On Long-Term Statistical Dependences in Channel Gains for Fixed Wireless Links in Factories

Markus Eriksson and Tomas Olofsson

***Abstract*— The reliability and throughput in an industrial wireless sensor network can be improved by incorporating the predictions of channel gains when forming routing tables. Necessary conditions for such predictions to be useful are that statistical dependences exist between the channel gains and that those dependences extend over a long enough time to accomplish a rerouting. In this paper, we have studied such long-term dependences in channel gains for fixed wireless links in three factories. Long-term fading properties were modeled using a switched regime model, and Bayesian change point detection was used to split the channel gain measurements into segments. In this way, we translated the study of long-term dependences in channel gains into the study of dependences between fading distribution parameters describing the segments. We measured the strengths of the dependences using mutual information and found that the dependences exist in a majority of the examined links. The strongest dependence appeared between mean received power in adjacent segments, but we also found significant dependences between segment lengths. In addition to the study of statistical dependences, we present the summaries of the distribution of the fading parameters extracted from the segments, as well as the lengths of these segments.**

***Index Terms*— Radio link, Wireless sensor networks, Fading, Channel models, Mutual information, Parameter estimation.**

I. Introduction

A. Background and Motivation

THE modern process industry depends on a well functioning data traffic for monitoring and control and this traffic is generally sent over cables. However, there is a growing interest to use wireless sensor and actuator nets (WSANs) to reduce installation costs and gain flexibility when modifying the monitoring and control implementations. By using a wireless communication medium, new issues such as fading and time varying throughput in wireless communication channels must be addressed. Generally, closed loop control and discrete signaling are identified as applications with the least tolerance for latency and outage, where a few lost packets can lead to a long standstill in production [1].

The better we understand the fading characteristics of the radio links, the better are our possibilities to reliably communicate over the network, e.g., by routing traffic in a way so that packet losses and congestion are avoided as much as possible. In particular, if we can *predict* the channel gains ahead of time, then the routing tables can be updated before such problems occur.

Communication standards for WSANs, such as, WirelessHART and ZigBee, usually advocate centralized scheduling of contention free periods, or time division multiple access (TDMA) with super frames to meet latency requirements [2]–[4]. They target radios that implement the low power, low bit rate, IEEE 802.15.4 physical layer, with a packet transmission duration of approximately 5 ms. The centralized scheduling imposes constraints on the update rate of the routing tables to be on the order of seconds, since the network manager needs to collect information on topology, link quality and sampling rate from the connected devices. As a consequence, prediction horizons must also be on the order of seconds or longer to be useful.

Such long horizons are in contrast to those usually considered in the literature for mobile communication [5], [6], where they involve a few symbols, say 10-100, which corresponds to micro- to milliseconds. The need for long time horizons is further emphasised by the fact that many of the use cases for industrial WSAN may generate relatively sparse data traffic [1], which means that some links will be probed at a relatively low rate.

Needless to say, long term prediction of channel gains requires that there exist long term statistical dependencies in the channel gains. For fixed wireless links, which generally constitute the majority of links in WSANs, such dependencies may be caused by regular movements of surrounding objects that strongly influence the radio propagation in terms of multi path fading or shadowing. If the movement of such influential objects is predictable, then we have reason to believe that their effect on the fading can be predictable as well.

The main purpose of the paper is to prove that channel gains observed in factory environments are predictable on a long time horizon and we do this by showing that there exist statistical dependencies between channel gains over long horizons. Our results indicate that such dependencies do exist for a majority of the links. Note, however, that finding long term dependencies in relatively highly resolved time series is generally considered to be difficult. An often used approach to handle this problem, and the one used in this study, is to set up a hierarchical model where some parameters describe the short term dependencies and where the long term behaviour of these parameters are described by another set of parameters.

Below we outline the approach we have used to detect the statistical dependencies.

Manuscript received December 22, 2015; revised March 18, 2016; accepted April 30, 2016. Date of publication May 5, 2016; date of current version July 12, 2016. The associate editor coordinating the review of this paper and approving it for publication was D. Niyato.

The authors are with the Department of Engineering Sciences, Signals and Systems, Uppsala University, Uppsala 75020, Sweden (e-mail: markus.eriksson@signal.uu.se; tomas.olofsson@angstrom.uu.se).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCOMM.2016.2563431

B. Outline of the Method for Detecting Statistical Dependencies in Channel Gain

Our analysis of the long term dependencies is based on a hierarchical switched regime model that captures the essential behavior of the channel gains at fixed wireless links that has been reported in the literature and that we have also observed in our studies, see Section II for details. The model separates the short term- and long term dynamics of the fading by allowing the time series to be split in a number of time segments, delimited by random *change points* (CPs).

All fading values that reside in the same segment are assumed to be independent and identically distributed (i.i.d.) and this serves as a model for the short term dynamics. Fading behavior on a longer time scale is determined by the dynamics of the parameter vector which changes value at every CP. Dependencies between fading values at two different time instants may have one of the following two reasons: (A) either the time instants are close enough such that the fading values are drawn from the same distribution, or (B) they belong to different segments, with corresponding parameter vectors that are statistically dependent. Reason (A) motivates a study on the distribution of the fading parameters and on segment length, and reason (B) motivates a study on statistical dependencies between parameters describing different segments.

In this work we perform the studies motivated by (A) and (B) using data that have been collected from three different sites in the process industry, one rolling mill in Sandviken, one paper mill in Iggesund and one flotation plant in Garpenberg. The following method is used: (i) We use a CP detection (CPD) algorithm to split link data series into non overlapping segments. The CPD algorithm has been specially tailored for a regime switching model where an i.i.d gamma distribution is used to model the variations in power gain in each segment. (ii) The parameter vectors associated with each segment are estimated and, together with the estimated CPs, they form a batch of data for each link that consists of time series of parameter vectors and the corresponding segment lengths. The length of each segment is defined as the separation in time between the two CPs that confine the segment. (iii) We study statistical dependence in the time series by combining parameters into pairs, resulting in observations that are two dimensional, between which it is possible to estimate mutual information (MI). In particular, we study statistical dependence between the length of segments and the variability in received power, as well as between parameters describing adjacent segments. However, the proposed method can of course be generalized to concern more complex patterns extending over several segments. (iv) A statistical test is performed on each estimated MI value to determine, with a 95% significance level, whether or not the parameters are dependent.

Note that the identified parameters, see (ii) above, provide important information about the short term fading distributions. In particular, we obtain the Nakagami-m figure[1] from which we, via a simple parameter transformation, can obtain estimates of the K-factor and thus enable a comparison with findings in other studies on fixed wireless links in factories.

The CPD algorithm in (i) was originally developed by Fearnhead in [7], but has been modified in this work for observations generated from a gamma distribution. It uses the Viterbi algorithm to find the, a posteriori, most probable sequence of CPs conditioned on a series of observations, see Section III for details. A similar work was presented in [8], based on an on-line algorithm developed by Adams and MacKay in [9], with the intent of reacting to changes in real time. In our application an off-line algorithm is more appropriate since we are studying batches of data from which we want to infer statistical dependencies.

C. Contributions

Our main contributions and findings in the paper can be summarized as follows:

- We have modified the CPD algorithm developed by Fearnhead in [7], so that the observations are assumed to be generated by a gamma distribution, with a piecewise constant parameter vector.
- We apply the CPD algorithm to measurements of received signal strength (RSS), obtained from three different factories. Our results show that the CPs are, typically, separated in time by more than one minute, up to several hours as the longest. We also uncover a statistical dependence between K-factors and duration of segments.
- We show the existence of statistical dependencies between parameter vectors describing adjacent segments. For example, at a significance level of 95%, more than 90% of the links show a statistical dependence in mean received power between segments.
- We confirm earlier reports on high values of K for fixed wireless links compared to mobile scenarios. In fact, we note values that are slightly higher compared to the reported results from previous work, and we provide a plausible explanation based on differences in how the studies were performed.

The paper is organized as follows: In Section II we summarize earlier studies on fixed wireless links and complement with own results to motivate why the regime switching model is well suited for describing the fading process. Details of the corresponding CPD algorithm are presented in Section III and the method for detecting statistical dependencies is explained in Section IV. This is followed in Section V by a description of how the experimental data was acquired at the different sites. In Section VI we present the results and, finally, in Section VII we conclude and discuss the results.

II. Fixed Wireless Links

In this work we are concerned with fixed wireless links. Such links have been studied in applications, such as, broadband wireless access (BWA) [10]–[13], indoor/office environments [14]–[16], and factories [17]–[19]. Most studies indicate that the fading can be well described as Rician, which can be explained as follows [17], [18]: The transmitter establishes a spatial field acting as a reference value corresponding to

[1]Please note that a gamma distribution for describing power gain corresponds to a Nakagami-m distribution for describing the amplitude gain.

a stationary channel gain. Movement of reflecting objects will perturb the multipath components, causing variations from the stationary value. If the number of moving reflectors are large enough and if their contribution to the scattering is of comparable magnitude, the central limit theorem states that the random component associated with the moving reflectors becomes complex Gaussian in the phasor domain and the resulting fading distribution thus becomes Rice. See, e.g., [20] for a brief summary on basic fading models and their underlying assumptions.

The K-factor associated with the Rician distribution depend, for fixed wireless links, on how much the moving reflectors contribute to the received power in comparison to the contribution from the fixed reflectors. Thus, these K-factors are specific to the site, the relative position of the transceivers, and other factors influencing the wave propagation paths. For instance, in [11] it was shown that the K-factor depends on season, antenna height, antenna beam width, and antenna separation. In [13] it was shown for a metropolitan area that antennas in line-of-sight (LOS) and with obstructed view of surrounding streets resulted in very high K-factors. For links with antennas in non-LOS (NLOS), and at street level in close proximity to vehicles, the K-factors were much lower. In the latter case, they varied significantly with traffic conditions; the heavier traffic, the lower K-factors were observed.

