Signal and Data Processing Techniques for Industrial Cyber-Physical Systems

G. Tzagkarakis[1], G. Tsagkatakis[1], D. Alonso[3], C. Asensio[3], E. Celada[3], A. Panousopoulou[1], P. Tsakalides[1], and B. Beferull Lozano[2,3]

[1] *Foundation for Research & Technology - Hellas (FO.R.T.H.), Institute of Computer Science (I.C.S.)-Signal Processing Laboratory, Greece*

[2] *University of Agder, Centre for Integrated Emergency Management (CIEM), Norway*

[3] *Universidad de Valencia, Group of Information and Communication Systems, Spain*

Abstract: Recent advances in computing and communications have raised a constantly increasing challenge of modernizing and decentralizing industrial processes, by introducing dynamic architectures of cyber-physical systems (CPS). The resulting industrial CPS (iCPS) exploit their inherent relationship with wireless sensor networks (WSN) for providing cost-effective, scalable and easily-deployed solutions in industrial spaces, whilst formulating new paradigms for industrial data acquisition and control. Employing iCPS in industrial processes poses characteristics which are not dominant in conventional monitoring scenarios. In specific, the deployment of the WSN components is highly coupled to the objectives of the industrial process, thus introducing logical, along with spatial and temporal, correlations between data streams originated by different positions. Moreover, typical WSN imperfections, such as limited bandwidth, computational complexity, and lifetime, should be treated as inseparable part the processes responsible for minimizing the WSN maintenance efforts. Finally, the quality of sensing is strongly coupled to the objectives and the performance of the industrial control laws, necessitating for intensified reliability and robustness in signal processing and data acquisition, as the complexity of the industrial process increases.

The objective of this book chapter is to present a framework of signal and data processing for treating different layers of information abstraction in an iCPS. By taking into account the limitations and imperfections of WSN employed for iCPS, we will focus on three complementary, yet equally important aspects: (a) signal processing-driven performance optimization for industrial WSN; (b) in-network signal processing techniques for distributed field estimation; (c) high-level data management and analysis for detecting abnormal behavior in the recorded iCPS data. The developed algorithms aim at formulating an integrated framework for signal and data processing in iCPS. With a strong emphasis on providing the essential theoretical background, the efficacy of the resulting framework will be ultimately evaluated in different aspects of real-life iCPS, designed for the autonomous monitoring of water treatment plants. It is anticipated that the methods herein presented and the accompanying discussions on the obtained results, will yield novel directions for iCPS standardization.

I. Introduction

Cyber-physical systems (CPS) are large-scale interconnected systems of heterogeneous, yet collaborating, components that are envisioned to provide integration of computation with physical processes. Nowadays, a first generation of cyber-physical systems can be found in areas as diverse as aerospace, civil infrastructures, energy, healthcare, manufacturing, transportation, entertainment, and consumer appliances.

Unlike traditional *embedded systems*, a full-fledged cyber-physical system is typically designed as a network of interacting elements with physical input and output instead of being simply a combination of standalone devices. The inherent heterogeneity and integration of different components pose new challenges to traditional data analysis, communication, control, and software theories. This often makes system design inefficient with current technologies.

Advances in the cyber world, such as communications, networking, sensing, computing, storage, and control, as well as in the physical world, such as materials, hardware, and renewable energy sources, are all converging rapidly to dramatically increase the adaptability, autonomy, efficiency, functionality, reliability, safety, and

usability of cyber-physical systems. This aspires to broaden the economic and societal potential of such highly collaborative computational systems in various dimensions, such as intervention (*e.g.,* collision avoidance), precision (*e.g.,* robotic manufacturing), operation in dangerous or inaccessible environments (*e.g.,* search and rescue), efficiency (*e.g.,* energy and cost reduction in water treatment plants), and augmentation of human capabilities (*e.g.,* healthcare monitoring and delivery).

A major difference between CPS and a typical control system or an embedded system is the use of communications, which adds reconfigurability and scalability, as well as complexity and potential instability. Furthermore, CPS have significantly more intelligence in sensors and actuators, as well as substantially stricter performance and energy constraints, which are critical for its efficiency and lifetime. On the other hand, cyber capabilities are embedded in every physical process and component, networking is employed at multiple scales, complexity lies at multiple temporal and spatial scales, and high heterogeneity is seen across devices and protocols.

The design and realization of the complex interface between cyber and physical worlds for seamless interactions is by no means a non-trivial task. Implacable, concurrent laws of physics govern the physical world, as opposed to the discrete and asynchronous nature of the cyber world. Timing and spatial precision, uninterrupted connectivity, predictability and repeatability are extremely critical for the cyber-physical interface. It is hence vital to build new theoretical foundations, scientific models, abstractions, and explicit dissociation between cyber and physical worlds for the interface, and rethink or reinvent interface functions, such as coordination, integration, monitoring and control.

Focusing on a demanding paradigm of interrelation between the cyber and physical worlds, *industrial control systems* are widely used to provide autonomous control through appropriate control loops dedicated to performing specific tasks. Such systems monitor an industrial environment through sensors deployed around the product line and interact with the various processes through proper actuators. Moreover, the complexity of modern industrial settings is usually simplified by dividing the overall infrastructure into individual subsystems containing separate processing and control modules. When interactions between distinct subsystems are required, usually skilled system operators or simple communication methods are exploited.

To this end, cyber-physical systems can provide broad control over complex and large industrial environments through heterogeneous network architectures of sensors, actuators, and processors. However, coverage and connectivity should be redefined in the framework of *industrial cyber-physical systems* (iCPS). Such systems will usually consist of both wired and wireless sensor and actuation networks with different capacities and reliability. Furthermore, emphasis is put on real-time operations, whereas sensing, processing, communication and actuation will be handled by different components in the iCPS infrastructure. To model such heterogeneous issues, an innovative technological turn is necessary.

To compromise all those critical aspects of iCPS, sophisticated signal and data processing techniques, coupled with novel network and communication protocols, should be designed to provide unprecedented performance and efficiency levels for industrial cyber-physical systems.

In this book chapter, we present an integrated framework of signal and data processing for treating different layers of information abstraction, ranging from raw samples at the front-end of the cyber-physical space, to data semantics for extracting high-level patterns. By taking into account the limitations and the imperfections of the sensor network infrastructure employed for iCPS, we focus on three complementary, yet equally important aspects: (a) signal processing-driven performance optimization for industrial sensor networks; (b) in-network signal processing techniques for estimation, detection and tracking for iCPS; (c) knowledge management for detecting behavior variations in the recorded iCPS data.

Concerning the first aspect, a novel technique is introduced in the framework of matrix completion for recovering missing information due to network or sensor failures, as well as in the case of adapting to higher temporal resolution than the operating resolution of the available sensors. With respect to the second aspect, an efficient iterative consensus method is introduced for distributed estimation and tracking, by employing a cross-layer link scheduling protocol. Finally, regarding the third objective, an uncertainty-aware framework for extreme events detection, in conjunction with a fast and robust correlations monitoring are analyzed. Doing so, we improve upon existing data management systems by proposing an integrated framework spanning all the stages of the data acquisition and processing in iCPS infrastructures.

While the corresponding methods reflect on different angles of the iCPS architecture, the common factor is the exploitation of the inherent information content of raw sensor streams, towards the formulation of an integrated mathematical and algorithmic framework for signal and data processing in iCPS. With a strong emphasis on providing the essential theoretical background, the efficacy of the resulting framework will be ultimately evaluated in different aspects of real-life iCPS, designed for the autonomous monitoring and decentralized control of water treatment plants [1]. It is anticipated that the methods herein presented and the accompanying discussions on real-life results, will yield novel directions for iCPS standardization.

The rest of the chapter is organized as follows: Section II introduces the main concepts of data-driven architectures for industrial cyber-physical systems, while Section III describes novel signal processing methods based on the theories of compressed sensing and matrix completion for recovering missing information in sensor streams. Section IV presents techniques for performing in-network signal processing tasks, such as parameter estimation and tracking, in a distributed fashion, and discusses their main performance issues. In Section V, a set of novel algorithmic tools is introduced for producing intelligent reasoning over the data by supporting advanced operations, such as, querying, uncertainty-aware high-level data analysis, and alerting. Finally, in Section VI the performance of the introduced signal and data processing techniques is evaluated in the real-world scenario of the HYDROBIONETS project, whose key objective is the autonomous monitoring and control of industrial water treatment and desalination plants, while Section VII summarizes the main achievements and gives directions for further enhancements.

II. Data-Driven Architectures for Industrial CPS

The transition from cyber to cyber-physical architectures is dictated by the tight coordination between cyber and physical resources; while traditional cyber systems observe a constantly changing physical world, in CPS the information processing components are inseparable from the physical procedures. As such, despite the inherited heterogeneity and complexity, CPS architectures should respond effectively to unexpected conditions, whilst exhibiting an increased level of resiliency and adaptability to system failures.

This necessity intensifies for the case of industrial processes, as the harshness of the operational environment is combined with strict performance requirements in terms of data acquisition, estimation, and control. In particular, the practical realization of wireless sensor-actuator networks (WSAN), which are considered the enablers of CPS architectures [2], is characterized by periods of severe impairments due to propagation phenomena, noise, and interference, dictated by the characteristics of the operational space [3]. In RF-harsh environments, like those met in industrial plants, the surrounding environment becomes even less mindful of the wireless communications performance. The sources of noise and interference increase significantly due to the presence of heavy machinery, obstacles with large volume and highly reflective characteristics [4]. In such operational spaces, the phenomenon of link disruption due to inter-symbol interference has, along with multipath fading and signal absorption, a direct impact on the network performance that cannot be treated with conditional methods, such as increased transmission power [5]. In addition, whilst electromagnetic interference affects the hardware characteristics of the network components [6], the presence of impulsive noise can lead to short periods of excessively weak channel conditions [7].

These limitations have a direct impact on the performance and reliability of the network, which are typically expressed in terms of packet losses, and excessive delays. Modern, networked approaches for industrial processes that rely on WSAN consider these imperfections of the network performance as an inseparable aspect of the control procedures. As such, the respective, system-wide architectures, emphasize on guarding the control requirements against the uncertainties imposed by network. Subsequently, data-driven schemes for iCPS often coincide with Networked Control Systems (NCS) [8], [9], which feature the communication of spatially distributed sensors, actuators and controllers over a shared, data network (Figure 1(a)).

During the last decade NCS have faced several challenges related to the Quality-of-Service (QoS) provided by the underlying network backbone. Representative examples include the preservation of the required bandwidth, the network synchronization, and the compensation from information losses. Depending on the network parameter that is modelled as part of the industrial process, the control loop can be viewed as an *asynchronous dynamic system*, for which the stability margins are guaranteed by retaining the information loss below the threshold dictated by the characteristics of the open loop system [10].

The recent advances on wireless communication standards for industrial processes, such as WiHart [11], ISA100 [12], and IEEE 802.15.4 [13], have given a different spin in the area of data-driven information processing and control. The scale and the complexity of the problem has been dramatically increased, thereby seeking cross-layer communication and control solutions, employing multi-hop topologies, and, ultimately, distributing the intelligence of the system across different parts of the iCPS architecture. Along this direction, in [14] an integrated control-communication framework is presented. It comprises of a communication protocol that enables self-triggered control actions, and an optimization algorithm based on simulated annealing, whilst considering both packet losses, as well as the physically constrained actions of the actuators. In parallel, the authors in [15] propose a mathematical framework that can support the operation of an iCPS, over a multi-hop WSAN. The resulting architecture is abstracted as a switched system, while link failures introduce random switching signals. Transiting this approach into a real-world industrial process provides a scalable architecture according to which each control loop is analyzed separately and associated to the maximum delay between sensing and actuation. All control loops are then integrated and a scheduling policy is applied, allowing data transmission from the sensors to the controller and from the controller to the actuators for each control loop, within the specified time bounds. This approach is characterized as compositional, as it enables applying the set of all schedules that satisfy the temporal boundaries for a single control loop to additional control loops, added at a later stage in the iCPS architecture.

