variables. The variance for each such variable is

$$\text{Var}(e_{k+p}^y e_k^x) = \text{Var}(e_{k+p}^y) \text{Var}(e_k^x) = \sigma^2 \sigma^2 = \sigma^4 \quad (45c)$$

and since  $e_{k+p}^y e_k^x$  are all independent and  $\sigma^4 < \infty$ , the Central Limit Theorem (see, for example, [34]) yields

$$\begin{aligned} \frac{1}{p} \sum_{k=1}^N e_{k+p}^y e_k^x &\xrightarrow{d} \mathcal{N}\left(0, \frac{N\sigma^4}{p^2}\right) \\ \frac{1}{p} \sum_{k=1}^N e_k^y e_{k+p}^x &\xrightarrow{d} \mathcal{N}\left(0, \frac{N\sigma^4}{p^2}\right) \end{aligned} \quad (45d)$$

as  $N \rightarrow \infty$ .

Consequently,  $\hat{f}_p$  is a sum of one deterministic constant and four zero mean normal distributed variables and since a sum of (possibly dependent) normal distributed variables is also normal distributed we have that

$$\hat{f}_p \sim \mathcal{N}(f_p, \sigma_{\hat{f}_p}^2) \quad \text{when } N \rightarrow \infty. \quad (46)$$

□

# APPENDIX B FUSION OF CONDITIONAL BERNOULLI RANDOM VARIABLES

Given multiple conditional PDFs  $p(x|y_1), \dots, p(x|y_J)$  it is of interest to find the joint-conditional PDF  $p(x|y_1, \dots, y_J)$ . Using Bayes rule and assuming that all  $y_1, \dots, y_J$  are statistically independent given the true  $x$  (conditional independence) yields

$$p(x|y_1, \dots, y_J) = \frac{p(y_1, \dots, y_J|x)p(x)}{p(y_1, \dots, y_J)} = p(x) \frac{\prod_{j=1}^J p(y_j|x)}{p(y_1, \dots, y_J)}$$

and using Bayes' rule on each  $p(y_j|x)$  gives

$$\begin{aligned} p(x|y_1, \dots, y_J) &= \frac{p(x)}{p(y_1, \dots, y_J)} \prod_{j=1}^J \frac{p(x|y_j)p(y_j)}{p(x)} \\ &= \frac{\prod_{j=1}^J p(y_j)}{p(x)^{J-1}p(y_1, \dots, y_J)} \prod_{j=1}^J p(x|y_j). \end{aligned}$$

Assuming the uninformative prior  $p(x) \propto 1$  yields that  $p(x|y_1, \dots, y_J) \propto \prod_{j=1}^J p(x|y_j)$  and thus

$$p(x|\mathbf{y}) = \frac{1}{\eta} \prod_{j=1}^J p(x|y_j) \quad \text{where} \quad \eta = \int_{\mathcal{D}} \prod_{j=1}^J p(x|y_j) dx, \quad (47)$$

with  $\mathbf{y} = [y_1 \ y_2 \ \dots \ y_J]^T$  and  $\eta$  being a normalization constant.

Using (47) for a set of Bernoulli random variables described by the PDF

$$p(k|p_j) = p_j^k(1-p_j)^{1-k} \quad \text{for} \quad k \in 0, 1, \quad (48)$$

where  $p_j$  is the probability of success, the fused PDF becomes

$$p(k|\mathbf{p}) = \frac{\prod_{j=1}^J p_j^k(1-p_j)^{1-k}}{\prod_{j=1}^J (1-p_j) + \prod_{j=1}^J p_j}. \quad (49)$$