Long term fading behavior for links in a rural area was studied in [21] where links obstructed by trees experienced temporal fading in excess of 10 dB through changes in wind conditions. Turbulence caused fast fading through wavering motion of the leaves, whereas the mean wind speed affected shadowing conditions by displacing the tree tops. The effect of the wind speed was not as prominent in the off-leaf season, giving rise to regular behavior in the fading process with a very long time cycle.

Similar mechanisms as described above hold also for indoor/office environments [14]. However, due to higher attenuation caused by denser partitioning from nonconductive walls and inventories, the overall indoor fading is more influenced by shadowing compared to outdoor environments.

Fixed links inside factories share fading characteristics with both the metropolitan area and the typical indoor/office environment [18]. Factories resemble metropolitan areas in the sense that attenuation is not severe. This is because factories often have large open areas where aisles are arranged in an orderly, orthogonal, intersecting fashion. Moreover, heavy machinery readily facilitate multiple paths to illuminate the receiver. On the other hand, since many factories are sectioned into localized work areas, they resemble indoor channels in the sense that perturbations of the channel due to motion often are confined to one or two main paths.

Relatively short term fading over links in four factories was examined in [19]. The links were measured for five minutes and then sorted into different groups based on the composition of personel and machinery in the vicinity of the radio nodes. From these short time series of channel gain, they showed that temporal fading is Rician, and topographies with a mix of moving personel and machinery tends to have the lowest K-factors.

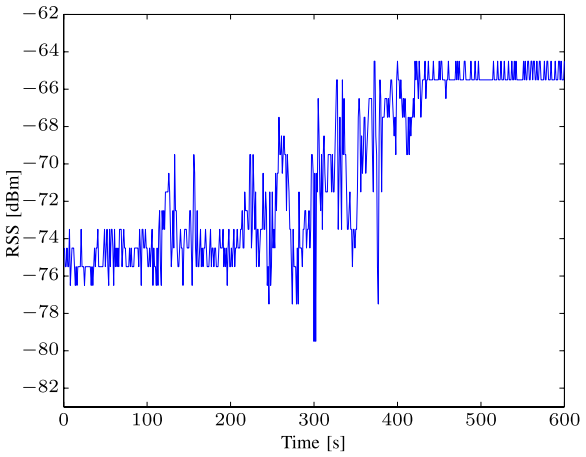

Fig. 1. RSS, sampled at 1 Hz, between fixed nodes deployed inside a flotation plant.

More long term evolution of channel statistics in factories was studied in [20] where it was observed that fading over a link can be described as switching between different states caused by major events in the vicinity of the nodes. Such events could for instance be the movement of an overhead crane or a fork lift. Movement of this type of equipment typically occur only for a relatively short time, followed by long time spans when they remain at fixed positions. When large objects are moving, causing radical changes either in the phase or magnitude of some of the multipath components that contribute significantly to the link, the distribution describing the fading before and after the movement of the object will be different. Such changes may occur both in mean level and/or dynamic range of the fading.

One example of this behaviour is given in Figure 1, where the RSS, sampled at 1 Hz, at a link inside a flotation plant is presented. Between $t = 0$ and $t \approx 210$ s, the link is nearly static. Between $t \approx 210$ s and $t \approx 410$ s, the fading fluctuation is approximately 10 dBm, and after $t \approx 410$ s, the fading goes back to a nearly static behavior, but now at a different mean level.[2] We can thus consider the time series to consist of three time segments, delimited by two CPs. According to the earlier discussion, the nearly static distributions correspond to situations with a small influence of moving reflectors and the large RSS fluctuations corresponds to a larger influence of such reflectors. In this particular example, it is quite plausible that the fluctuations is caused by one single large object close to the transceivers that begins to move at $t \approx 210$ s and stops at $t \approx 410$ s. See also Figure 3 for another example where the measured RSS changes abruptly between different distributions.

The fading dynamics within the segments is closely connected to the speed of the moving reflectors that give rise

[2]The behavior could be explained by a change in shadowing conditions due to a nearby moving object. Another plausible explanation is that the low RSS region corresponds to a scenario with very little activity and with multiple radio paths interfering destructively. Between $t \approx 210$ and $t \approx 450$ s some reflecting object, responsible for one or more of the multiple paths, is moving and it stops at a position where the contributions from the object now instead add constructively to the other paths.

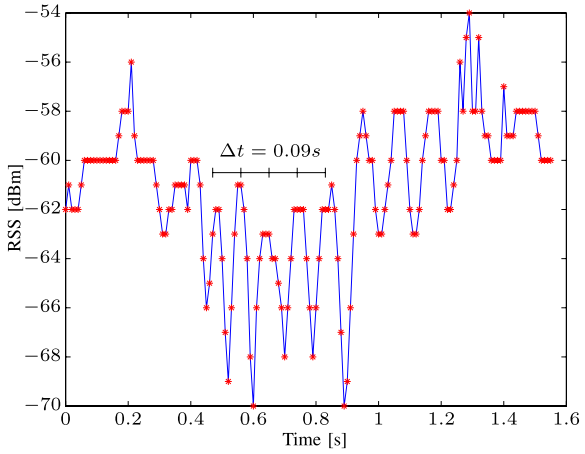

Fig. 2. Measured RSS between fixed nodes sampled at 100 Hz. The ruler marks time differences of $\Delta = 0.09$ s which coincide with the peaks in the measured RSS.

to the fading variations. If only one single nearby reflecting object is moving, then its contribution to the fading will roughly be a harmonic oscillation with a frequency that will be determined by how many wave lengths the radio path is changed per second. According to the sampling theorem, the RSS values must be observed at a rate at least twice the oscillation frequency to be faithfully monitored. Machinery, cranes, and vehicles typically move at walking speed, i.e., in the order of meters per second, and the frequencies of the oscillations associated with such objects should then be in the order of 10 Hz for a wavelength of 12.5 cm.[3] Thus, the RSS must be sampled at rates of at least 20 Hz to faithfully monitor the oscillations.

To examine the short term fading dynamics and confirm the above reasoning, we conducted a separate experiment where the RSS values were sampled at 100 Hz. During the measurement, one overhead crane in the vicinity of the transceivers was moving sporadically, giving rise to sporadic oscillations in the RSS time series. One example is shown in Figure 2 where the fading suddenly experiences oscillations with peaks that are separated in time by approximately 0.09 s corresponding to a frequency of approximately 11 Hz.

We realize from the above example that, to capture the fading dynamics on a short time scale, covering for instance the oscillative behavior caused by large reflecting objects moving at walking speeds, we must observe the RSS values at a relatively high sampling rate, say at least 20 Hz. If our sampling rate is below this, then the oscillations between $t \approx 0.4$ and $t \approx 1$ s in Figure 2 would be indistinguishable and instead appear as random fluctuations. Therefore, at low sampling rates it is more practical to assume that the observations are i.i.d. on the short time scale.[4] This assumption simplifies the mathematical treatment of the statistical model considerably.

In summary, the fixed wireless links in factories can be well described as Rician on a short time scale but on a longer time scale we must also take into account that the parameters of the distribution will change, often quite abruptly and unexpectedly. Furthermore, short term dependencies are difficult to capture unless the RSS values are observed at relatively high sampling rates. If these high rates are not met, an i.i.d. model is more relevant. We therefore propose a regime switching model to describe the fading on these links, where the random CPs models the abrupt and unexpected switching behavior and the assumption of i.i.d samples within each time segment stems from the difficulties of describing the short time fading dynamics using undersampled RSS data.

Note that such a regime switching model can, with proper choices of the transition probabilities, be written as a special case of a hidden Markov model (HMM). HMMs have been used to model channel fading over links with mobile transceivers, for instance in [22]. See [23] for a survey on Markov modeling of channel fading and its applications.

As outlined in Section I-B, this regime switching model is further used to segment the RSS measurements using a CPD algorithm.

Note that the Rice distribution can be approximated by the Nakagami distribution with high accuracy [14]. As was shown in [24], the receiver will experience Nakagami distributed fading when the received signal consists of a sum of i.i.d. Rayleigh-fading signals, which resembles the assumptions that led to the Rice distribution. Since the Nakagami distribution results in significantly simpler mathematical treatment in the CPD algorithm, we have chosen in this work to model channel gains as Nakagami distributed instead of Rician. This corresponds to power gains that are gamma distributed.

Note also that parameters from the Nakagami distribution can be transformed into the corresponding Rice parameters, see (28) below. This allows our results to be compared with those from studies that have fitted Rician distributions to empirical measurements.

III. Change Point Detection

In this section we describe the CPD algorithm that was originally developed by Fearnhead [7], with the modification that observations are assumed to be generated by a gamma distribution.

We treat CPD as a longest path problem, where the cost function is determined by the marginal likelihood and a prior assumption about the rate of CPs. The problem is solved with the Viterbi algorithm and the solution consists of the maximum a posteriori (MAP) set of CPs.

The section starts with an introduction to the data generation model and we show, among other things, how to compute the marginal likelihood for an arbitrary sequence of observations.

A. Data Generation Model

Consider a sequence of RSS observations, $r_{1:T} \triangleq \{r_1, r_2, \ldots, r_T\}$, where we use the colon notation for defining which indices that belong to the sequence. A segmentation of $r_{1:T}$ is defined by the set of CPs, $\tau_1, \ldots, \tau_{M+1}$, where we define $\tau_1 = 0$ and $\tau_{M+1} = T$. We then have M segments where the observations belonging to the mth segment are $r_{(\tau_m+1):\tau_{m+1}}$.

[3]The carrier frequency was approximately 2.4 GHz in our experiments.

[4]The maximum sampling rate used in our experiments was 8 Hz, see Section V.