While the aforementioned approaches emphasize on the extreme ends of the iCPS architecture, and in specific how the controller can treat the network imperfections as part of the process under control, recent theoretical results have indicated that increasing the intelligence on the sensors and intermediate networked components can improve the stability and performance of the entire system [16], [17]. Building upon this statement, whilst considering a multi-hop, arbitrary topology between the sensor and the remote estimator/controller, an iCPS architecture that enables in-network processing is proposed in [18], originated by the point of sensor and recursively adopted by each node located on the network path between the sensor and the estimator/controller (Figure 1(b)). In specific, under the assumption of perfect synchronization between all involved nodes, when the sensor collects a new measurement its encoder applies a Kalman filter for updating the estimation of the physical process, updates the time instant, and transmits the updated estimation to all outgoing edges. Upon reception of this information, each intermediate node checks the timestamp on the incoming data and keeps the one corresponding to the latest measurement, which is in turn re-directed to its outgoing edges. The destination node, located at the side of estimator/controller will collect all incoming information and retain the one corresponding to the most recent measurement for either calculating the best possible estimate in terms of minimum mean square error, or minimizing a quadratic cost.

The mechanisms described so far rely on the existence of both a single sensor-actuator pair, as well as a designated estimator / controller which is located somewhere in the industrial plant. While this convention corresponds to many of existing industrial processes, its practical realization is often subject to the computational and energy constraints of the embedded platforms, which are often employed for implementing iCPS architectures due to their cost-effectiveness and miniatured size. In addition, modern industrial processes, such as Smart Grid and Smart Water Networks, require multiple sensor and actuator points, distributed in a wide geographical area and characterized by heterogeneous signals travelling from and to different points of interest. Driven by this motivation Pajic et al. introduced in [19], [20], a novel paradigm of iCPS architectures that enables the entire network acting as a controller, instead of assigning this role to designated nodes of the network. The resulting architecture, entitled Wireless Control Network (WCN), is considered an extension of traditional approaches, that builds upon the collaboration of the operational nodes with their 1-hop neighbors. As shown in Figure 1(c), under the assumption of a a-priori known topology, some nodes have access to the sensor measurements, while others are located in the communication range of the actuators. Based on a linear iterative strategy, each node periodically updates its state, which results from the linear combination of the states of its 1-hop neighbors. Similarly, the actuators apply linear combinations of the states of the nodes in their neighborhood. The resulting overall scheme has been proven capable of controlling continuous-time physical processes, whilst preserving the stability of the system under the existence of independent link failures.

The discussion thus far highlights the emphasis that has been given on improving feedback control laws

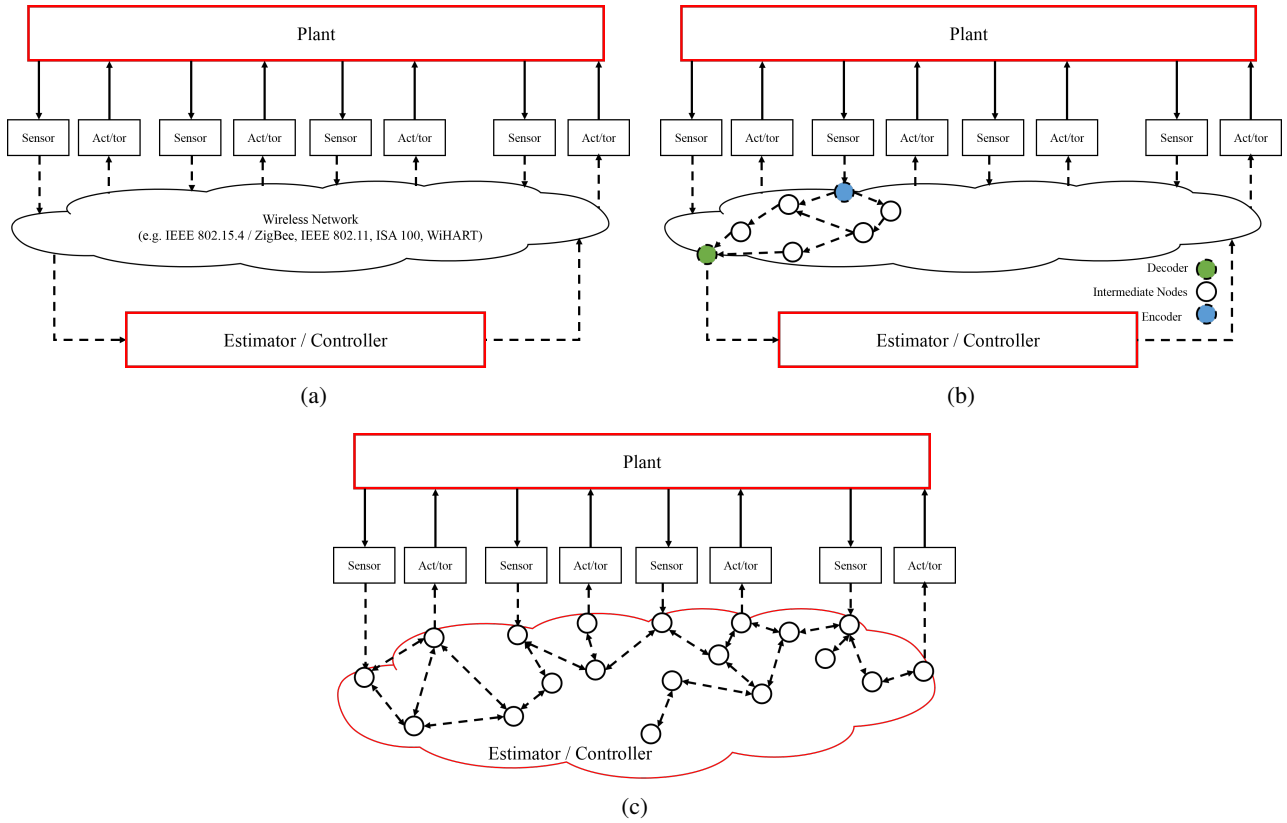


Fig. 1. Data-driven iCPS architectures: (a) traditional NCS approach – sensors and actuators communicate with the controller/estimator of the system over a wireless network; (b) a more sophisticated approach – encoders and decoders collaborate with the intermediate wireless nodes for improving the high-level control/estimation procedure [18] over multi-hop network paths; (c) architecture of the Wireless Control Network ([19], [20]) – control/estimation process is distributed inside the fully connected network.

against the presence of network imperfections in industrial processes. While this is an essential requirement for iCPS, challenges associated to the system-wide *information management*, capable of covering the entire span of the industrial procedure, remain open.

More specifically, the utilization of inexpensive sensing nodes is crucial for the proliferation of CPS in every-day environments. In industrial settings, the challenges associated with this specific type of environment, such as excessive temperature or humidity, can hinder the performance of even high-end sensing nodes. As a consequence, node failures are frequent and the systems must be able to handle the dynamic insertion and removal of sensors. Although catastrophic failures do happen, node failure is typically attributed to energy depletion. The design of iCPS platforms must therefore maximize the amount of information that can be extracted from the data, while minimizing the usage of valuable resources, such as power, bandwidth and storage.

In addition to the fragility of the infrastructure, real-life architectures are characterized by periods of communication failures due to changes in the environmental conditions, such as the occurrence of moving objects, interference from other devices, and conjunction due to network traffic overload. As a consequence, packets that encode the captured data will be lost, while disseminating information through the network to other nodes or local sinks. System designers must therefore account for such communications impairments and provide robustness without consuming unnecessarily the networks' resources.

These properties and characteristics suggest that modern data-driven architectures iCPS should consider:

- Data-driven techniques for improving the low-level network and sensing imperfections in the industrial environment. According to the current trend, such shortcomings on the joint sensing-communicating performance are assimilated into the controlled process.
- Mechanisms capable of distributing the intelligence of the system between sensors, and thereby extracting

correlations associated to the quality of sensing in the industrial plant. Related approaches that consider a distributed approach elaborate on collaboration between sensing and pure network components, for achieving the higher-level control and estimation objectives.
- Algorithms for high-level decision making capable of keeping the end-user in the control loop, based on heterogeneous sources of information. According to [21], such mechanisms are essential for the system-wide reliability and robustness of CPS.

These aspects are considered a necessity for modern iCPS architectures, capable of adopting a knowledge engineering approach [22]; they would enable the layered contextualization of real-time streamed data according to the level of abstraction and the different perspectives of the same infrastructure. As such, in parallel to the control perspective, three additional views and respective levels of data abstractions are recognized:

- Front-end data representations, provided by sensors deployed at physical frontier of the industrial process and capable of improving data acquisition and sampling;
- In-network correlations, enabled by the decentralized collaboration between sensing components, for extracting useful observations on the quality of sensing and the industrial-driven relationships between physically distant components of the same industrial process;
- High-level data abstractions, resulting from distilling information from the raw data streams into laconic notifications on the status and the quality of the underlying sensing and actuation procedures.

With these considerations in mind, in the remaining of this section we will introduce our framework for signal and data processing for iCPS, capable of treating different layers of information abstraction, ranging from raw samples at the front-end of the cyber-physical space, to data semantics for extracting high-level patterns.

A. An integrated framework for signal and data processing for iCPS

The herein proposed framework yields essentially a virtual multi-layered architecture, corresponding in a straightforward manner to the data-abstraction layers described above. In an attempt to magnify the benefits of sensing at the industrial processes, our emphasis is explicitly on the signal and data processing techniques that have impact on improving the *intelligence* of the information that flows from the sensors towards the controlling processes. Therefore, our approach can act as a complementary tool for sophisticated networked controlled mechanisms, whilst decoupling the limited and imperfect access to the sensor data streams from the specific characteristics of the controlled plant and / or the design of the controller.

In a nutshell, our framework considers the following levels of information processing:

- **Level 1: Signal-modeling of the front-end industrial data representations**, focusing on two recently developed yet extremely influential, signal-processing paradigms, namely Compressed Sensing, and Matrix Completion;
- **Level 2: In-network processing** for estimation, detection and tracking for iCPS, while taking into account the impact of network imperfections;
- **Level 3: High-level data analysis and early warning**, focusing on uncertainty management, notification mechanisms for extreme events, and extraction of high-level pairwise correlations for improving decision support making systems in industrial processes.