## REFERENCES

- [1] N. Wahlström, R. Hostettler, F. Gustafsson, and W. Birk, "Classification of driving direction in traffic surveillance using magnetometers," *Intelligent Transportation Systems, IEEE Transactions on*, vol. 15, no. 4, pp. 1405–1418, 2014.
- [2] R. Wood, M. Palmer, L. Walter, and I. Rillie, "Driver Interaction with Temporary Traffic Management," Transport Research Laboratory, UK, Tech. Rep. RPN 1813, 2011.
- [3] Federal Highway Administration, "Traffic Monitoring Guide," U.S. Department of Transportation, Tech. Rep. FHWA-PL-13-015, Sep. 2013.
- [4] L. A. Klein, M. K. Mills, and D. R. Gibson, "Traffic detector handbook," Federal Highway Administration, Tech. Rep. Volume I, FHWA-HRT-06-108, October 2006.
- [5] W. Birk, E. Osipov, and J. Eliasson, "iRoad – cooperative road infrastructure systems for driver support," in *Proc. 16th ITS World Congress and Exhibition on Intelligent Transport Systems and Services*, Stockholm, Sweden, Sep. 2009.
- [6] S.-Y. Cheung and P. Varaiya, "Traffic surveillance by wireless sensor networks: Final report," California Path Program, Institute of Transportation Studies, University of California, Berkeley, Tech. Rep. UCB-ITS-PRR-2007-4, 2007.
- [7] J. Chinrungrueng, S. Kaewkamnerd, R. Pongthornseri, S. Dummin, U. Sunantachaiakul, S. Kittipiyakul, S. Samphanyuth, A. Intarapanich, S. Charoenkul, and P. Boonyanant, "Wireless sensor network: Application to vehicular traffic," *Advances in Wireless Sensors and Sensor Networks*, pp. 199–220, 2010.
- [8] G. Girban and M. Popa, "A glance on WSN lifetime and relevant factors for energy consumption," in *Proc. International Joint Conference on Computational Cybernetics and Technical Informatics (ICCC-CONTI)*, Timisoara, Romania, May 2010, pp. 523–528.
- [9] S. Giannechini, M. Caccamo, and C.-S. Shih, "Collaborative resource allocation in wireless sensor networks," in *Proc. 16th Euromicro Conference on Real-Time Systems, ECRTS*, Catania, Sicily, Italy, Jun. 2004, pp. 35–44.
- [10] Y. Yu and V. K. Prasanna, "Energy-balanced task allocation for collaborative processing in wireless sensor networks," *Mobile Networks and Applications*, vol. 10, pp. 115–131, 2005.
- [11] G. Hegeman, A. Tapani, and S. Hoogendoorn, "Overtaking assistant assessment using traffic simulation," *Transportation Research Part C: Emerging Technologies*, vol. 17, no. 6, pp. 617–630, 2009.
- [12] S. Moler, "Stop. You're going the wrong way!" *Public Roads*, vol. 66, no. 2, pp. 24–29, Sep./Oct. 2002.
- [13] National Highway Traffic Safety Administration, "Traffic safety facts 2010: A compilation of motor vehicle crash data from the fatality analysis reporting system and the general estimates system," 2010. [Online]. Available: <http://www-nrd.nhtsa.dot.gov/Pubs/811659.pdf>
- [14] X. Zhang and M. Forshaw, "A parallel algorithm to extract information about the motion of road traffic using image analysis," *Transportation Research Part C: Emerging Technologies*, vol. 5, no. 2, pp. 141–152, 1997.
- [15] Y. Goyat, T. Chateau, L. Malaterre, and L. Trassoudaine, "Vehicle trajectories evaluation by static video sensors," in *Proc. 9th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, 2006, Toronto, Canada, Sep. 2006, pp. 864–869.
- [16] M. Pucher, D. Schabus, P. Schallauer, Y. Lypetsky, F. Graf, H. Rainer, M. Stadtschnitzer, S. Sternig, J. Birchbauer, W. Schneider, and B. Schalko, "Multimodal highway monitoring for robust incident detection," in *Proc. 13th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, Madeira Island, Portugal, Sep. 2010, pp. 837–842.
- [17] L. E. Y. Mimbela and L. A. Klein, "Summary of vehicle detection and surveillance technologies used in intelligent transportation systems," Federal Highway Administration, Tech. Rep., 2000. [Online]. Available: <http://www.fhwa.dot.gov/ohim/tvtw/vdstits.pdf>
- [18] C. Sun, G. Arr, R. Ramachandran, and S. Ritchie, "Vehicle reidentification using multidetector fusion," *IEEE Trans. on Intelligent Transportation Systems*, vol. 5, no. 3, pp. 155–164, 2004.
- [19] K. Kwong, R. Kavalier, R. Rajagopal, and P. Varaiya, "Arterial travel time estimation based on vehicle re-identification using wireless magnetic sensors," *Transportation Research Part C: Emerging Technologies*, vol. 17, no. 6, pp. 586–606, 2009.
- [20] N. Pantazis and D. Vergados, "A survey on power control issues in wireless sensor networks," *Communications Surveys & Tutorials, IEEE*, vol. 9, no. 4, pp. 86–107, 2007.
- [21] N. Wahlström, R. Hostettler, F. Gustafsson, and W. Birk, "Rapid classification of vehicle heading direction with two-axis magnetometer," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Kyoto, Japan, Mar. 2012, pp. 3385–3388.
- [22] J. D. Jackson, *Classical Electrodynamics*, 2nd ed. John Wiley and Sons, Inc., 1975.
- [23] N. Wahlström, J. Callmer, and F. Gustafsson, "Magnetometers for tracking metallic targets," in *Proc. 13th International Conference on Information Fusion (FUSION)*, Edinburgh, Scotland, Jul. 2010.
- [24] M. Birsan, "Non-linear Kalman filters for tracking a magnetic dipole," in *Proc. of International Conference on Maritime Electromagnetics, (MARELEC)*, London, UK, 2004.
- [25] M. Rakijas, A. Saglambeni, K. K. Kohnen, and H. C. Gilbert, "Magnetic object tracking based on direct observation of magnetic sensor measurements," US Patent 6 269 324, July 31, 2001.
- [26] N. Wahlström and F. Gustafsson, "Magnetometer modeling and validation for tracking metallic targets," *IEEE Trans. on Signal Processing*, In press.
- [27] W. M. Wynn, "Detection, localization, and characterization of static magnetic dipole sources," in *Detection and Identification of Visually Obscured Targets*. Taylor and Francis, 1999, pp. 337–374.
- [28] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [29] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*. Prentice Hall, 1998.
- [30] W. L. Root, "An introduction to the theory of the detection of signals in noise," *Proceedings of the IEEE*, vol. 58, no. 5, pp. 610–623, May 1970.