We assume the observations within a segment to be conditionally i.i.d. gamma defined by the parameter vector $\theta \triangleq (\alpha, \bar{r})$, where α and $\bar{r}$ denote the shape parameter, and the mean power, respectively, having the pdf,[5]

$$f(r|\theta) = \frac{\alpha^\alpha}{\Gamma(\alpha)\bar{r}^\alpha} r^{\alpha-1} \exp\left(-\frac{\alpha r}{\bar{r}}\right). \quad (1)$$

For some general segment $r_{t:s}$, the pdf in (1) yields the likelihood

$$f(r_{t:s}|\theta) = \left(\frac{\alpha^\alpha}{\Gamma(\alpha)\bar{r}^\alpha}\right)^{s-t+1} P^{\alpha-1} \exp\left(-\frac{\alpha}{\bar{r}} S\right), \quad (2)$$

where

$$P = \prod_{i=t}^{s} r_i \quad \text{and} \quad S = \sum_{i=t}^{s} r_i. \quad (3)$$

ML parameter estimates of α and $\bar{r}$ that are considered in Section V are obtained as

$$\hat{\theta} = \arg\max_{\theta} f(r_{t:s}|\theta). \quad (4)$$

The parameter vector θ is assumed to be piecewise constant between the CPs $\tau_1, ..., \tau_{M+1}$. Let θ_m denote the parameters associated with the mth segment and assume that θ_m and θ_n are, a priori, independent for $m \neq n$. The prior distribution for these parameters is for practical reasons here chosen to be conjugate with respect to (2) [25]. This means that the posterior density function and the prior will share the same functional shape and this choice allows for an efficient algorithm implementation for CPD that utilizes simple updating rules of sums and products of the type shown in (3).

The conjugate prior over θ with respect to (2) is given by [25],

$$f(\theta|a, b, c, d) = \frac{1}{Z} \frac{a^{\alpha-1}}{\Gamma(\alpha)^b} \left(\frac{\alpha}{\bar{r}}\right)^{\alpha c} \exp\left(-\frac{\alpha d}{\bar{r}}\right), \quad (5)$$

where Z is a normalization constant that depends on the real, positive, hyperparameters a, b, c and d. To avoid lengthy notation below, the conditioning on these parameters is not written out explicitly.

The posterior probability distribution of θ is proportional to the product of the likelihood function (2) and the prior distribution (5),

$$\begin{aligned} f(\theta|r_{t:s}) &\propto f(r_{t:s}|\theta) f(\theta) \\ &= \frac{1}{Z} \frac{a'^{\alpha-1}}{\Gamma(\alpha)^{b'}} \left(\frac{\alpha}{\bar{r}}\right)^{\alpha c'} \exp\left(-\frac{\alpha d'}{\bar{r}}\right), \end{aligned} \quad (6)$$

with a', b', c', d' defined as,

$$\begin{aligned} a' &= aP \quad b' = b + s - t + 1, \\ c' &= c + s - t + 1 \quad d' = d + S. \end{aligned} \quad (7)$$

We see that (6) and (5) share the same functional form in θ. The marginal likelihood, $\mathcal{L}(t, s)$, is given by,

$$\mathcal{L}(t, s) = \int_{\alpha=0.5}^{\infty} \int_{\bar{r}=0}^{\infty} f(r_{t:s}, \theta) d\bar{r} d\alpha, \quad (8)$$

[5]Several different parameterizations exist in the literature for the gamma distribution and with the one chosen in this work, α corresponds directly to the parameter m in the Nakagami-m distribution.

where integration from $\alpha = 0.5$ is due to the fact that this is the smallest shape parameter that is allowed in the corresponding Nakagami distribution [24]. The inner integration over $\bar{r}$ in (8) is valid for $\alpha c' - 1 > 0$, $\alpha d' > 0$ and from [26, eq. (3.381:4)] we get,

$$\mathcal{L}(t, s) = \frac{1}{Z} \int_{0.5}^{\infty} g(\alpha) d\alpha, \quad (9)$$

where $g(\alpha)$ is given by,

$$g(\alpha) = \frac{\alpha a'^{\alpha-1}}{\Gamma(\alpha)^{b'}} \frac{\Gamma(\alpha c' - 1)}{d'^{\alpha c'-1}}, \quad (10)$$

and a', b', c' and d' are defined in (7).

The integral over α in (9) has no analytical solution and it is evaluated numerically using the Laplace approximation, as described in Appendix D.

The CPs $\tau_{2:M}$ are, a priori, assumed to be uniformly distributed over the sequence of observations. This noninformative prior assigns equal probability that r_t is a CP, for all $t = 1, \ldots, T$,

$$Pr(\text{CP at } t) = \lambda, \quad (11)$$

where λ is set to some small number that reflects the expected rate of CPs,

$$\lambda = \frac{\text{Expected number of CPs}}{\text{Total number of observations}}. \quad (12)$$

This leads to a geometric distribution of segment length,

$$Pr(\tau_{m+1}|\tau_m) = \lambda(1-\lambda)^{\tau_{m+1}-(\tau_m+1)}. \quad (13)$$

As will be evident from the results presented in section VI-A, a specific value of λ results in segments with a wide range of different length. Thus, the resulting segmentation is relatively insensitive to the choice of λ.

B. MAP Change Point Detection

The MAP segmentation is defined as the set of CPs $\tau_2^*, \ldots, \tau_M^*$ that fulfill,

$$\{M, \tau_2^*, \ldots, \tau_M^*\} = \arg\max_{M, \tau_2, \ldots, \tau_M} Pr(\tau_2, \ldots, \tau_M|r_{1:T}). \quad (14)$$

Note that we also maximize over the number of segments, M. Thus, the optimization is not performed over a predetermined number of CPs.

This segmentation is obtained using the Viterbi algorithm which relies on an auxiliary variable, $Q^*(t)$, that is defined as [7]

$$\begin{aligned} Q^*(t) = Pr(&r_{t:T}|\text{CP at } t-1, \\ &\text{MAP estimate for } t < \tau_m^*, \ldots, \tau_M^*), \end{aligned} \quad (15)$$

for $t = 2, \ldots, T$ and,

$$Q^*(1) = Pr(r_{1:T}|\text{MAP estimate for } \tau_2^*, \ldots, \tau_M^*). \quad (16)$$

$Q^*(1)$ can be interpreted as the likelihood that the model described in Section III-A generated the sequence $r_{1:T}$, conditioned on the most likely sequence of CPs. Using the above stated definitions, (15) can be calculated recursively as [7],

$$Q^*(t) = \max_{s>t} \{\mathcal{L}(t, s) Q^*(s+1) \lambda^{I(s \neq T)} (1-\lambda)^{s-t}\}, \quad (17)$$

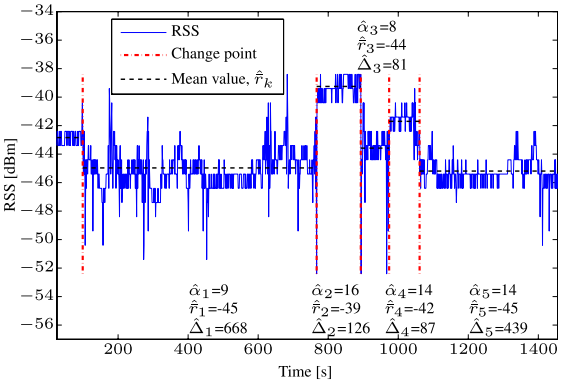

Fig. 3. Measured RSS from a link that shows a regime switching behavior. The output from the CPD algorithm and parameter estimation is included in the figure.

where $t \leq s \leq T$ and $Q^*(T+1) = 1$. The indicator function $I(s \neq T) = 1$ when $s \neq T$ and 0 otherwise. Let $s^*(t)$ be the value that maximises (17),

$$s^*(t) = \arg\max_{s>t} \{\mathcal{L}(t,s) Q^*(s+1) \lambda^{I(s \neq T)} (1-\lambda)^{s-t}\}. \quad (18)$$

When $Q^*(t)$ has been calculated for $t = 1, \ldots, T$, the MAP CPs $\tau^*_{2:M}$ in (14) are calculated with the following recursion:

1) Set $\tau^*_1 = 0, m = 1$,
2) While $\tau^*_m < T$; set $\tau^*_{m+1} = s^*(\tau^*_m + 1)$; $m = m + 1$.

C. Parameter Estimation

With a segmented sequence of RSS observations, the ML estimate, $\hat{\theta}$, for the mth segment $r_{\tau^*_m+1:\tau^*_{m+1}}$ is obtained through (4), where $f(r_{\tau^*_m+1:\tau^*_{m+1}}|\theta)$ is given in (2). The length of the segment will be denoted $\hat{\Delta}_m$ and is calculated as,

$$\hat{\Delta}_m = \tau^*_{m+1} - \tau^*_m. \quad (19)$$

The algorithm is illustrated in Figure 3 where the red lines mark the estimated CPs that separates a sequence of RSS measurements into five segments, each with a parameter triplet.

In Section V-B, the CPD algorithm will be applied to RSS measurements from L different links. For future reference, let M_l denote the number of segments in the lth link, and let φ^l_m denote the parameter triplet associated with the mth segment,

$$\varphi^l_m = \{\hat{\alpha}^l_m, \hat{\bar{r}}^l_m, \hat{\Delta}^l_m\}, \quad (20)$$

where $m = 1, \ldots, M_l$. The total set of parameter triplets from the lth link are referred to as Ψ_l,

$$\Psi_l = \{\varphi^l_1, \ldots, \varphi^l_{M_l}\}. \quad (21)$$

D. Choice of Hyperparameters

The processing in Section V-B, was conducted with hyperparameters set to values indicated by table I. By inspection of the update rules (7) for the posterior parameters, we note that the prior parameters a, b, c and d, can be viewed as the sufficient statistics for a number of fictitious samples, where b and c sets the number of fictitious samples, and a, d correspond to their product and sum, respectively. In applications where prior information is available, this can be incorporated by setting a, d to reflect the behavior of the signal, and b, c can be considered as weights. When b and c are large in comparison to the number of observed data points, the resulting segmentation will be strongly influenced by the prior.