The proposed architecture is presented in Figure 2, highlighting the positioning of each information processing tier with respect to the generalized industrial control process. It is considered important to highlight the particular interest on providing to the system administrator the necessary means for qualitative supervision of the industrial procedure. At the level of raw sensing from the plant, the proposed framework considers the front-end data handlers, which are responsible for improving the data acquisition and sampling according to the statistical and mathematical attributes of the iCPS data. These handlers are in turn combined to sensing agents (Level 2), which travel within the wireless network for the characterization of the quality of sensing, the recovery of an accurate sample from the target field, and tracking of time varying signals. This approach operates in a distributed manner, thereby transiting the tasks of noise detection, signal estimation and tracking to the level of low-level sensing, and increasing the intelligence of the iCPS architecture, in order to account for failures at the level of sensing and/or networking. Finally, the sensor streams are fed into Level 3, which is

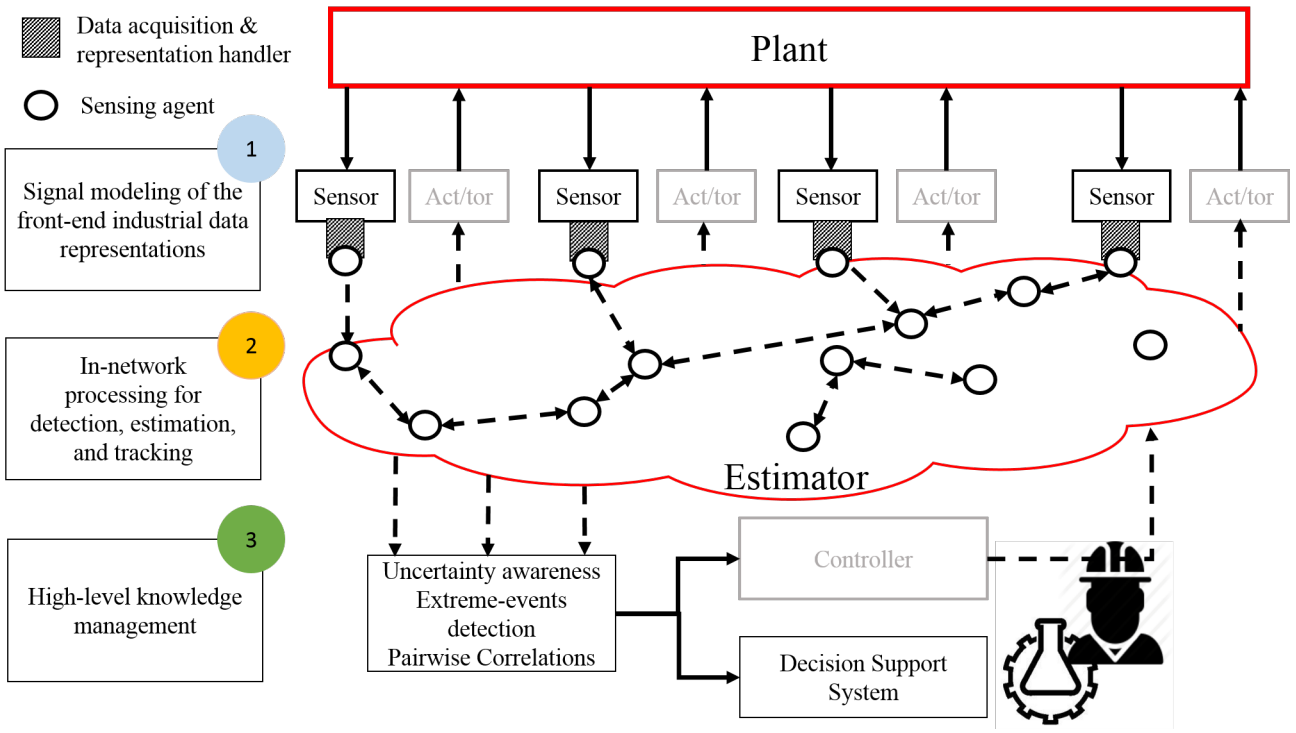


Fig. 2. The proposed framework for signal and data processing in industrial CPS, featuring the three information processing tiers.

responsible for further data analysis towards the qualitative and quantitative characterization of the industrial sensing process. The result can be directed towards both the controller components for adjusting the control actions in the presence of unexpected conditions, as well as the user-oriented decision support systems for facilitating the supervision of the overall system.

While the corresponding methods reflect on different angles of the iCPS architecture, their common factor is the exploitation of mathematical and statistical characteristics of the information, and not the information itself, thereby remaining agnostic with respect to the data-driven details for each industrial process. As such, while addressing aspects complementary to the control of industrial processes, it is anticipated that the proposed framework will provide a universal methodology for signal and data processing techniques that can be employed at different industrial scenarios.

In the remaining of this chapter we will elaborate on the theoretical background of each information processing tier, followed by a discussion on the evaluation of the proposed framework in different aspects of real-life iCPS, designed for the autonomous monitoring and decentralized control of water treatment plants.

III. Industrial Data Representations: From Dense to Sparse Sampling

A. *Properties and issues of iCPS data*

Data acquisition and processing in iCPS relies on the presence of a Wireless Sensor Network (WSN) composed of distributed nodes that communicate their measurements to other nodes with similar or higher capabilities. Efficient acquisition and communication of these measurements is a critical aspect of WSN systems that determine directly the lifetime and usability of the infrastructure. To address the issues associated with data sampling and processing, one must consider both the physical constraints of such networks, ranging from node failures to communications break-downs, as well as the properties and recovery capabilities of cutting-edge signal sampling and processing algorithms.

To achieve the strict requirements and overcome the limitations of such environments, one can exploit data redundancies and employ signal processing algorithms to guarantee the performance for numerous types of signals. In the following, we will focus on two particular characteristics of iCPS data, namely, sparsity and

low rankness. These characteristics are linked directly with properties of WSN architectures including spatio-temporal correlations, predictable behaviour, and physical constraints. Sparsity can refer to either the presence of a very small number of large-valued measurements or to the ability in expressing a complex signal using a small number of representative examples. While the former case reflects a specific type of signals, such as biological, seismic and astronomic data, the latter corresponds to a rather large number of signals that can be accurately represented using a sparse collection of fundamental signals encoded in a so-called *dictionary*. Intelligent exploitation of sparsity can offer numerous advantages from a WSN point of view.

Furthermore, we will consider the low rankness of various matrices that can be found in WSNs, such as measurement matrices, where each row of the matrix corresponds to a specific sensor, and each column represents a sampling instance in time. The rank of such matrices is indicative of the amount of correlation that exists within the data, since highly correlated measurements lead to low rank measurement matrices. The rank of a matrix is manifested by the number of non-zero singular values, in which case low-rank matrices can be described by sparse singular values. By exploiting the low rank property, significant benefits can be achieved for WSN architectures, such as efficient sampling and robust storage.

The sparsity and low rankness of iCPS data is demonstrated in Figure 3, which presents iCPS data collected by a WSN (part of the Intel-Berkeley dataset [1]) along with certain properties of the matrix. More specifically, Figure 3(b) shows the magnitude of the signal values (in blue). One can observe that a very small number of singular values capture most of the signals' energy, while the rest correspond to noise and outliers. Furthermore, Figure 3(c) shows the magnitude of the representations coefficients for three sampling instances, based on a mapping in a dictionary generated using data (measurements vectors) collected from the previous day. One may also observe that only a small number of coefficients are non-zero leading to a sparse representation of the new vectors in terms of the dictionary atoms. Two innovative signal processing algorithms are presented, namely, *compressed sensing* (CS) and *matrix completion* (MC), which can exploit efficiently the sparsity and low rankness of iCPS data captured via WSNs.

B. Compressed sensing

Compressed sensing (CS) is a radically novel approach in signal acquisition and processing [23], [24]. The main underlying concept of CS is that a complex signal can be recovered from a small number of random measurements, far below the traditional Nyquist-Shannon limit. The key assumption in CS is that either the signal itself is sparse or that it can be *sparsely* represented in an appropriate dictionary, and that enough *random* measurements are collected. Formally, a signal $\mathbf{s} \in \mathbb{R}^N$ is called k-sparse if $\|\mathbf{s}\|_0 < k$, where $\|\mathbf{s}\|_0 = \#$ non-zero elements of $\mathbf{s}$. This signal can be reliably recovered from a low-dimensional representation $\mathbf{y} = \mathbf{\Psi s} \in \mathbb{R}^M$, where $M \ll N$ by solving an ℓ_0-constrained minimization problem given by:

$$\min \|\mathbf{s}\|_0 \quad \text{subject to} \quad \mathbf{y} = \mathbf{\Psi s} \; . \tag{1}$$

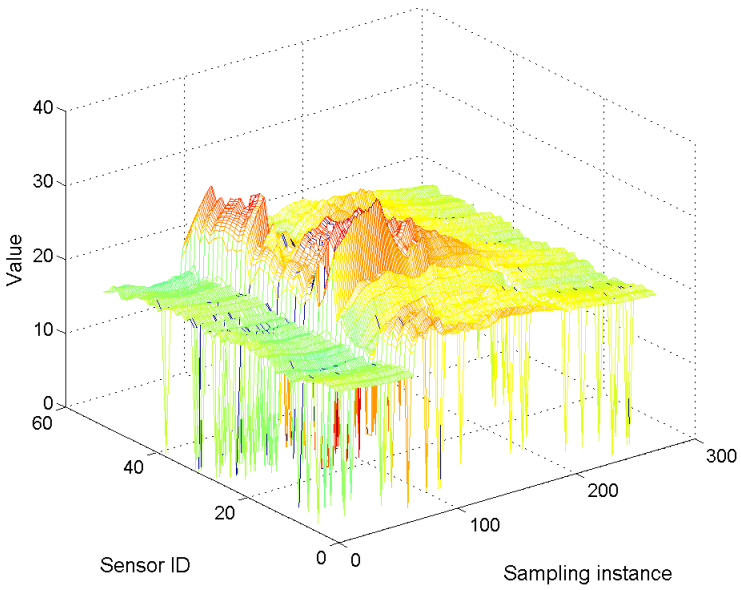
To guarantee the stable recovery of the original signal, the $M \times N$ sensing matrix $\mathbf{\Psi}$ must satisfy the so-called restricted isometry property (RIP). A sensing matrix $\mathbf{\Psi} \in \mathbb{R}^{M \times N}$ satisfies the RIP with isometry constant $0 \leq \delta < 1$ if for all k-sparse signals, $\mathbf{s}$, it holds that:

$$(1 - \delta)\|\mathbf{s}\|_2^2 \leq \|\mathbf{\Psi x}\|_2^2 \leq (1 + \delta)\|\mathbf{s}\|_2^2 \; . \tag{2}$$

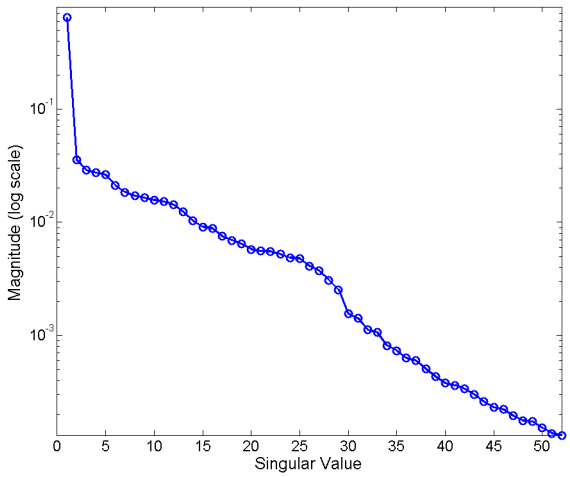
Designing such a sensing matrix is proven to be a challenging task. However, it has been proven that matrices whose elements are drawn randomly from appropriate distributions satisfy the RIP with high probability. Examples of such distributions include normalized mean bounded variance Gaussian [23] and Rademacher [25] distributions.