- [31] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice Hall, 1993.
- [32] Honeywell, *2-Axis Magnetic Sensor Circuit HMC6042*, datasheet.
- [33] Y. Bar-Shalom and X.-R. Li, *Multitarget-Multisensor Tracking: Principles and Techniques*. Storrs, CT: YBS Publishing, 1995.
- [34] P. Billingsley, *Probability and measure*. John Wiley & Sons, 1995.



**Niklas Wahlström** (S'11) received the M.Sc. degree in applied physics and electrical engineering in 2010 and the Tech. Lic. degree in automatic control in 2013, both from Linköping University, Sweden.

Since 2010 he is working towards the Ph.D. degree at the Division of Automatic Control, Department of Electrical Engineering, Linköping University. His research interests include sensor fusion, localization and mapping, especially using magnetic sensors.



**Roland Hostettler** (S'10) received the Dipl.Ing. degree in electrical and communication engineering from Bern University of Applied Sciences, Switzerland in 2007 and the M.Sc. degree in electrical engineering from Luleå University of Technology, Sweden in 2009.

He is currently working toward the Ph.D. degree in automatic control at Luleå University of Technology, Sweden. His research interests include parameter estimation, system identification, and adaptive filtering.



**Fredrik Gustafsson** (S'91–M'93–SM'05–F'12) is professor in Sensor Informatics at Department of Electrical Engineering, Linköping University, since 2005. He received the M.Sc. degree in electrical engineering 1988 and the Ph.D. degree in Automatic Control, 1992, both from Linköping University. During 1992–1999 he held various positions in automatic control, and 1999–2005 he had a professorship in Communication Systems. His research interests are in stochastic signal processing, adaptive filtering and change detection, with applications to communication, vehicular, airborne, and audio systems. He is a co-founder of the companies NIRA Dynamics (automotive safety systems), Softube (audio effects) and SenionLab (indoor positioning systems).

He was an associate editor for IEEE Transactions of Signal Processing 2000–2006 and is currently associate editor for IEEE Transactions on Aerospace and Electronic Systems and EURASIP Journal on Applied Signal Processing. He was awarded the Arnberg prize by the Royal Swedish Academy of Science (KVA) 2004, elected member of the Royal Academy of Engineering Sciences (IVA) 2007, elevated to IEEE Fellow 2011 and awarded the Harry Rowe Mimno Award 2011 for the tutorial “Particle Filter Theory and Practice with Positioning Applications”, which was published in the AES Magazine in July 2010.



**Wolfgang Birk** (M'97) is currently Associate Professor at the Control Engineering Group at Luleå University of Technology. He holds a M.Sc. degree in Electrical Engineering from University of Saarland, Germany (1997) and a Ph.D. degree in Automatic Control from Luleå University of Technology (2002). Birk has a background in the development of Active Safety system of passenger cars, relating to collision avoidance, collision mitigation, and driver state estimation and warning. For his work relating to the development of Volvo Cars Driver Alert, he

has received the Henry Ford Technology Award (2007). Birk is co-founder of several spin-off companies from research projects and is actively involved in research and development activities relating to intelligent transport systems, control of complex interconnected systems and process control applications.