TABLE I
PARAMETER VALUES USED FOR DEFINING THE PRIOR DISTRIBUTIONS IN (5) AND (13)

Parameter	Value
a	1
b	2
c	2
d	7
λ	0.001

The conjugate prior used in this work was difficult to tune to the vast span of possible $\bar{r}$ for the collection of links. To handle this we normalised the received power to unity on each link.[6] This simplified the choice of a common prior that could be used for all links.

IV. Detecting Statistical Dependencies Using Mutual Information

The main goal of our study is to assess whether or not the RSS values show statistical dependencies over time, which in the switched regime model corresponds to dependencies between the parameters describing the segments. Since the priors used do not induce statistical dependencies, any such dependencies that are observed in the estimated parameters associated with the extracted segments, must have been induced by data.

MI provides means to measure general statistical dependencies between variables and the authors of [27] developed a framework for nonparametric estimation of MI that avoids assumptions about the underlying distribution. These estimated MI values can subsequently be used in a hypothesis test to determine whether the parameters are dependent or not. The approach in [27] is summarized below, followed by a description of the hypothesis tests and how to choose appropriate thresholds in the tests.

A. Computation of MI

For general continuous random variables, X and Y, taking on values x and y respectively, their MI [28] is defined as,

$$I(X,Y) = \int_Y \int_X p(x,y) \log \frac{p(x,y)}{p(x)p(y)} dx dy, \quad (22)$$

where integration is performed over the support of X and Y. For independent variables, we have $p(x,y) = p(x)p(y)$ and the integral becomes zero. For dependent variables, (22) evaluates to some positive value corresponding to the amount of information that Y holds about X or vice versa.

[6]Note that a scaled gamma distributed variable is also gamma distributed.

Throughout the rest of this work, log, denotes the natural logarithm, thus, the unit for I is nat.

The unknown distributions of X and Y, are assumed to generate M two dimensional observations, (x_i, y_i), for $i = 1, \ldots, M$. The first step in estimating MI from the samples $(x_{1:M}, y_{1:M})$, is to construct a kernel estimate of $p(x, y)$,

$$\hat{p}(x, y) = \frac{1}{M} \sum_{i=1}^{M} \frac{1}{h^2} G\left(\frac{1}{h} \|(x - x_i, y - y_i)\|_2\right), \quad (23)$$

where G is the Gaussian function,

$$G(v) = \frac{1}{\sqrt{2\pi}} e^{\frac{-v^2}{2}}, \quad (24)$$

h is the scalar kernel width, and $\|\cdot\|_2$ denotes the 2-norm. From $\hat{p}(x, y)$, estimates $\hat{I}(x_{1:M}, y_{1:M})$ can be computed as,

$$\hat{I}(x_{1:M}, y_{1:M}) = \frac{1}{M} \sum_{i=1}^{M} \log\left(\frac{\hat{p}(x_i, y_i)}{\hat{p}(x_i)\hat{p}(y_i)}\right), \quad (25)$$

where $\hat{p}(x)$ and $\hat{p}(y)$ are estimates of marginal densities obtained as

$$\hat{p}(x) = \sum_{i=1}^{M} \hat{p}(x, y_i) \quad \text{and} \quad \hat{p}(y) = \sum_{i=1}^{M} \hat{p}(x_i, y). \quad (26)$$

In addition to the above steps of obtaining $\hat{I}(x_{1:M}, y_{1:M})$, we make use of a preprocessing step where the samples $(x_{1:M}, y_{1:M})$ are copula transformed [29], leaving them uniformly distributed on the interval $[0, 1]$, which simplifies the choice of kernel width h.[7]

B. Test for Statistical Dependencies

The estimates $\hat{I}(x_{1:M}, y_{1:M})$ are always non-negative and to state that a dependency has been detected we require the estimates to deviate significantly from zero. Thus, we set up a hypothesis test where the null hypothesis, H_0, states that the parameters are independent, and, accordingly, hypothesis H_1 states that they are dependent. The test consists in comparing $\hat{I}(x_{1:M}, y_{1:M})$ to a threshold, I_0^M, and it rejects H_0 if $\hat{I}(x_{1:M}, y_{1:M}) > I_0^M$, where I_0^M is chosen to give a rate of type I errors of $\rho = 5$ %, i.e.,

$$Pr(\hat{I}(x_{1:M}, y_{1:M}) > I_0^M | H_0) = 0.05, \quad (27)$$

We will in the following refer to I_0^M as the *independence threshold*, and in the section below we describe how to compute it by means of Monte Carlo (MC).

C. Computation of the Independence Threshold

We study two different types of dependencies between parameters that require a somewhat different treatment when computing I_0^M. The first type is examined in Section VI-B, where x and y are replaced with parameter estimates $(\hat{\alpha}_m, \hat{\Delta}_m)$. Hence, parameters from the same segment are regarded as a two dimensional observation, from a set of such observations we can estimate MI. The second type of dependence is treated in Section VI-C, where x and y are replaced with parameter estimates from adjacent segments, $(\hat{\alpha}_m, \hat{\alpha}_{m+1})$, $(\hat{\bar{r}}_m, \hat{\bar{r}}_{m+1})$ and $(\hat{\Delta}_m, \hat{\Delta}_{m+1})$, which again form sets of two-dimensional observations from which we can estimate MI.

The reason for treating the two types differently is that CPD will introduce excess dependencies between the estimated parameters in adjacent segments, but not between parameters within the same segment. This causes an upward bias in the estimated MI-values for the second type of dependence. This bias stems from an increased risk of missed detections in CPD when the difference between parameters associated to adjacent segments is small. Thus, consecutive parameter estimates with relatively similar values will be underrepresented in the batches. We have no reason to believe that this mechanism will be present when parameters associated with the same segment are considered.

In the former case, the MC evaluation involves running the CPD on a simulated data set, generated by the model described in Section III-A, with segments having independent parameters. The latter case allows for simplifications in computing I_0^M. The MC methods for the two cases are detailed below.

1) Tests for Parameters Within the Same Segment: When considering parameters belonging to the same segment, we can find the distribution of MI estimates using an MC method as follows: (i) Generate $N_{MC} = 10,000$ sets of x and y, each of size M, from normally distributed independent X and Y. (ii) Copula transform each set and compute $\hat{I}(x_{1:M}, y_{1:M})$. (iii) Set I_0^M as the 95% percentile of the distribution of estimated MI values.

2) Tests for Parameters Between Adjacent Segments: To take into account the above mentioned bias in MI estimates, the CPD algorithm is incorporated into the MC method, as illustrated in Figure 4. The algorithm executes as follows: (i) generate $M_{sim} > M$ independent θ vectors[8] from the prior distribution (5) and M_{sim} independent Δ values from (13). (ii) Use these parameters to simulate M_{sim} segments and concatenate these segments into one vector, $r_{1:T}$. (iii)) Apply the CPD algorithm to $r_{1:T}$ using the framework in Section III-B and estimate the parameters corresponding to each segment. (iv) Select the M first parameter triplets, $\hat{\varphi}_{1:M}$, and for each parameter, estimate MI with a time shifted version of itself, i.e., $\hat{I}(\hat{\alpha}_{1:M-1}, \hat{\alpha}_{2:M})$, $\hat{I}(\hat{\bar{r}}_{1:M-1}, \hat{\bar{r}}_{2:M})$, and $\hat{I}(\hat{\Delta}_{1:M-1}, \hat{\Delta}_{2:M})$. Repeat steps (i)-(iv) to obtain $N_{MC} = 300$ estimates of MI for each parameter. Set I_0^M as the 95% percentile of the distribution of estimated MI values.

The computed I_0^M resulted in slightly different values for the three different parameters, $\hat{\alpha}$, $\hat{\bar{r}}$ or $\hat{\Delta}$ and the threshold that corresponded to the parameter that was currently being tested was used.

Note that the approach for obtaining the threshold for parameters in adjacent segments is computationally expensive and we therefore limited the computations to sample sizes

[7] Since MI is invariant under monotonic transformations of x and y, this does not affect estimates $\hat{I}((x_{1:M}, y_{1:M}))$ [29].

[8] Since some true CPs may be missed in the detection we need to simulate a larger number of segments than is actually required in the batch. In our work, choosing $M_{sim} = 1.3M$ was found to yield a sufficient number of segments.

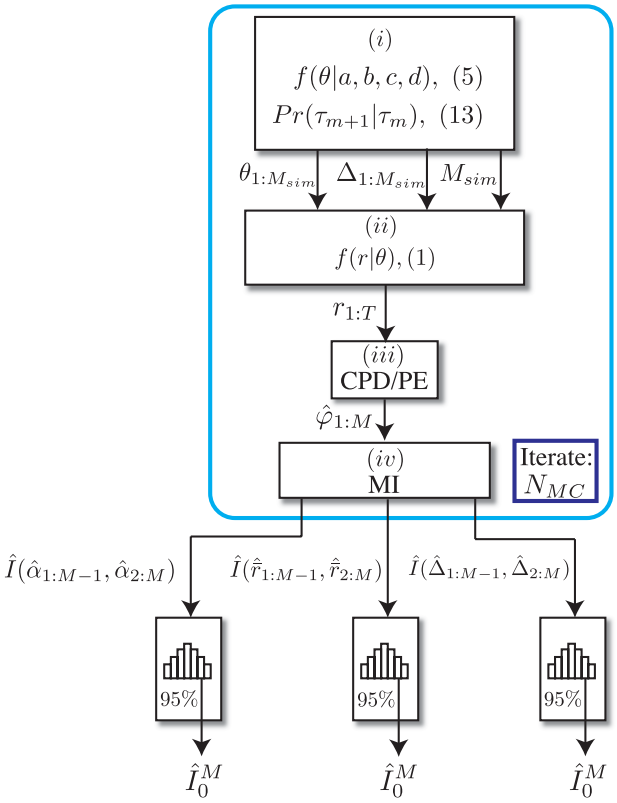

Fig. 4. Monte Carlo method for computing the I_0^M when testing for statistical dependence between parameters in adjacent segments.

of $M = \{50, 100, 400, 600\}$ as well as a smaller N_{MC} to find I_0^M. When the hypothesis tests were applied to a time series of consecutive parameter estimates of arbitrary size M_l, the independence threshold with the closest, smaller M in the above list was used.