The formulation of CS expressed by (1) assumes that the signals in question are naturally sparse, that is, they consist of a small number of non-zero elements. However, a large class of signals do not belong to this category. To tackle this issue, the CS theory has been extended through the use of a dictionary of elementary examples as a sparsifying transform. During the early stages of CS theory formulation, well known orthogonal transforms, including the discrete Fourier transform (DFT), the discrete cosine transform (DCT),

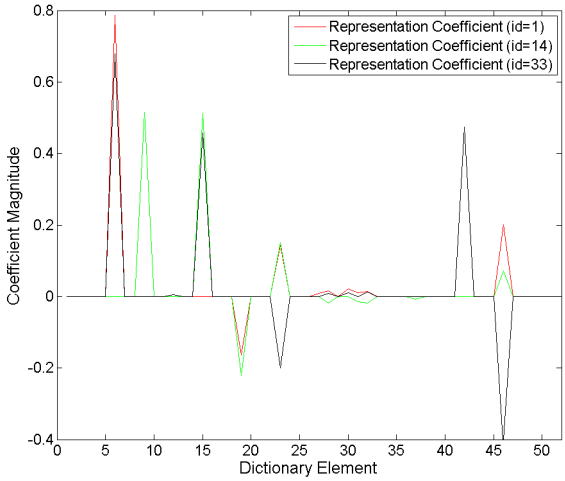
[1] P. Bodik, C. Guestrin, W. Hong, S. Madden, M. Paskin, and R. Thibaux, http://select.cs.cmu.edu/data/labapp3/index.html



(a) Measurements



(b) Properties



(c) Properties

Fig. 3. Graphical illustration of (a) a WSN measurements matrix, (b) the corresponding singular values, and (c) the representation coefficients. We observe that i) a small number of singular values captures most of the signals' energy and ii) only a small number of non-zero coefficients suffice for representing the measurements vectors.

and wavelets were employed as sparsifying dictionaries. In [24] it was shown that the theory of CS is also applicable in cases where the signal is sparse over coherent and redundant dictionaries, including overcomplete DFT, wavelet frames, and concatenations of multiple orthogonal bases. By incorporating the dictionary in the reconstruction problem, the dictionary-based ℓ_0 minimization is formulated according to:

$$\min \|\mathbf{s}\|_0 \quad \text{subject to} \quad \mathbf{y} = \mathbf{\Psi D s} \ . \tag{3}$$

Even though solving the ℓ_0 minimization in (1) and (3) will produce the correct solution, this is an NP-hard problem and therefore impractical for moderate-sized scenarios. To address this issue, greedy methods, such as the orthogonal matching pursuit (OMP) [26], have been proposed among other approaches for solving (3). OMP tries in a greedy way to identify the elements of the dictionary that contain most of the signal energy by selecting iteratively the element of the dictionary exhibiting the highest correlation with the residual, and updating the current residual estimate. One of the main breakthroughs of the CS theory is that under the sparsity constraint and the incoherence of the sensing matrix, the reconstruction of the original signal, $\mathbf{x}$, and

the coefficient vector, $\mathbf{s}$, from $\mathbf{y}$, can be found by solving a tractable ℓ_1 optimization problem, the so-called *Basis Pursuit*, given by:

$$\min \|\mathbf{s}\|_1 \quad \text{subject to} \quad \mathbf{y} = \mathbf{\Psi D s} \ . \tag{4}$$

For compressible signals, the goal is not the exact reconstruction of the signal, but the reconstruction of a close approximation of the original signal. In this case, the problem is reduced to a *Basis Pursuit Denoising* and (4) takes the following form:

$$\min \|\mathbf{s}\|_1 \quad \text{subject to} \quad ||\mathbf{y} - \mathbf{\Psi D s}||_2 < \epsilon \ , \tag{5}$$

where ϵ is a bound on the residual error of the approximation, which is related to the amount of noise in the data. The optimization in (5) can be solved efficiently via the LASSO [27] algorithm for sparsity regularized least squares.

The number of measurements required for the signal reconstruction is dictated by the *mutual coherence* between the sensing matrix $\mathbf{\Psi}$ and the dictionary $\mathbf{D}$, which is defined as the maximum of the inner product between columns of the dictionary and the sampling matrix:

$$\mu(\mathbf{\Psi}, \mathbf{D}) \doteq \max_{\substack{1 \le i \le M \\ 1 \le j \le N}} \sqrt{N} \left| \langle \boldsymbol{\psi}_{\cdot,i}, \mathbf{d}_{j,\cdot} \rangle \right| \ , \tag{6}$$

where $\boldsymbol{\psi}_{\cdot,i}$ and $\mathbf{d}_{j,\cdot}$ denote the i-th column of $\mathbf{\Psi}$ and the j-th row of $\mathbf{D}$, respectively. For a specific mutual coherence, recovery is possible from $M \geq C \cdot \mu^2(\mathbf{\Psi}, \mathbf{D}) \cdot K \cdot \log(N)$ random measurements. As a consequence, having low coherence between the dictionary and the sampling matrix is beneficial in terms of performance.

C. Matrix completion

Matrix completion (MC) [28]–[30] is a recently proposed framework, which builds on the concepts of CS and extends the sparsity framework to the case of sub-sampled matrix-valued data. More specifically, MC considers a measurement matrix $\mathbf{M} \in \mathbb{R}^{N \times S}$, which can encode various CPS data, such as sensors measurements over time or spatial data in a single time instance, where a large number of its entries are missing. In general, one cannot recover the $N \cdot S$ entries of the matrix $\mathbf{M}$ from a smaller number of K entries, where $K \ll N \cdot S$, unless some characteristics about the measurement matrix are known. MC theory suggests that such a recovery is possible if the matrix is characterized by a rank smaller than its dimensions and a sufficient number of randomly selected entries of the matrix is available. The rank of the matrix indicates the number of linearly independent columns (or rows) and thus serves as a proxy to the correlations that exists within the data.

More specifically, one can recover an accurate approximation $\mathbf{X}$ of the matrix $\mathbf{M}$ from $K \geq C \cdot N^{6/5} \cdot r \cdot \log(N)$ random measurements, where $rank(\mathbf{M}) = r$, by solving the following minimization problem:

$$\min_{\mathbf{X}} rank(\mathbf{X}) \quad \text{subject to} \quad \mathcal{A}(\mathbf{X}) = \mathcal{A}(\mathbf{M}) \ , \tag{7}$$

where $\mathcal{A}$ is, in general, a linear map from $\mathbb{R}^{S \times N} \mapsto \mathbb{R}^K$. The theory of matrix completion suggests that recovery is possible when the linear map $\mathcal{A}$ is defined as a random sampling operator that records a small number of entries from the matrix $\mathbf{M}$, that is $\mathcal{A}_{ij} = \{1, \text{ if } (ij) \in \mathcal{S} \ \mid \ 0, \text{otherwise}\}$, where $\mathcal{S}$ is the sampling set. In the context of WSN sampling, the set $\mathcal{S}$ specifies the collection of sensors that are active at each specific sampling instance. In general, the solution of the MC problem requires the linear map $\mathcal{A}$ to satisfy a modified restricted isometry property, which is the case when uniform random sparse sampling is employed in both rows and columns of matrix $\mathbf{M}$ [31].

Unfortunately, the rank minimization in (7) is an NP-hard problem and therefore cannot be directly employed for data recovery. According to MC, one can resort to a relaxation capable of producing arbitrary accurate results, by replacing the rank constraint by the tractable nuclear norm, which represents the convex envelope of the rank. The minimization in (7) can then be reformulated as follows,

$$\min_{\mathbf{X}} \|\mathbf{X}\|_* \quad \text{subject to} \quad \mathcal{A}(\mathbf{X}) = \mathcal{A}(\mathbf{M}) \ , \tag{8}$$

where the nuclear norm is defined as $\|\mathbf{X}\|_* = \sum \|\lambda_i\|_1$, that is, the sum of absolute values of the singular values. For the noisy case, an approximate version is given by

$$\min_{\mathbf{X}} \|\mathbf{X}\|_* \quad \text{subject to} \quad \|\mathcal{A}(\mathbf{X}) - \mathcal{A}(\mathbf{M})\|_F^2 \leq \epsilon \ , \tag{9}$$

where $\|\mathbf{X}\|_F^2 = \sum \lambda_i^2$ denotes the Frobenius norm and ϵ is the approximation error. To solve the nuclear norm minimization problem (8) and (9), various distinct approaches have been proposed, including the singular value thresholding (SVT) [32], the augmented Lagrange multipliers (ALM) [33], and the so-called OptSpace [34]. In the following, the technique based on the ALM is reviewed briefly, since it has been shown to offer exceptional performance both in terms of processing complexity and reconstruction accuracy, and because it serves as a basis for our extended scheme.

D. Applications in iCPS

CS and MC have been successfully employed in various tasks related to iCPS data acquisition, processing, and management. This success can be attributed to various characteristics of these algorithms. Both CS and MC employ lightweight encoders, while shifting the computational complexity and the associated resources to the decoder side. Furthermore, both CS and MC offer scalable signal recovery capabilities, where more measurements contribute positively to the reconstruction performance. The benefits of CS have been explored for efficient compression and transmission of many complex cyber-physical data, such as video and audio in wireless multimedia sensor networks [35], [36], vehicle information in vehicular networks [37], and ECG in wireless body sensor networks [38] to name a few. Moreover, CS offers the ability to perform independent encoding and joint decoding of the data, while MC does not require a specific sampling architecture but instead relies on the random sub-sampling of the measurements themselves.

- *iCPS Data Sampling & Compression:* CPS data acquisition is a prominent case, where intelligent signal processing can greatly support network operation and increase the usability of the sensing infrastructure. The concept of MC has been successfully employed in the efficient sampling of the spatio-temporal iCPS data acquired by WSNs [39], [40], as well as data from Internet-of-Things platforms [41]. For instance, [42] investigates the co-design of the sampling pattern with the network channel access for recovering spatio-temporal fields monitored by WSNs. MC has been also considered for the coupled reconstruction of missing measurements and data classification of WSN data [43], where it was shown that both objectives can be achieved through the introduction of a dictionary in the low-rank matrix estimation process. Recently, a robust compression scheme was introduced and evaluated on real WSN data [44] based on the introduction of both CS and MC for data compression and recovery of lost packets. On the other hand, certain properties of CS, such as the lightweightness and the universality of the encoding stage, make it a good candidate for efficient distributed data compression in WSNs [45], [46]. In this setting, spatial transformations including DFT, graph-wavelets and diffusion-wavelets can be utilized for storage and retrieval of network data. CS was also investigated as a rateless distributed coding scheme, offering reduced communications cost independently of the routing algorithm and the network topology [47]. A CS-guided architecture for decentralized recovery of sparse signals in WSNs was proposed in [48], where the authors considered a random node sleeping pattern, in conjunction with a consensus algorithm for achieving global signal reconstruction from the local estimates. Introducing CS into data sampling and compression has also been supported by novel hardware architectures [49].
- *Aggregation & Routing:* CS has been also applied recently for data aggregation in multi-hop WSNs, where typically the objective is to collect the full set of measurements to a centralized location, such as a sink node or a gateway. CS-driven data aggregation techniques utilize a random encoding process as nodes forward measurements to a central processing unit, reducing the amount of packets that have to be communicated [50], [51]. While CS requires a specific random encoding process, MC, which relies on randomized sub-sampling, has been recently explored as an alternative data gathering scheme [52], where MC was supported from a temporal prediction process to recover completely empty measurement vectors. The authors also provided evidence that a large class of CPS signals, such as temperature and

humidity, are indeed characterized by the low rank property, while other signals, such as illumination conditions, cannot be reliably approximated by low-rank measurement matrices.
The efficient interaction between CS encoding/decoding and routing in WSNs was investigated in [53], where the authors showed that the high coherence between the data sparsifying transform and the routing can limit significantly the straightforward applicability of CS in networked data. A remedy to this issues was proposed in [54], where an optimal adaptive forwarding scheme was considered for network lifetime maximization. The CS framework has also been combined with another communications paradigm, the network coding. NetCompress [55], for instance, proposes the simultaneous transmission of measurement packets and the encoding via the random projection step of CS.

- *Sensor Localization:* Localization information is an important task in a large class of iCPS monitoring scenarios in both outdoors and indoors scenarios. For instance, one can consider the situation where sensing nodes are mobile, or the case where the data collection unit possesses mobility capabilities Cutting-edge signal processing paradigms, such as CS and MC can offer substantial benefits with respect to training time, positioning accuracy, algorithmic complexity and adaptability. One class of approaches is founded upon the low rank property of the Euclidean distance matrix, and employ MC for the recovery of the complete set of distances from a small number of noisy measurements [56]–[58] and allow tracking of mobile devices [59]. Another class utilizes CS and MC in order to improve the efficiency of fingerprinting approaches in challenging environments by reducing the training requirements and offering dynamic adaptation mechanisms [60], [61].
- *iCPS Data Classification & Event Detection:* Recently, a combination of low rank and sparse signal recovery was introduced for traffic anomaly detection in large scale networks [62], distributed temporal pattern detection in WSNs [63], and target localization and counting [64] among others. The concept of decentralized estimation of missing data has also been investigated in [65] and is applicable to WSNs.

E. Potential of intelligent signal processing in iCPS

Despite their youth, CS and MC have shown great potential in iCPS data acquisition and processing via WSNs. Performance gains have been observed in various aspects including sampling, compression, routing, and detection to name a few. The connection between WSN and high-performance distributed computing platforms such as *cloud computing* can serve as an excellent paradigm for the next generation of CPS architectures. Issues related to the coupled design of such infrastructure include the robust distributed storage and the practical implementation of linear random sampling measurements acquisition.

When considering iCPS, one must not only consider the sensing platform, but the actuation components as well. In wireless sensing and actuation networks, the physical environment is responsible for closing the loop between sensing and actuation. Application of CS and MC for control of CPS processes remains an open scientific and technical challenge that calls for immediate attention.

IV. In-Network Processing: Distributed Estimation and Tracking for iCPS

Throughout this section, we present how various in-network signal processing tasks, such as parameter estimation and tracking, can be accomplished in an industrial environment. Since these approaches are based on *iterative average consensus*, we focus in the distributed implementation of this specific approach, while also considering the generic case of *non-uniform* deployment of the sensing devices. A distributed framework eliminates the need of performing all the computations at one or more sink nodes, thus, reducing congestion around them and increasing the robustness of the WSN. Moreover, we study how the randomness and asymmetry of instantaneous communications, occurred in real iCPS, affect the performance of both estimation and tracking tasks. To alleviate this impact, we present a cross-layer approach based on a link scheduling protocol that deals with the particularities of the industrial environment, providing a suitable framework for the in-network processes to be executed with a reduced error.

A. Statistical signal processing background

In general, the sensors that compose an iCPS are covering the monitored area by being placed on a non-uniform grid, whilst observing a distorted version of the target data, which is usually corrupted by random

noise. Let $\mathbf{y} = f(\mathbf{x}, \mathbf{w})$ be the vector whose elements are the distorted observations of the S nodes, $\mathbf{x} \in \mathbb{R}^M$ is the parameter vector of interest, and $\mathbf{w} \in \mathbb{R}^S$ is a random vector modeling the noise component. Then, the data observed by the sensors can be expressed as follows,

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w} \ , \tag{10}$$

where the observation matrix $\mathbf{H}$ models the spatial distortion of the data, and $\mathbf{w}$ is a spatially uncorrelated additive noise, which is modeled as a zero-mean Gaussian with covariance matrix $\mathbf{Q}_w = \sigma^2\mathbf{I}$, where $\mathbf{I}$ denotes the identity matrix. Each component of the vector $\mathbf{s} \triangleq \mathbf{H}\mathbf{x}$ can be viewed as the signal or field present at the location of each node. Based on these noisy and random observations, the iCPS needs to draw conclusions and actuate accordingly within the control loop. Thus, the more accurate the estimation of the target data is, the better this actuation can be performed. In general, the estimation process can be viewed as a problem of data selection from a continuous space that minimizes a certain cost function.

If the data to be estimated is a time-invariant quantity, the process is reduced to a *parameter estimation* problem. On the other hand, if the data evolves in time according to a stochastic equation, the process corresponds to a *state estimation* problem. Although the term *tracking* may be applied only to the specific estimation of the state of a moving object, throughout this section we will use the terms state estimation and tracking indistinctly.

For the state estimation problem, the objective is to estimate the value of a deterministic data vector as accurately as possible, without relying on previous information. The general approach is to maximize the probability density function (PDF) of the observations conditioned on the data, which is simply the maximum likelihood estimate given by

$$\mathbf{x}_{\text{ML}} \triangleq \underset{\mathbf{x}}{\operatorname{argmax}}\, p(\mathbf{y}|\mathbf{x}) \ .$$

If we assume the linear Gaussian model of (10) for the observations, it can be shown that the maximum likelihood estimator is obtained as follows [66]

$$\mathbf{x}_{\text{ML}} = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{y} \ . \tag{11}$$

Furthermore, it can be shown that this estimator is unbiased, that is, $\mathbb{E}[\mathbf{x}_{\text{ML}}] = \mathbf{x}$, with covariance matrix

$$\mathbf{C}_{\text{ML}} = \sigma^2(\mathbf{H}^T\mathbf{H})^{-1} \ . \tag{12}$$

This estimation process can be explained in a geometrical sense; specifically, from (11), we can deduce that the optimal estimation of $\mathbf{s}$, given the observations $\mathbf{y}$, is the projection of $\mathbf{y}$ onto the subspace spanned by $\mathbf{H}$, with the projection matrix defined as follows,

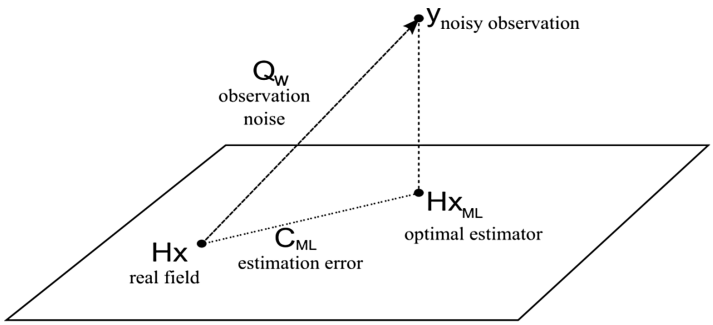
$$\mathbf{P} = \mathbf{H}(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T$$

Moreover, if $\mathbf{H}^T\mathbf{H} = \mathbf{I}$, that is, if $\mathbf{H}$ is an orthonormal matrix, then, the estimation process is the orthogonal projection of the observation vector onto the above-mentioned subspace with $\mathbf{P} = \mathbf{H}\mathbf{H}^T$. Figure 4(a) illustrates this geometrical representation for $M = 2$. Besides, it can be seen that the estimator is always improving the initial observations, which is stated formally as $\operatorname{trace}(\mathbf{C}_{\text{ML}}) = M\sigma^2 \leq \operatorname{trace}(\mathbf{Q}_w) = S\sigma^2$, as long as the dimension of the target vector M is smaller than the number of observations S.

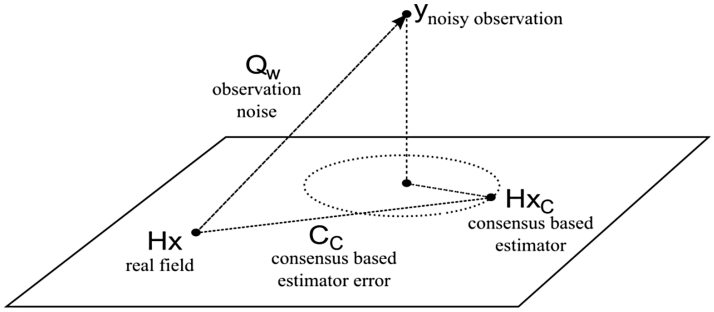
On the other hand, if the parameters to be estimated evolve discretely and stochastically across time, the parameter estimation problem reduces to a state estimation process. To this end, successive observations are acquired, $\mathbf{y}[k] = \mathbf{H}[k]\mathbf{x}[k] + \mathbf{w}[k]$, and the prior information ($p(\mathbf{x})$) about the evolution of the process is used to refine the estimations in a Bayesian framework. Specifically, the system is assumed to evolve according to a Markov-Gaussian model as follows,

$$\mathbf{x}[k+1] = \mathbf{A}[k]\mathbf{x}[k] + \mathbf{v}[k] \ , \tag{13}$$

where $\mathbf{x}[k]$ is the state vector at time k, $\mathbf{A}[k]$ is a $M \times M$ time-varying matrix that rules the evolution of the state, and $\mathbf{v}[k]$ is the noise of the system, which is considered to be white, Gaussian and uncorrelated with $\mathbf{w}[k]$, with covariance matrix $\mathbf{Q}_v[k]$. Under the assumption of a Markov-Gaussian model, the optimal



(a)



(b)

Fig. 4. Geometrical interpretation of the parameter estimation when the observation matrix is orthonormal, for $M = 2$. (a) the optimal estimation is the orthogonal projection of the observations onto the subspace; (b) the projection is computed in a distributed way by means of iterative average consensus, yielding a sub-optimal estimator.

estimator can be computed recursively by means of recursive least-squares. This is obtained via the dynamics of a Kalman filter [67], which is given by

$$\hat{\mathbf{x}}[k+1] = \mathbf{A}[k]\hat{\mathbf{x}}[k] + \mathbf{K}[k]\left(\mathbf{y}[k] - \mathbf{H}[k]\mathbf{A}[k]\hat{\mathbf{x}}[k]\right) \ . \tag{14}$$

The gain, $\mathbf{K}[k]$, of the filter is given by

$$\mathbf{K}[k] = \mathbf{M}[k]\mathbf{H}^T[k]\mathbf{Q}_w^{-1}[k] \ , \tag{15}$$

where $\mathbf{M}[k]$ is defined as

$$\mathbf{M}^{-1}[k] = \mathbf{P}^{-1}[k] + \mathbf{H}^T[k]\mathbf{Q}_w^{-1}[k]\mathbf{H}[k] \ . \tag{16}$$

In the above expression, $\mathbf{P}[k]$ denotes the covariance error of the estimator, whose dynamic evolution is described by

$$\mathbf{P}[k+1] = \mathbf{A}[k]\mathbf{M}[k]\mathbf{A}^T[k] + \mathbf{Q}_v[k] \ . \tag{17}$$

The basis of this filter is to achieve a trade-off between the optimal state estimator (MLE) computed with each new observation and the previous estimation, or between the previous estimation and the innovation computed from the new observations. The weight assigned to each term of the trade-off is given by the gain of the filter, which is computed in order to minimize the covariance error of the estimator at each time step. As stated before, this methodology can be tackled in the general case of non-uniformly deployed sensing devices, working in a distributed fashion by employing an iterative consensus algorithm. However, an industrial environment imposes certain communication constraints that must be taken into account in a real implementation.

B. Distributed average consensus under realistic constraints

The harsh environmental conditions of an industrial scenario provoke random packet losses in the communications between the non-uniformly located iCPS devices, thus affecting the performance of consensus-based applications, such as estimation and tracking. This performance loss is mainly due to the resulting randomness and asymmetry of the links, which affect the convergence time and error of the process.