V. Experiments

Measurements were conducted at three fully equipped and operational factories, one rolling mill, one paper mill, and one flotation plant. The deployment areas typically included heavy machines, overhead cranes, and a large amount of metal objects, all acting as reflectors. This is illustrated in Figure 5, showing the deployment area at the rolling mill, where nodes were placed on both sides of the production line.

A. Measurement Setup

Measurements were conducted with a net of Zolertia Z1 nodes which use the IEEE 802.15.4 compliant CC2420 transceiver, operating at the 2.4 GHz band. They were deployed in an area of approximately 1000 m^2 with the majority of nodes in a close proximity to machines to mimic a realistic wireless control scenario. An example is shown in Figure 6 where two of the Z1 nodes are mounted on a paper machine. The distances between nodes in the deployment areas varied between three and approximately 50 meters. Furthermore, some links were in LOS most of the time, whereas others were close to disconnected, e.g., due to strong shadowing. This resulted

Fig. 5. Overview of the deployment area of the nodes at the rolling mill in Sandviken.

Fig. 6. Two Zolertia Z1 nodes deployed at the paper mill.

in a vast span in the links' measured RSS. Measurements were gathered in two different campaigns summarized below.

Campaign I: The first campaign involved transmitting packets every second[9] using a National Instruments signal generator. The nodes received the packets and stored the corresponding RSS-values on internal memory cards. The measurements at the rolling mill and the paper mill were gathered over periods of approximately 20 hours, using 24 and 34 nodes, respectively. This resulted in 7×10^4 RSS samples per link, with an amplitude resolution of 1 dBm. Measurements at the flotation plant were conducted on two occasions of approximately 20 hours each, using 23 nodes. Altogether, Campaign I resulted in RSS measurements from 104 links. The packet loss at different links is presented in the middle column in Table II.

Campaign II: Campaign II involved a more extensive measurement conducted at the rolling mill. Eighteen nodes were used and transmissions were done in a round-robin fashion where all pairwise links were monitored during each round. This resulted in 153 possible unique links. Each round was completed in 0.125 seconds[10] during 90 consecutive hours,

[9]The RSS observation rate in Campaign I was thus 1 Hz.
[10]The RSS observation rate in Campaign II was thus 8 Hz.

TABLE II

THE PACKET LOSS FOR THE TWO CAMPAIGNS. THE MIDDLE AND RIGHT COLUMNS OF THE TABLE SHOWS THE NUMBER OF LINKS EXPERIENCING A PACKET LOSS WITHIN THE INTERVALS SPECIFIED IN THE LEFT COLUMN

	Number of links	
Packet loss (%)	Campaign *I*	Campaign *II*
[0 5]	66	52
[5 10]	7	23
[10 15]	3	15
[15 20]	9	14
[20 25]	5	10
[25 30]	1	9
[30 55]	13	7
L	104	130

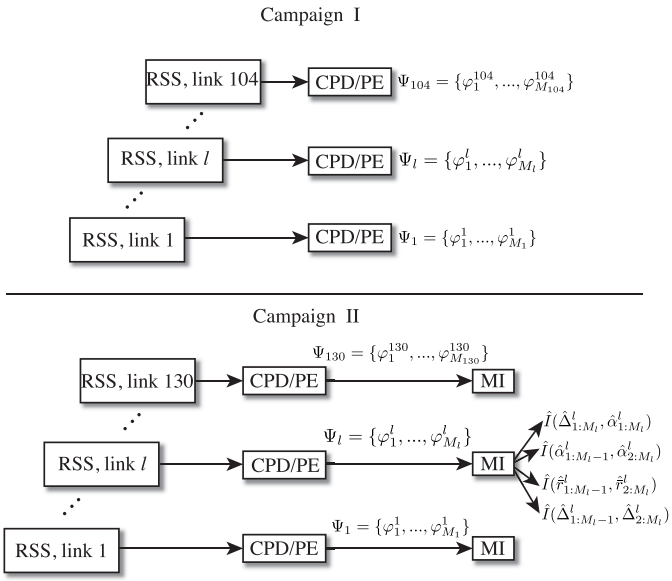

Fig. 7. Overview of the data processing.

resulting in approximately 2.6×10^6 RSS samples per link and these samples were stored at logging units.

Out of the 153 possible links, 23 proved to be disconnected, thus resulting in 130 links containing meaningful link RSS data. The rightmost column in Table II summarizes the packet loss at the different links.

B. Processing

The RSS measurements from Campaigns I and II were processed as illustrated in Figure 7. CPD was applied to each of the L links in the campaign followed by parameter estimation (PE).[11] This resulted in L batches, $\Psi_1, ..., \Psi_L$, see (21), consisting of parameter triplets with φ_m^l defined in (20). The batches from Campaign I were used solely to overview the distributions of fading parameters and segment lengths. The number of segments obtained from Campaign I was found to be insufficient for reliably estimating MI between the parameters. These were instead computed using batches obtained from the more extensive Campaign II.

[11] $L = 104$ in Campaign I and $L = 130$ in Campaign II.

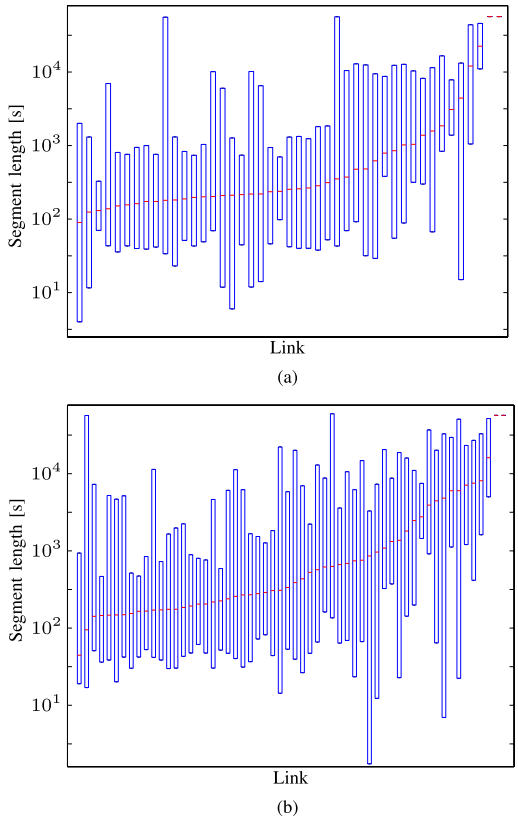

Fig. 8. The distribution of segment lengths for each link. The red line in the middle of each box indicates the median and the blue edges represent the 10th and 90th percentiles. The links have been sorted by ascending median. (a) Links in LOS. (b) Links in NLOS.

VI. RESULTS

In this section we present distributions of the parameters estimated using data from Campaign I and results from the dependency test using data from Campaign II. The parameter estimation summaries are given in Section VI-A, and Sections VI-B and VI-C concern the results from the dependency tests. In Section VI-B, the dependencies between parameters belonging to the same segments are presented, whereas VI-C presents results on dependencies between parameters in adjacent segments.

A. Distribution of Estimated Parameters

1) Segment Lengths: The distribution of segment lengths are presented in logarithmic scale in Figure 8 using box plots. Links in LOS and NLOS are presented separately and to simplify visual inspection, the links have been sorted by ascending median.

We see that a large amount of the segments last for 100 to 1000 seconds. However, the spread is large and the median at different links may differ by more than a factor 100. There is no evident difference between the distributions for links in LOS and NLOS. The shortest observed segments only

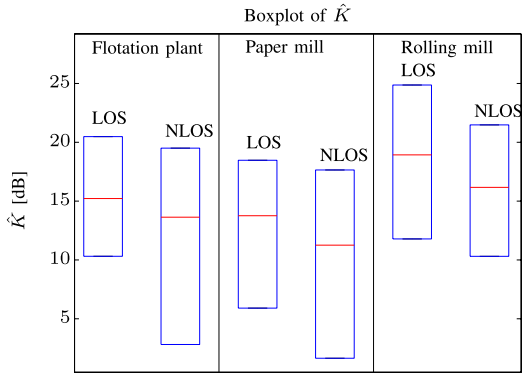

Fig. 9. The distribution of K-factors, sorted by site and presence of LOS. The red line in the middle of each box indicates the median and the blue edges represent the 10th and 90th percentiles.

lasted a few seconds whereas the longest lasted 24 hours. Two of the links were static enough to constitute only one single long segment, marked by red lines in the upper right corner of the figure.

2) K-Factors: The parameter α influences the spread of the gamma distribution, with high and low values yielding small and large spread, respectively. Since most studies on fixed wireless links have involved the Rice distribution and presented results in terms of K-factors, we have here chosen to transform the estimates of α-values to corresponding estimates of K-factors according to [30]

$$\hat{K} = \frac{\sqrt{\hat{\alpha}^2 - \hat{\alpha}}}{\hat{\alpha} - \sqrt{\hat{\alpha}^2 - \hat{\alpha}}}, \tag{28}$$

and present these values instead. The resulting distribution of estimates are shown in Figure 9. To further facilitate a comparison with other studies, each estimate has been weighted with its corresponding segment length prior to computing the histogram that underlies the box plot.[12] The results are presented separately for the three sites and the links are sorted by LOS and NLOS.

Much as expected, links in LOS typically yield higher K-factors than those in NLOS; Links in LOS have K-factors that are approximately twice as large as for links in NLOS. The majority of the estimated K-factors are found in the interval $[5, 25]$ dB, which is slightly higher than what has been reported by previous similar studies. For instance, in [19] the K-factors were found to reside between 4 and 19 dB, with an overall average of 12 dB, and in [18], a K-factor of 10 dB adequately described variability in their measurements.