Let $\mathcal{S}$ be a set of S autonomous nodes with initial measurements $x_i[0]$, $i = 1, \ldots, S$ following a normal distribution with mean x_{avg} and variance σ^2. Then, the *distributed consensus* (or agreement) problem consists of successive iterations, where each node i refines its own value by exchanging information only with those nodes belonging to the set of its neighbors $\mathcal{S}_i$. This procedure continues until the nodes agree asymptotically on a global common value α, where the asymptotics is expressed in terms on infinite time, $\lim_{k\to\infty} \mathbf{x}[k] = \alpha\mathbf{1}$, where $\mathbf{1} \in \mathbb{R}^N$ is the vector of all ones. Let $\mathbf{W}$ be the weight matrix that rules the mixing of information at each iteration. Then, the state evolution is expressed by the following process,

$$\mathbf{x}[k] = \mathbf{W}\mathbf{x}[k-1] = \mathbf{W}^k\mathbf{x}[0] = \mathbf{M}_k\mathbf{x}[0] \ ,$$

with $[\mathbf{W}]_{ij} \neq 0$ if and only if $j \in \{\mathcal{S}_i \cup i\}$. Moreover, the Perron-Frobenius theorem states that if $\mathbf{W}$ is row-stochastic, that is, $\mathbf{W}\mathbf{1} = \mathbf{1}$, and irreducible, then $\lim_{k\to\infty} \mathbf{W}^k = \frac{\mathbf{1}\mathbf{q}^T}{\mathbf{q}^T\mathbf{1}}$, where $\mathbf{q}$ is the left eigenvector of $\mathbf{W}$ corresponding to the eigenvalue 1. Consequently, all rows of the matrix $\mathbf{M}$ are asymptotically equal to a vector $\mathbf{m}$, where $m_i = q_i / \sum_{i=1}^{S} q_i$. Therefore, the nodes achieve a consensus, which corresponds to the value

$$\alpha = \sum_{i=1}^{S} m_i x_i[0] \ . \tag{18}$$

If, in addition, $\mathbf{W}$ is column-stochastic, that is, $\mathbf{1}^T\mathbf{W} = \mathbf{1}^T$, then $\mathbf{m} = \frac{1}{S}\mathbf{1}$, and $\alpha = \frac{1}{S}\sum_{i=1}^{S} x_i[0]$, which is exactly the average of the initial values. Nevertheless, since the weight matrix $\mathbf{W}$ should be compatible with the underlying topology of the network, we have that in a realistic industrial scenario, where interference, fading and packet losses may occur, each instantaneous topology is totally random, and, in general, different. Based on that, we also consider instantaneous matrices, $\mathbf{W}[k]$, in the time-evolving state equation:

$$\mathbf{x}[k] = \mathbf{W}[k] \ldots \mathbf{W}[0]\mathbf{x}[0] = \mathbf{M}[k]\mathbf{x}[0] \ . \tag{19}$$

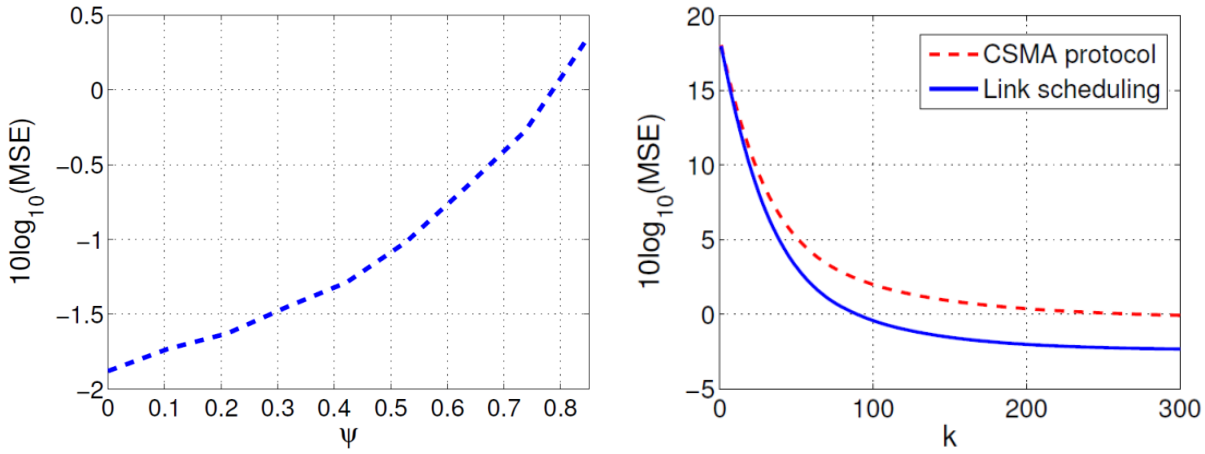
By construction, we can force every $\mathbf{W}[k]$ to be row-stochastic and guarantee the nodes to reach a consensus, $\lim_{k\to\infty} \mathbf{M}[k] = \mathbf{1}\mathbf{m}^T$. However, in this case $\mathbf{m}$ becomes a random vector which is, in general, different from $\frac{1}{S}\mathbf{1}$, and whose first two moments are computed in [68]. The moments of $\mathbf{x}[k]$ can be asymptotically computed as follows,

$$\begin{aligned} \mathbb{E}[\mathbf{x}] &= x_{\text{avg}}\mathbf{1} \\ \mathbf{C}_x &= \sigma^2\mathbb{E}[\mathbf{m}^T\mathbf{m}]\mathbf{1}\mathbf{1}^T \ . \end{aligned} \tag{20}$$

Thus, $\lim_{k\to\infty} x_i[k]$ can be viewed as the unbiased estimator of x_{avg} computed by node i, with variance $\sigma_x^2 = \sigma^2\mathbb{E}[\mathbf{m}^T\mathbf{m}]$.

If $\mathbf{\Sigma}$ is the matrix whose entry ij is the activation probability of the link between nodes i and j, then the variance σ_x^2 can be reduced by enforcing $\mathbf{\Sigma}$ to be as symmetric as possible. This is motivated by the fact that its symmetry entails the symmetry of matrix $\mathbb{E}[\mathbf{W}[k]]$, yielding $\sum_{i=1}^{S} \mathbb{E}[m_i]^2 = \frac{1}{S}$, which corresponds to the minimum possible value. Figure 5(a) shows how the mean squared error (MSE), which is equal to the variance since the estimator is unbiased,, increases with the asymmetry of $\mathbf{\Sigma}$. This means that the randomness and asymmetry of communications, occurring in an industrial environment due to packet losses, affect the error of the consensus.

The approach described in [69] deals with this imperfection of communications, ensuring, on average, the symmetry of the links. This implies having less consensus error than by applying traditional approaches, such as the protocol implemented by default in most of the motes, which follows a CSMA strategy and, as a consequence, it only focus on reducing collisions. Moreover, the protocol introduced in [69] employs connectivity patterns that are as dense as possible, so as to favor the convergence time, which is also crucial towards enabling the iCPS to actuate as fast as possible. The implementation of this new protocol is based



(a) (b)

Fig. 5. (a) MSE of the average consensus as a function of the asymmetry of the connection probability matrix. (b) Evolution of the MSE across the iterations of the consensus process. The proposed link scheduling protocol outperforms the general CSMA when computing distributed average consensus. The parameter ψ represents the asymmetry of the underlying graph.

on a cross-layer scheme in which the decisions taken by the MAC layer about whether to transmit or not, besides providing collision avoidance, favors the performance of the consensus process. To this end, at each scheduling step, the protocol activates randomly a link and creates an associated inhibition area that contains the links that are inhibited when the current link is activated. In order to ensure that this inhibition radius guarantees a collision-free communication pattern, the worst case scenario is assumed: 1) every transmitter is at maximum distance from its intended receiver, and 2) every interferer is assumed to be as close as possible to the receivers. By locating these inhibition areas at the center of each link, every pair of nodes includes the same number of potential inhibitors inside those areas, leading to the same probability of inhibition. Moreover, since every link in the network presents the same probability of being considered for activation, symmetric probabilities of connection for each pair of nodes are also ensured.

Figure 5(b) shows the efficiency of this protocol, which outperforms the general CSMA protocol when computing the average consensus. This cross-layer approach is applied in the following sections to reduce the error of the proposed in-network processing algorithms for iCPS, which are all based on iterative average consensus.

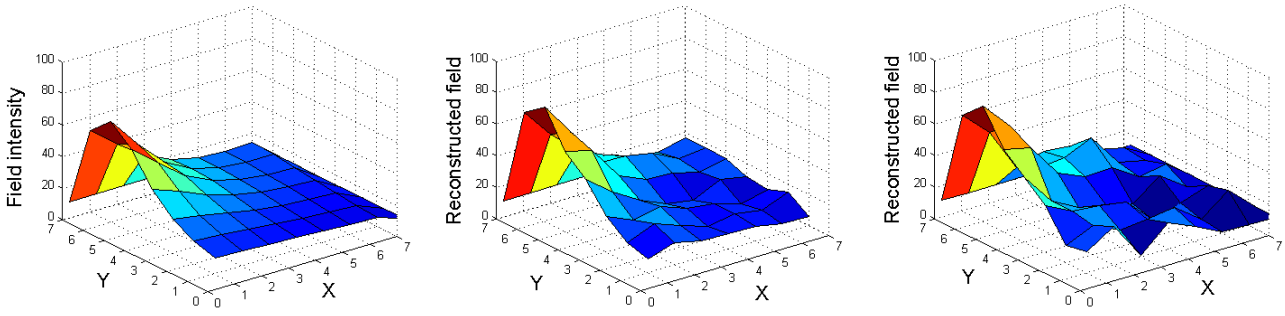
C. Consensus-based in-network processing

The cross-layer protocol proposed in the previous section supports the implementation of consensus-based estimation techniques in iCPS with reasonably low error. In this section, we describe in detail how the distributed parameter estimation and the distributed state estimation techniques are both favored by the application of this cross-layer technique, demonstrating its performance in the case of reconstructing a generic two-dimensional field.

1) Distributed parameter estimation: The maximum likelihood estimator given by (11) can be also expressed as follows,

$$\mathbf{x}_{\text{ML}} = \left(\sum_{i=1}^{S} \mathbf{h}_i \mathbf{h}_i^T \right)^{-1} \sum_{i=1}^{S} \mathbf{h}_i y_i \ , \tag{21}$$

where $\mathbf{h}_i$ is the i-th row of $\mathbf{H}$. It is straightforward to show that each node is able to compute both the matrix $\frac{1}{S}\sum_{i=1}^{S} \mathbf{h}_i \mathbf{h}_i^T$ and the vector $\frac{1}{S}\sum_{i=1}^{S} \mathbf{h}_i y_i$ in a distributed fashion by means of two iterative average consensus processes, and consequently to compute the ML estimate asymptotically. We emphasize again that our proposed approach does not make any assumption for the network topology, thus it can be used successfully in both uniform and non-uniform sensor deployments. Nevertheless, due to the randomness of



(a) Real field

(b) Reconstructed field using the link scheduling protocol proposed in [69]

(c) Reconstructed field using a generic CSMA protocol

Fig. 6. Two-dimensional field reconstruction using distributed parameter estimation based on iterative average consensus. The nodes form a 7×7 grid.

the iterative processes, there exists a deviation from the average, with the actual estimator being given by

$$\mathbf{x}_c \triangleq \left(\sum_{i=1}^{S} m_i \mathbf{h}_i \mathbf{h}_i^T \right)^{-1} \sum_{i=1}^{S} m_i \mathbf{h}_i y_i = (\mathbf{H}^T \Delta_m \mathbf{H})^{-1} \mathbf{H}^T \Delta_m \mathbf{y} \ , \tag{22}$$

where $\mathbf{m}$ is a random vector, and Δ_m is the diagonal random matrix with the elements of $\mathbf{m}$ on its main diagonal. The covariance matrix of this estimator takes the following form:

$$\mathbf{C}_c = \sigma^2 \mathbb{E}\left[(\mathbf{H}^T \Delta_m \mathbf{H})^{-2} \mathbf{H}^T \Delta_m^2 \mathbf{H} \right] \ . \tag{23}$$

The expression of $\mathbf{x}_c$ in (22) may be seen as a noisy (random) version of the expression of $\mathbf{x}_{\text{ML}}$ in (11), since the deviation from the average in the iterative consensus involves an additional random error. This is shown in Figure 4(b), where following the geometrical interpretation of the estimation, the consensus-based estimation error is the sum of the optimal estimation error and the consensus error. It can be shown that $\text{trace}(\mathbf{C}_c) \geq \text{trace}(\mathbf{C}_{\text{ML}})$, where equality holds if and only if the average in both consensus processes is always achieved. Therefore, the consensus-based estimator is sub-optimal in probabilistic terms. In fact, it may be the case that the consensus process is so inaccurate that the consensus-based estimator is worse than the initial observations, that is, $\text{trace}(\mathbf{C}_c) \geq \text{trace}(\mathbf{Q}_w) = \sigma^2 N$.