3) Mean Received Power: As discussed in Section II, the fading in the considered environments can be described as switching between different states. The distribution of the mean received power for a certain link with this regime switching behaviour then describes one aspect of the variability of the states, namely how the mean power differ between different states.

[12]More specifically, the histogram is computed based on a set of estimated K-factors where each estimate has been duplicated as many times as the corresponding segment's length. Thus, values associated with a long segment are upweighted in proportion to the segment length.

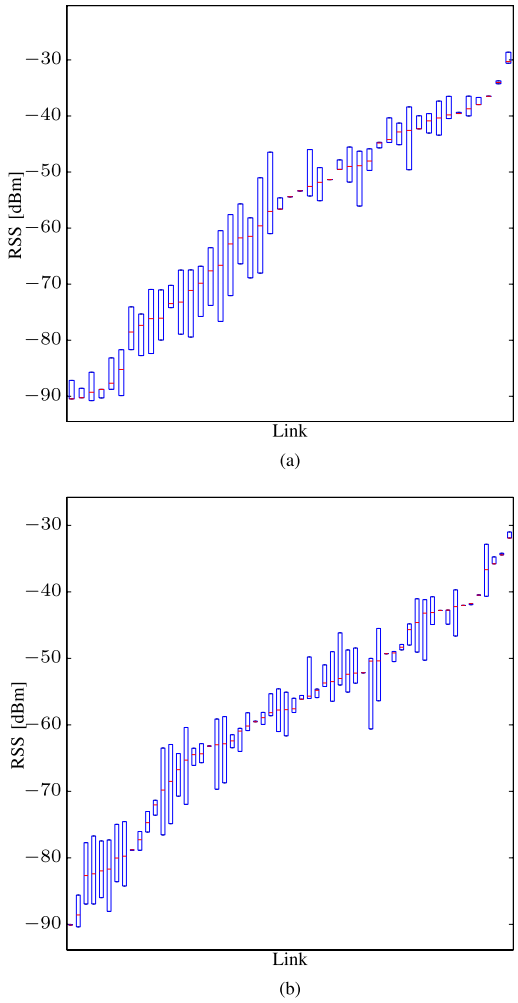

Fig. 10. The distribution of $\hat{\bar{r}}$ for each link. The red line in the middle of each box indicates the median and the blue edges represent the 10th and 90th percentiles. (a) Links in LOS. (b) Links in NLOS.

The distribution of mean received power is presented in Figure 10. The results are presented separately for links in LOS and NLOS and they have been sorted by ascending median in the same way as in Figure 8.

We observe a vast spread in the median values. These range between -90 dBm up to -30 dBm. This large difference is a direct consequence of the spread in antenna separation, in combination with the fact that some links are in LOS and some are in NLOS.

The variability of mean received power seldom exceeds 10 dBm, and several links have an almost indistinguishable variability in the parameter. Please note that the overall fading variability is always higher than the variability of *mean* received power. For example, if a link consists of only one single stationary fading state that runs over the entire measurement and with a low K-factor, then the fading variability is high, but the mean received power is constant and thus shows no variability.

We observe a slight tendency for both links in LOS and NLOS that weak links generally experience a higher variability in mean received power. One possible explanation is that the strong links often have one significant strong contributing

radio path and a moving reflector that is not responsible for this path will not be able to alter the mean power much, although it may be enough to trigger a change in fading state. A weak link has, with higher probability, several contributing radio paths of comparative strengths and a moving reflector may than influence the mean more strongly. Thus, weak links are comparatively more susceptible to perturbations due to moving reflectors.

B. Dependencies Between Parameters Belonging to the Same Segments

We have limited the study of parameter dependencies within segments to the combination α and Δ, which we believe to be the most important. If such dependencies exist, we may potentially be able to predict the next CP based on an estimate of α that can be obtained from a few RSS-samples.

The test for dependencies consists in combining the segmentation algorithm in Section III-B with the framework for estimating MI in Section IV. The different steps in the processing are outlined in the lower half of Figure 7. Before presenting the outcome of the dependency test, we illustrate the method of computing the MI in more detail by applying it to one of the links and presenting some intermediate steps.

1) Illustration of the Method for Computing MI: The method for computing MI is illustrated in Figure 11. The link considered in the example experienced nearly static fading over long periods of time, separated by shorter periods with larger variability. In the upper plot we see a section of the measured RSS, where six CPs were detected, marked with red lines. The CPD for this particular link, resulted in overall $M = 624$ segments. Parameters α and Δ were estimated for each such segment and these parameter pairs are presented as circles in a scatter plot (middle).

These M parameter pairs are used to compute a joint density estimate, $\hat{p}(\hat{\Delta}_m, \hat{\alpha}_m)$, using (23). This density is shown as a contour plot overlaying the scatter plot. Please note that the variables have been copula transformed prior to presentation. The true parameter values can be read from the lower and left axis, whereas the copula transformed values are given by the upper and right axis. Recall from Section IV-A that the copula transform yields parameters that are uniformly distributed in the interval $[0, 1]$. This implies further that independent copula transformed parameters will be uniformly distributed in a unit square and any deviation from this indicates dependent parameters.

In the plot we can see a linear trend in the data, where high $\hat{\alpha}$ correspond to large $\hat{\Delta}$, and vice versa. Thus, the plot indicates that the parameters are dependent, implying further that the MI should deviate significantly from zero.

To estimate the MI between α and Δ we use (25), with the marginal density estimates $\hat{p}(\hat{\alpha})$ and $\hat{p}(\hat{\Delta})$, computed using (26) resulting in $\hat{I}(\hat{\Delta}_{1:M}, \hat{\alpha}_{1:M}) = 0.25$ nat. For this case we have that $I_0^M = 0.11$ nat and since the estimate is well above this threshold we consider α and Δ to be dependent for this link.

We can examine this dependence in more detail in the bottom plot that shows a comparison between the conditional density $\hat{p}(\hat{\Delta}|\hat{\alpha} = 8 \text{ dB})$ and the marginal density $\hat{p}(\hat{\Delta})$.

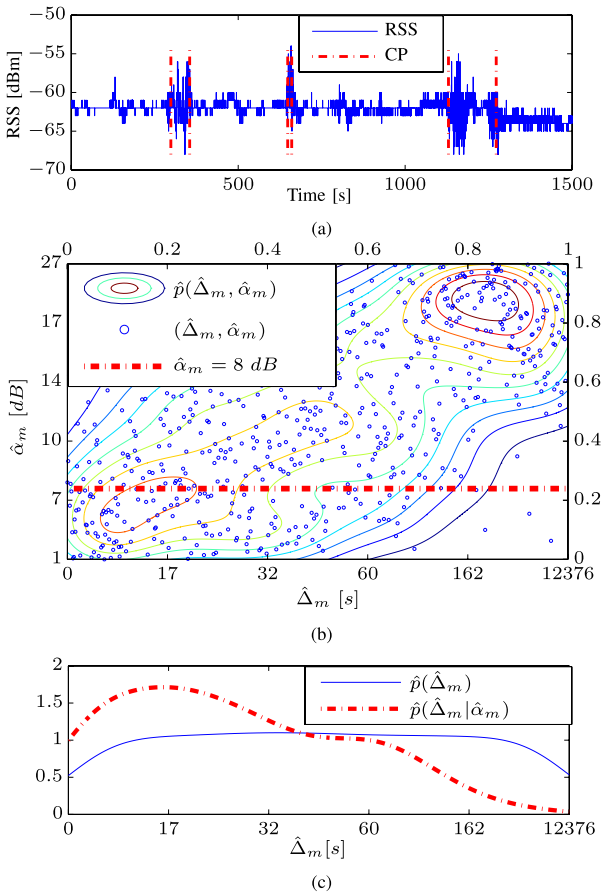

Fig. 11. (a) Measured RSS for the example link. (b) ML estimates $\hat{\alpha}_{1:624}$ and $\hat{\Delta}_{1:624}$ with the corresponding joint density function, $\hat{p}(\hat{\Delta}_m, \hat{\alpha}_m)$. The copula transformed values are given by the upper and right axis. (c) Comparison between $\hat{p}(\hat{\Delta})$ and $\hat{p}(\hat{\Delta}|\hat{\alpha} = 8 \text{ dB})$, the latter given by the cross section along the red line in (b).

The conditional density is given as the cross section along the red line in Figure 11(b). We see that, given information that $\alpha = 8$ dB, we are in a better position to predict Δ than if we have not received this information. In this particular example, $E[\Delta|\alpha = 8 \text{ dB}] = 16$ s would probably be a better estimate than the unconditional mean, $E[\Delta] = 44.8$ s.

2) MI for All Links: Each link in the data set is processed in a similar way as in the example above, resulting in $L = 130$ MI values. Note that the number of segments, M, which is the number of samples used to compute the MI, varies between links. Furthermore, since I_0^M is a function of M, different thresholds are used for different links in the test.

The L MI values are presented in Figure 12 as a scatter plot. Each blue circle represents one link by its M and estimated MI and the red curve shows I_0^M. For every link appearing above this curve, we consider α and Δ to be dependent.

We see that all links passes the test. Thus, we can, at a significance level of 95%, reject the hypothesis that α and Δ are independent. Note, however, that this does not necessarily imply that the existing dependence is very strong.

C. Dependencies Between Parameters in Adjacent Segments

Dependencies between parameters in adjacent segments were detected in a similar manner as in Section VI-B. In this

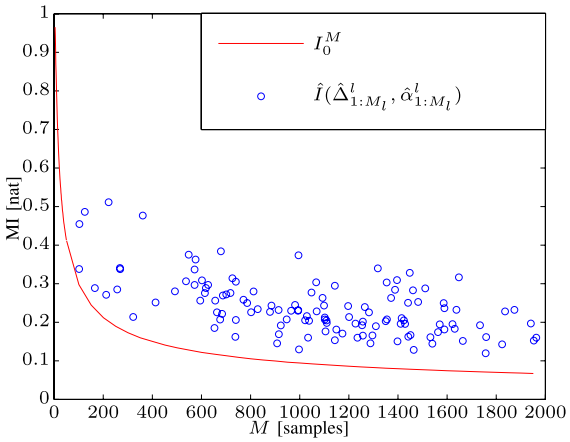

Fig. 12. Scatter plot of $\hat{I}(\hat{\Delta}^l_{1:M_l}, \hat{\alpha}^l_{1:M_l})$. The red line marks I_0^M.