However, interestingly, from Figure 4(b) it can be also seen that there exists a possibility that some realizations of the consensus-based estimator improve the MLE. Specifically, for $M = 1$ the probability is 0.50, but it decreases quickly as M increases. The performance of the consensus-based estimator can be improved if the connection probability matrix approximates a symmetric matrix. To accomplish that, the link scheduling protocol defined in [69] can be applied in a cross-layer scheme at the link layer, instead of the standard CSMA protocol, when a distributed parameter estimation is performed using iterative consensus. As an illustration, Figure 6 shows the distributed reconstruction of a two-dimensional field by a network of $S = 49$ sensors deployed uniformly on a grid structure. Notably, it can be seen that the symmetry of $\mathbf{\Sigma}$ improves the performance of the estimator.

2) Distributed state estimation: For the distributed computation of the Kalman filter notice that (14) can be expressed as [70]

$$\hat{\mathbf{x}}[k+1] = \mathbf{A}[k]\hat{\mathbf{x}}[k] + \frac{1}{\sigma^2}\mathbf{M}[k]\left(\mathbf{H}^T[k]\mathbf{y}[k] - \mathbf{H}^T[k]\mathbf{H}[k]\mathbf{A}[k]\hat{\mathbf{x}}[k]\right) \ ,$$

where $\mathbf{Q}_w[k] = \sigma^2 \mathbf{I}$. The necessary computations for the distributed implementation of this filter are very similar to the ones performed in the parameter estimation problem. In fact, at each time k, every node has to compute $\mathbf{H}^T[k]\mathbf{y}[k] = \sum_{i=1}^{S} \mathbf{h}_i y_i$ and $\mathbf{H}^T[k]\mathbf{H}[k] = \sum_{i=1}^{S} \mathbf{h}_i \mathbf{h}_i^T$, which are the two terms that are present in (21). Furthermore, each node has also to compute $\mathbf{M}[k]$ in order to obtain the gain of the filter, and perform the corresponding weighting between the optimal estimation at the current time instant and the

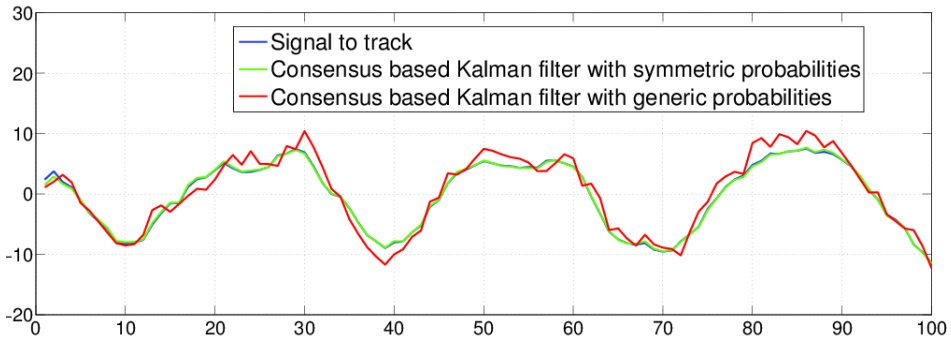


Fig. 7. Consensus-based distributed Kalman filter. The performance of the filter obtained via the algorithm in [69] is compared against the performance of the filter obtained when generic CSMA is used.

previous estimation. From (16) and (17) we deduce that for the computation of $\mathbf{H}^T[k]\mathbf{H}[k]$ it suffices for each node to compute $\mathbf{M}[k]$ in a distributed way at time k. Thus the estimator at time k is expressed as a function of the individual observations as follows,

$$\hat{\mathbf{x}}[k+1] = \mathbf{A}[k]\hat{\mathbf{x}}[k] + (\sigma^2 S\mathbf{P}^{-1}[k] + \sum_{i=1}^{S} \mathbf{h}_i\mathbf{h}_i^T)^{-1} \left(\sum_{i=1}^{S} \mathbf{h}_i y_i - \sum_{i=1}^{S} \mathbf{h}_i\mathbf{h}_i^T \mathbf{A}[k]\hat{\mathbf{x}}[k] \right) , \tag{24}$$

which can be computed in a completely decentralized way. Again, in practice we have to consider a deviation from the average in the computation of the iterative consensus, thus obtaining a sub-optimal filter,

$$\hat{\mathbf{x}}[k+1] = \mathbf{A}[k]\hat{\mathbf{x}}[k] + (\sigma^2\mathbf{P}^{-1}[k] + \mathbf{H}^T[k]\Delta_m\mathbf{H}[k])^{-1} \left(\mathbf{H}^T[k]\Delta_m\mathbf{y}[k] - \mathbf{H}^T[k]\Delta_m\mathbf{H}[k]\mathbf{A}[k]\hat{\mathbf{x}}[k] \right) .$$

Concerning the performance of this filter, all the statements presented in the previous section with respect to the parameter estimation are also valid here. Consequently, the performance of this sub-optimal filter can be improved significantly by employing the same cross-layer scheme described in [69], which enforces $\mathbf{\Sigma}$ to approximate a symmetric matrix, as shown in Figure 7.

V. High-Level Data Analysis and Early Warning

In an industrial CPS setting, the associated distributed autonomous sensing introduced before, is further exploited to produce intelligent reasoning over the data by supporting advanced operations, such as, querying, high-level analysis, and alerting. In particular, a *high-level data management and analysis* (HDMA) module is an integral part towards an efficient decision making. Typically, an HDMA component comprises of collaborating computational nodes, which observe and control distinct physical entities and dynamic phenomena. Rather than relying on single sensor stream statistics, such as average and standard deviation, which is the customary approach in most data analysis systems, an efficient HDMA module focuses on finding and extracting inherent information for detecting behavioral variations in the acquired data. This is crucial, especially in an industrial CPS framework, since the accurate and timely detection of abnormal changes in sensor measurements will enable early actuation aiming at minimizing operational and maintenance costs.

Usually WSN nodes do not handle any quality aspect of physical device data, but rather interface with a high-level representation of the sensed physical world. In practice, the recorded sensor measurements are often incomplete, imprecise, or even misleading, thus impeding the task of an accurate and reliable decision making. Motivated by this, a powerful HDMA system should also cope with what we call *uncertain* data. *Uncertainty-aware data management* [71] presents numerous challenges in terms of collecting, modeling, representing, querying, indexing and mining the sensor data. Since many of these issues are interrelated, we address them jointly wherever possible. In contrast to most of the existing industrial CPS, a versatile HDMA module considers uncertainty as an additional source of information that could be valuable during data analysis and thus it should be preserved.

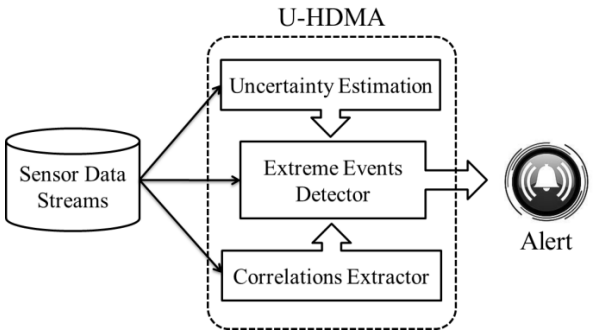


Fig. 8. Building blocks of an uncertainty-aware high-level data analysis system.

Another major functionality assigned to a modern data analysis system is to perform *high-level operations*, such as the notification of extreme events from raw sensor data. Since the detection of abnormal behavior is affected by the underlying uncertainty, incorporation of its estimated value is expected to yield more meaningful results. More specifically, widely used methods for extreme events detection can be enhanced by incorporating the inherent data uncertainty, yielding an integrated uncertainty-aware HDMA (U-HDMA) system capable of identifying, quantifying, and combining the individual uncertainties corresponding to the most significant sources of uncertainty for providing *early warning notifications* of extreme events.

On the other hand, extracting *highly correlated* pairs of data streams acquired by distinct sensors is another important issue. Doing so, we aim at revealing interrelations between seemingly independent physical quantities, or guaranteeing the validity of a detected extreme event. Whereas traditional statistical machine learning provides well-established mathematical tools for monitoring and analyzing multiple data streams by exploiting potential pairwise correlations [72], [73], their performance is limited when processing heterogeneous and uncertain data streams. More specifically, [74] studies the problem of maintaining data stream statistics over sliding windows, with the focus being only on single stream statistics, while [75] introduced an extension for monitoring the statistics of multiple data streams, but the computation of correlated aggregates is limited to a small number of monitored sensor streams. On the other hand, [76] introduced a successful data stream monitoring system, which enables the computation of single- and multiple-stream statistics. However, its performance diminishes in an industrial environment, since the sensor streams we manage describe dynamic phenomena, whose distribution is not known a priori. Such limitations of previous approaches can be overcome by designing an appropriate stream correlation engine based on a computationally efficient similarity function, which enables fast and accurate monitoring of pairwise correlations between time-synchronized (high-dimensional) sensor data streams.

Figure 8 summarizes our utmost goal in this section, which is to provide an insight into the design and implementation principles of efficient and robust U-HDMA systems, integrating the above functionalities for industrial monitoring and surveillance applications, while emphasizing the importance of accounting for the underlying data uncertainty as an additional source of information, which should be preserved across all stages of the data processing chain. In particular, a generic U-HDMA module consists of the three building blocks shown in Figure 8, namely, (i) uncertainty estimation, (ii) correlations extraction, and (iii) detection of extreme events. Appropriate data services are provided to manipulate the sensor measurements, as well as to characterize the generated data quality. Computationally efficient extraction of correlations from uncertain data streams is then coupled with modified uncertainty-aware extreme event detectors to enable higher-level analysis, which form the basis for the development of an integrated U-HDMA system for monitoring dynamic sensor networks and providing early warning notifications in case of abnormal events.