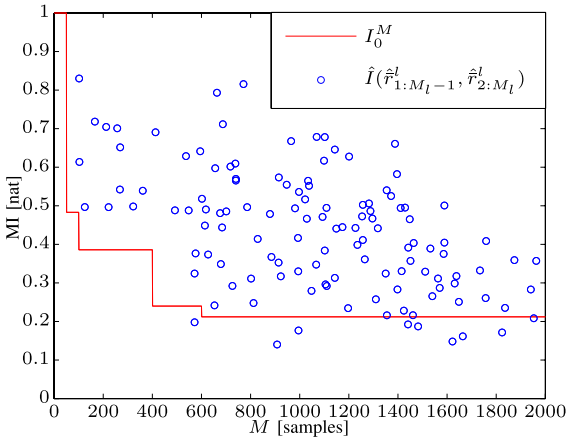

Fig. 13. Scatter plot of $\hat{I}(\hat{\tilde{r}}^l_{1:M_l-1}, \hat{\tilde{r}}^l_{2:M_l})$. Each blue circle correspond to a link and $\hat{I}(\hat{\tilde{r}}^l_{1:M_l-1}, \hat{\tilde{r}}^l_{2:M_l})$ is plotted against the number of segments, M, from which the estimate was obtained. The red curve shows the independence threshold that was used for different M.

case MI was estimated between parameters and shifted versions of themselves, leaving other cross-combinations to future work. Thus, we have restricted the study to MI estimates $\hat{I}(\hat{\alpha}_{1:M-1}, \hat{\alpha}_{2:M})$, $\hat{I}(\hat{\tilde{r}}_{1:M-1}, \hat{\tilde{r}}_{2:M})$, and $\hat{I}(\hat{\Delta}_{1:M-1}, \hat{\Delta}_{2:M})$.

The MI values between $\hat{\tilde{r}}_m$ and $\hat{\tilde{r}}_{m+1}$ for the different links are presented in Figure 13 as a scatter plot. As earlier, each blue circle represents one link by its M and estimated MI. When computing I_0^M, special care was taken to compensate for the bias in MI introduced by the CPD algorithm, as mentioned in Section IV. However, due to the high computational cost involved, I_0^M was only computed for a limited set of values, $M \in \{50, 100, 400, 600\}$. Tests involving arbitrary M used the I_0^M with the closest, smaller M in the above list. The red curve shows the independence threshold that was used for different M and for every link appearing above this curve, we consider $\hat{\tilde{r}}_m$ and $\hat{\tilde{r}}_{m+1}$ to be dependent.

The estimates $\hat{I}(\hat{\tilde{r}}^l_{1:M_l-1}, \hat{\tilde{r}}^l_{2:M_l})$ exceed I_0^M for 91% of the links and most links have a MI estimate well above I_0^M, indicating a significant statistical dependence.

Similar scatter plots for $\hat{\alpha}$ and $\hat{\Delta}$ are shown in Figures 14 and 15, respectively. Figure 14 suggests that $\hat{\alpha}$ has the weakest statistical dependence since the links are grouped close to I_0^M. Note, however, that most segments are fairly

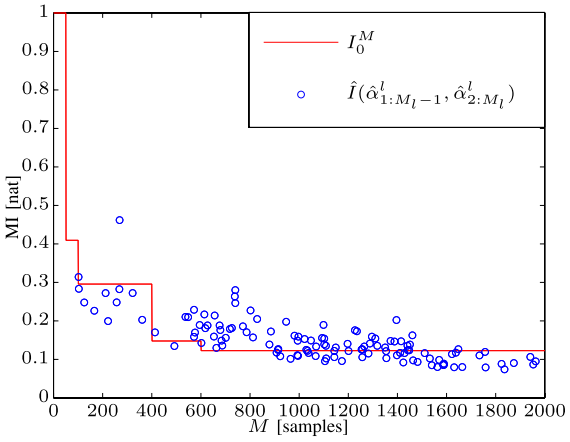

Fig. 14. Scatter plot of $\hat{I}(\hat{\alpha}^l_{1:M_l-1}, \hat{\alpha}^l_{2:M_l})$. The red curve shows the independence threshold that was used for different M.

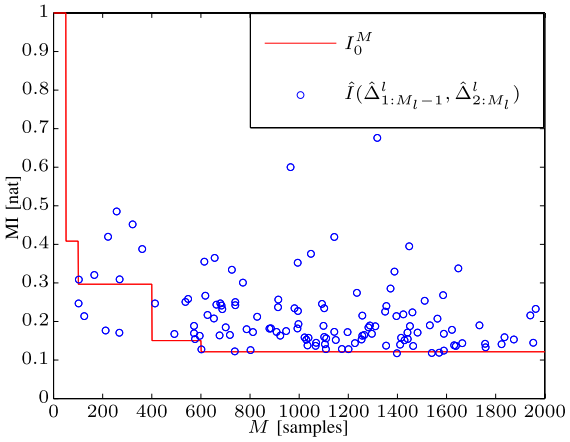

Fig. 15. Scatter plot of $\hat{I}(\hat{\Delta}^l_{1:M_l-1}, \hat{\Delta}^l_{2:M_l})$. The red curve shows the independence threshold that was used for different M.

TABLE III

PERCENTAGE OF LINKS SHOWING SIGNIFICANT STATISTICAL DEPENDENCE BETWEEN PARAMETERS CORRESPONDING TO ADJACENT SEGMENTS

Mutual information	%
$\hat{\alpha}_m, \hat{\alpha}_{m+1}$	54
$\hat{\tilde{r}}_m, \hat{\tilde{r}}_{m+1}$	91
$\hat{\Delta}_m, \hat{\Delta}_{m+1}$	95

static which corresponds to high values of α. With the coarse RSS resolution of 1 dBm, estimates of this parameter will suffer from erratic quantisation noise that, potentially, conceals statistical dependencies. Note that this is of minor concern in a routing context since the information on mean signal strength, $\bar{r}$, are more important for nearly static channels.

The results from the dependency tests are summarized in Table III. We see that only half of the links passes the test for α, but the majority of the links passes the test for $\bar{r}$ and Δ, with $\bar{r}$ yielding the largest estimates of MI.

VII. CONCLUSIONS AND DISCUSSION

We have introduced a regime switching model for fixed wireless links in the process industry and used this model for CPD to split the RSS time series into segments. From the

extracted segments, we presented statistics on their typical length as well as on the parameters that describe temporal fading within each segment. The segment lengths are typically 100-1000 s, but extreme values of only a few seconds or several hours have also been observed.

The bulk of the observed K-factors were found between 5 dB and 25 dB, which is slightly higher compared to results from previous similar studies. However, it should be noted that our method of estimating K-factors will typically yield higher values than methods that do not rely on a segmentation prior to the parameter estimation. The CPD will favor homogeneous segments, meaning that the fading variation within these segments will generally be less than for segments of predetermined length, which is the case in, e.g., [19].

In addition, we have used MI to identify dependencies in the fading process. For instance, the statistical dependencies were found between segment lengths and K-factors, where long segments are, typically, associated with close-to-static channels. This dependence could potentially be used to design a predictor of the next CP based on an estimate of the K-factor obtained from a few RSS samples and such predictions would allow the network manager to adapt the rate of beaconing transmissions. When a channel is close to static, a lower rate is preferable since further beaconing of a fading state that has already been observed contains little information and the resources are better spent sending data.

The fading distribution parameters also showed statistical dependencies between adjacent segments. For instance, more that 90% of the links proved to have a significant dependency for $\bar{r}$. This is interesting from a routing perspective since these parameters can be associated with a packet error rate, which often is the information that routing decisions are based upon.

The benefits of incorporating accurate information on link quality into the problem of state estimation over a WSAN, were emphasised by Quevedo et al. in [31]. Their method relies on a finite set of possible routing tables and full information on link quality. Developing methods to acquire this information was pointed out as an open research area, where our work may serve as a starting point.

Potential future work includes the design of predictors utilising these dependencies. Our work has been limited to studies of dependencies between parameters in adjacent segments. However, it is plausible that other parameter combinations may also yield strong dependencies, and predictors that use several parameters in combination, and over several segments, may prove superior.

Lastly, the algorithm that is outlined in Sections III-IV could be useful for other applications than the one considered in this paper. Whenever the regime switching model with gamma distributed observations is a reasonable model for the underlying process, the algorithm serves as a general tool for detecting regular behavior in the observed time series.

APPENDIX

A. Implementation of Viterbi

Running a direct implementation of Viterbi according to (17) on long data series, soon yield numbers that go below the machine precision on a computer. A simple way to avoid this problem is to implement the algorithm in log domain. By taking the log of,

$$\log Q^*(t) = \max_s \{\log \mathcal{L}(t,s) + \log Q^*(s+1) + I(s \neq T)\log\lambda + (s-t)\log(1-\lambda)\}. \quad (29)$$

Maximising (29) will lead to the same solution, $s^*(t)$, as maximising (17).

B. Efficient Computations of $\mathcal{L}(t, s)$

The maximisation over s in (17), requires the computation of $\mathcal{L}(t,s)$ for a fixed t with increments of $s = t+1, \ldots, T$. Since the sum, S, and product, P, in (3) form sufficient statistics to (2), recomputing for $\mathcal{L}(t, s+1)$ is a simple matter of updating S and P with r_{s+1}, followed by reiterating the Laplace method in (9). The Laplace method fits a Gaussian function to $g(\alpha)$, which involves finding the maxima α_0. In section D we show that $g(\alpha)$ has a strict global maxima, thus Newton's method quickly converges towards α_0.

C. Comments on the Choice of Conjugate Prior in (5)

For periods of time where the measured RSS are very static, small changes such as quantisation noise, can trigger the algorithm to form a new segment. This behavior can be counteracted by adjusting λ. Another way is to choose a prior, (5), that assigns low probability to high values for α. This will affect the algorithm's average resilience to form segments that consist of static RSS values. With parameters chosen as in Table I, the algorithm punish values of α that are larger than 5.

D. Laplace Approximation for Evaluating (9)

In the Laplace approximation, a Gaussian function is fitted to the posterior distribution in (9), for which the integral is analytically tractable [32]. This is done in two steps as follows,

Step 1: Find the mode, α_0, of the posterior, thus fulfilling $\frac{d}{d\alpha} g(\alpha_0) = 0$

Step 2: Compute the so called precision $A = -\frac{d^2}{d\alpha^2}\ln(g(\alpha_0))$

When α_0 and A has been computed, an approximation of the marginal likelihood is given by

$$\mathcal{L}(t,s) = \int_{0.5}^{\infty} f(\alpha)d\alpha \approx \int_{-\infty}^{\infty} \exp\left(\frac{A}{2}(\alpha-\alpha_0)^2\right)d\alpha = \sqrt{2\pi}\frac{g(\alpha_0)}{A^{\frac{1}{2}}}. \quad (30)$$

The derivative of the posterior is given by,

$$g'(\alpha) = (n+b)\psi_0(\alpha) - (n+c)\psi_0(\alpha(n+c)+1) + \ln\left(\frac{pa}{(s+d)^{n+c}}\right) \quad (31)$$

There is only one zero crossing of the above derivative and it is easily found through the numerical method of Newton-Raphson.

The precision A of the fitted Gaussian is given by the second order derivative of the logarithm of the posterior distribution. In this case given by,

$$g''(\alpha) = (n+c)^2\psi_1(\alpha(n+c)-1) - \frac{1}{\alpha^2} - (n+b)\psi_1(\alpha). \tag{32}$$

Results obtained by the Laplace approximation coincide almost completely with results obtained from trapezoidal integration indicating that the computed marginal likelihood values are accurate.

ACKNOWLEDGMENT

The authors would like to thank Mikael Gidlund at Mid Sweden University for granting us access to the factories. They would also like to thank the reviewers and the editor for their constructive comments.

REFERENCES

[1] M. Åkerberg, J. Gidlund, and M. Björkman, "Future research challenges in wireless sensor and actuator networks targeting industrial automation," in *Proc. 9th IEEE Int. Conf. Ind. Inform.*, Jul. 2011, pp. 410–415.

[2] D. Chen, M. Nixon, and A. Mok, *WirelessHART*. New York, NY, USA: Springer, 2010.

[3] J. Araujo, M. Mazo, Jr., A. Anta, P. Tabuada, and K. H. Johansson, "System architectures, protocols and algorithms for aperiodic wireless control systems," *IEEE Trans. Ind. Informat.*, vol. 10, no. 1, pp. 175–184, Feb. 2014.

[4] S. Tennina *et al.*, *IEEE 802.15.4 and ZigBee as Enabling Technologies for Low-Power Wireless Systems With Quality-of-Service Constraints*. Berlin, Germany: Springer, 2013.

[5] T. Ekman, "Prediction of mobile radio channels: Modeling and design," Ph.D. dissertation, Dept. Signals Syst., Uppsala Univ., Uppsala, Sweden, 2002.

[6] A. Duel-Hallen, S. Hu, and H. Hallen, "Long-range prediction of fading signals," *IEEE Signal Process. Mag.*, vol. 17, no. 3, pp. 62–75, May 2000.

[7] P. Fearnhead, "Exact Bayesian curve fitting and signal segmentation," *IEEE Trans. Signal Process.*, vol. 53, no. 6, pp. 2160–2166, Jun. 2005.

[8] N.-C. Wang and K. Yao, "Characterizing fading channel under abrupt temporal variations," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, May 2013, pp. 5056–5060.

[9] R. P. Adams and D. J. C. MacKay, "Bayesian online changepoint detection," Univ. Cambridge, Cambridge, U.K., Tech. Rep. arXiv:0710.3742, 2007.

[10] D. Niyato, E. Hossain, and J. Diamond, "IEEE 802.16/WiMAX-based broadband wireless access and its application for telemedicine/e-health services," *IEEE Wireless Commun.*, vol. 14, no. 1, pp. 72–83, Feb. 2007.

[11] L. J. Greenstein, S. S. Ghassemzadeh, V. Erceg, and D. G. Michelson, "Ricean K-factors in narrow-band fixed wireless channels: Theory, experiments, and statistical models," *IEEE Trans. Veh. Technol.*, vol. 58, no. 8, pp. 4000–4012, Oct. 2009.

[12] P. Soma, D. S. Baum, V. Erceg, R. Krishnamoorthy, and A. J. Paulraj, "Analysis and modeling of multiple-input multiple-output (MIMO) radio channel based on outdoor measurements conducted at 2.5 GHz for fixed BWA applications," in *Proc. IEEE Int. Conf. Commun. (ICCC)*, vol. 1. Apr. 2002, pp. 272–276.

[13] L. Ahumada, R. Feick, R. A. Valenzuela, and C. Morales, "Measurement and characterization of the temporal behavior of fixed wireless links," *IEEE Trans. Veh. Technol.*, vol. 54, no. 6, pp. 1913–1922, Nov. 2005.

[14] H. Hashemi, "The indoor radio propagation channel," *Proc. IEEE*, vol. 81, no. 7, pp. 943–968, Jul. 1993.

[15] C. Oestges, N. Czink, B. Bandemer, P. Castiglione, F. Kaltenberger, and A. J. Paulraj, "Experimental characterization and modeling of outdoor-to-indoor and indoor-to-indoor distributed channels," *IEEE Trans. Veh. Technol.*, vol. 59, no. 5, pp. 2253–2265, Jun. 2010.

[16] B. Bandemer, C. Oestges, N. Czink, and A. Paulraj, "Physically motivated fast-fading model for indoor peer-to-peer channels," *Electron. Lett.*, vol. 45, no. 10, pp. 515–517, May 2009.

[17] R. J. C. Bultitude, "Measurement, characterization and modeling of indoor 800/900 MHz radio channels for digital communications," *IEEE Commun. Mag.*, vol. 25, no. 6, pp. 5–12, Jun. 1987.

[18] T. S. Rappaport and C. D. McGillem, "UHF fading in factories," *IEEE J. Sel. Areas Commun.*, vol. 7, no. 1, pp. 40–48, Jan. 1989.

[19] E. Tanghe *et al.*, "The industrial indoor channel: Large-scale and temporal fading at 900, 2400, and 5200 MHz," *IEEE Trans. Wireless Commun.*, vol. 7, no. 7, pp. 2740–2751, Jul. 2008.

[20] P. Agrawal, A. Ahlén, T. Olofsson, and M. Gidlund, "Long term channel characterization for energy efficient transmission in industrial environments," *IEEE Trans. Commun.*, vol. 62, no. 8, pp. 3004–3014, Aug. 2014.

[21] T. H. Chua, I. J. Wassell, and T. A. Rahman, "Wind-induced slow fading in foliated fixed wireless links," in *Proc. IEEE 75th Veh. Technol. Conf.*, May 2012, pp. 1–5.

[22] E. Vinogradov, W. Joseph, and C. Oestges, "Measurement-based modeling of time-variant fading statistics in indoor peer-to-peer scenarios," *IEEE Trans. Antennas Propag.*, vol. 63, no. 5, pp. 2252–2263, May 2015.

[23] P. Sadeghi, R. A. Kennedy, P. B. Rapajic, and R. Shams, "Finite-state Markov modeling of fading channels—A survey of principles and applications," *IEEE Signal Process. Mag.*, vol. 25, no. 5, pp. 57–80, Sep. 2008.

[24] M. D. Yacoub, J. E. V. Bautista, and L. G. de Rezende Guedes, "On higher order statistics of the Nakagami-m distribution," *IEEE Trans. Veh. Technol.*, vol. 48, no. 3, pp. 790–794, May 1999.

[25] D. Fink, "A compendium of conjugate priors," Dept. Biol., Montana State Univ., Bozeman, MT, USA, Tech. Rep. DOE 95831, 1997.

[26] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, A. Jeffrey and D. Zwillinger, Eds., San Diego, CA, USA: Academic, 1980.

[27] A. A. Margolin *et al.*, "ARACNE: An algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context," *BMC Bioinform.*, vol. 7, no. 1, p. S7, 2006.

[28] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York, NY, USA: Wiley, 1991.

[29] R. S. Calsaverini and R. Vicente, "An information-theoretic approach to statistical dependence: Copula information," *Europhys. Lett.*, vol. 88, no. 6, p. 68003, 2009.

[30] M. Nakagami, "The m-distribution—A general formula of intensity distribution of rapid fading," in *Statistical Methods in Radio Wave Propagation*. New York, NY, USA: Pergamon, 1960.

[31] D. E. Quevedo, A. Ahlén, and K. H. Johansson, "State estimation over sensor networks with correlated wireless fading channels," *IEEE Trans. Autom. Control*, vol. 58, no. 3, pp. 581–593, Mar. 2013.

[32] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer-Verlag, 2006.

Markus Eriksson is currently pursuing the Ph.D. degree in electrical engineering and signal processing with the Signals and Systems Group, Uppsala University. He has a professional background as a Software Developer with Scania CV, Södertälje, Sweden. His research interests are in the fields of signal processing for wireless sensor networks, embedded systems, and energy harvesting.

Tomas Olofsson received the Ph.D. degree in signal processing from Uppsala University, in 2000. He is currently an Associate Professor with the Department of Engineering Sciences, Signals and Systems Division, Uppsala University. His research has been involved in inference problems, in particular inverse problems in ultrasonics and lately, work on communications within wireless sensor network, in particular channel characterization for wireless control applications.