A. Managing uncertainty in sensor measurements

In practice, the raw sensor data acquired by distinct sensors distributed across an industrial infrastructure are often unreliable, imprecise, or even misleading. This yields results of unknown quality, which may impede the task of an accurate and reliable decision making. To this end, the notion of *measurement uncertainty* arises

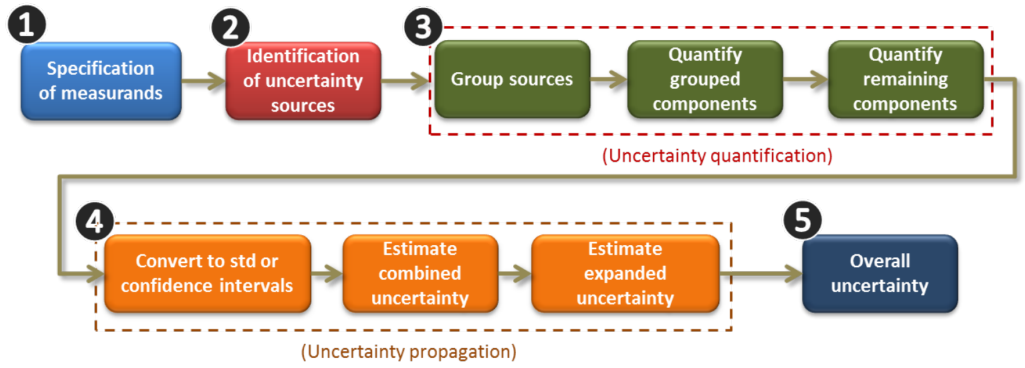


Fig. 9. Flow diagram for uncertainty estimation in sensor data streams.

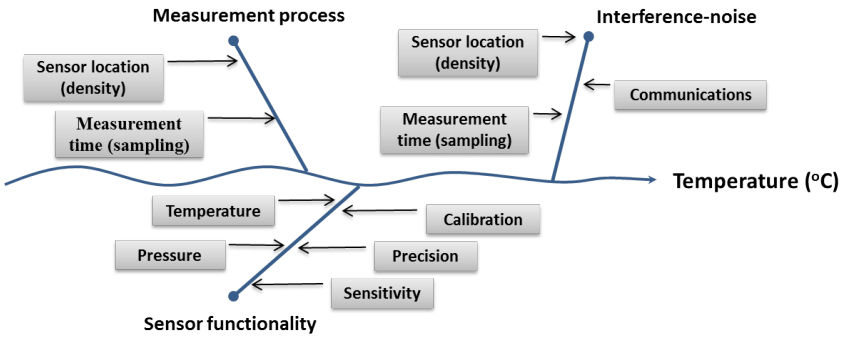


Fig. 10. Cause and effect diagram for a temperature sensor.

as an indicator of measurement quality. Speaking formally, the uncertainty is *"a parameter associated with the result of a measurement that characterizes the dispersion of the values that could reasonably be attributed to the measurand"*, where a measurand refers to a quantity to be measured.

The underlying uncertainty may arise due to several distinct sources, such as, an incomplete definition of the observed quantities, sampling effects and interferences, varying environmental conditions, or hardware defections of the equipment. The effects of all these factors can be observed and quantified from the recorded sensor data only. For this purpose, a set of ordered steps need to be performed in order to obtain an estimate of the uncertainty associated with a measurement result. Figure 9 presents the processing flow, which starts by identifying the measurands to be monitored and returns the overall estimated uncertainty.

Having specified appropriate measurands associated with our industrial application, such as, temperature (°C), pressure (bar), capacitance (F), and current (A), the next step is to identify the potential sources of uncertainty. A very convenient way to determine the most dominant uncertainty sources, along with their potential interdependencies, is to exploit the so-called *cause and effect* (or Ishikawa) diagram [77]. This diagram also ensures comprehensive coverage, while helping to group similar sources and avoid double counting. Figure 10 shows a typical cause and effect diagram for a temperature sensor. Its performance may be affected by several distinct factors, such as, its sensitivity and precision, calibration, and operating temperature. Furthermore, the accuracy of the acquired measurements depends also on the deployment density and location of the sensors, as well as on the sampling process. Possible misplacement or a very sparse time-sampling is expected to increase the uncertainty, especially when the monitored variables vary rapidly across time.

Despite the pervasive nature of computational analysis in nowadays engineering practice, an objective establishment of the confidence levels in the measurement procedure, as well as in the subsequent numerical processing, still remains a difficult task. This is due to the differences between a real device and the corresponding numerical models, and the lack of knowledge associated with the underlying physical processes.

Uncertainty quantification, which is the third step in the estimation chain, plays a fundamental role aiming at developing a rigorous framework to characterize the impact of variability and lack-of-knowledge on the final quantity of interest and provides the basis for certification in high-consequence decisions. However, it is important to notice that not all of the sources will make a significant contribution to the combined uncertainty. In practice, it is often likely that only a small number of them will contribute the major portion of the overall uncertainty. If possible, an initial estimate of the contribution of each separate source, or groups of sources, to the uncertainty could be made, so as to eliminate the less significant ones.

Towards assessing the underlying uncertainty component in a given raw sensor data stream we recall its distinction into two separate categories, namely, *type A* (aleatoric, statistical, or irreducible) and *type B* (epistemic, systematic, or reducible) uncertainty [78]. For instance, the physico-chemical properties of substances concentration, the operating conditions of the sensors and their manufacturing tolerances are typical examples associated with type A uncertainties, which cannot be reduced. On the other hand, the mathematical models, the calibration methods, and the inference techniques from experimental observations are typical sources of type B uncertainties, which can be reduced by improving the accuracy of our physical models or calibration methods.

Without going through much detail, in the following we introduce the main approaches for carrying out the steps 3 and 4 in Figure 9. Specifically, uncertainties of type A are characterized by the estimated variances σ_i^2 (or the standard deviations σ_i), which are obtained by statistical analysis of the observations in the raw sensor data streams. This is equivalent to obtaining a *standard uncertainty* from a probability density function (pdf) derived from an observed frequency (empirical) distribution. Let $\mathbf{y} = \{y_1, \ldots, y_N\}$ be a set of N sensor measurements, which correspond to a specific observed variable. Then, the standard uncertainty of $\mathbf{y}$, which is denoted by $u(\mathbf{y})$, is expressed in terms of the corresponding standard deviation σ_y, estimated directly from the observations y_i, as follows,

$$u(\mathbf{y}) = \frac{\sigma_y}{\sqrt{N}} \ . \tag{25}$$

For uncertainties of type B, the estimated "variance" s_j^2 is obtained from an assumed probability density function based on our prior knowledge for the corresponding source of uncertainty, which may include: a) data from previous measurements; b) experience or knowledge of the properties of instrumentation and materials used; c) manufacturers specifications; and d) calibration data. In general, concerning type B uncertainties, the quantification is performed either by means of an external information source, or from an assumed distribution.

Typical assumptions for the prior distributions include the Gaussian (*e.g.,* when an estimate is made from repeated observations of a randomly varying process, or when the uncertainty is given as a standard deviation or a confidence interval), the uniform (*e.g.,* when a manufacturers specification, or some other certificate, give limits without specifying a confidence level and without any further knowledge of the distributions shape), and the triangular distribution (*e.g.,* when the measured values are more likely to be close to a value α than near the bounds of an interval with mean equal to α) [79]. For instance, if a manufacturers specification, or some other certificate, give limits in the form of a maximum range, $y \pm \alpha$, without any further knowledge of the distributions shape, then the estimated standard uncertainty is equal to $u(\mathbf{y}) = \frac{\alpha}{\sqrt{3}}$, while if the maximum range is described by a symmetric distribution then $u(\mathbf{y}) = \frac{\alpha}{\sqrt{6}}$.

Having expressed the individual uncertainties as standard uncertainties, the next step (ref. Figure 9-Step 4) is to combine them in the form of a *combined standard uncertainty*. Although in practice there may exist correlations between the individual uncertainty sources, however, it is usually impossible to compute those correlations accurately. For this purpose, it is more convenient to rely on an assumption of independence between the individual uncertainty sources. In the following, let $y = f(x_1, \ldots, x_L)$ be an observed variable, which depends on L input variables x_l through a functional relation $f(\cdot)$. Then, the combined standard uncertainty of y, for independent input variables x_l, $l = 1, \ldots, L$, is given by

$$u_c(y) = \sqrt{\sum_{l=1}^{L} \left(\frac{\partial f}{\partial x_l}\right)^2 u^2(x_l)} \ , \tag{26}$$

where $u(x_l)$ denotes the standard uncertainty of the input variable x_l (either of type A, or of type B), while the partial derivatives $\frac{\partial f}{\partial x_l}$, the so-called *sensitivity coefficients*, quantify how much the output y varies with

TABLE I
COVERAGE FACTOR AS A FUNCTION OF CONFIDENCE LEVEL FOR THE GAUSSIAN DISTRIBUTION

Coverage factor (k)	**Confidence level** ($\%$)
$k = 1$	67%
$k = 1.96$	95%
$k = 2.576$	99%
$k = 2.3$	99.7%

changes in the values of the input variables x_l. It is also important to note that, before the evaluation of $u_c(y)$, we have to ensure that all the distinct standard uncertainties are expressed in the same units.

However, in practice, even for the modern sensing devices, it is usually extremely difficult to calculate the sensitivity coefficients accurately. To this end, the easiest way is either to consider a weighted scheme, that is, $\frac{\partial f}{\partial x_l} = w_l$ with $\sum_{l=1}^{L} w_l = c$, where c is a predefined constant, while the degree of contribution (or, equivalently, the weight value) of the individual input variables (uncertainty sources) is set in a rather empirical fashion. More details about the complex case of correlated input variables can be found in [78], we emphasize though that this assumption is usually avoided in the industrial practice due to the difficulties in computing accurately the interdependencies among the identified uncertainty sources. On the other hand, if such an assumption is not valid explicitly, the correlations themselves can be avoided if the common influences are introduced as additional independent input variables.

Finally, the combined standard uncertainty, which may be thought of as equivalent to one standard deviation, is transformed into an overall *expanded uncertainty*, U, which is the final output, via multiplication with a coverage factor k, that is,

$$U(y) = k \cdot u_c(y) \ , \tag{27}$$

where the value of k is determined in terms of the desired confidence level of a Gaussian distribution as shown in Table I.

The computation of U completes the first building block of the U-HDMA system shown in Figure 8. In the following, we focus on the second building block, namely, the detection of extreme events by employing appropriate extreme event detectors in order to account for the inherent data uncertainty.

B. Uncertainty-aware alerting notifications

When working in an industrial environment, various distinct rarely occurring "events" can be devastating to the proper operation of the whole infrastructure. For instance, in manufacturing industries we are interested in preventing potential defections of the production engines by monitoring critical structural parameters, while in a large-scale water treatment industrial plant a key task is to support early detection of high concentrations for several chemicals which may be harmful for the public health. Thus improving our understanding of such *extreme events* will further help to mitigate their effects.

Extreme events can occur at any phase and time instant of the infrastructure's life-cycle, which necessitates its continuous and efficient monitoring to achieve early detection of abnormal behavior. Although a typical industrial setting is generally intended to operate autonomously, however, in extreme events it is of high significance to anticipate the impact of the detected events by triggering appropriate actuators in time. To this end, designing fast and accurate extreme events detectors for providing *early warning notifications* is a strong demand in order to guarantee the smooth operation of critical industrial infrastructures.

Among the several approaches, which have been introduced in the literature,*extreme value theory* (EVT) provides efficient algorithmic tools to assess, from a given ordered sample of a given random variable, the probability of events that are more extreme than any previously observed. Two approaches are the most widely used in practice for extreme value analysis, namely, the method of *block maxima* (BMax) [80] and the method of *peaks-over-threshold* (POT) [81], [82]. Depending on the application, each method has its own advantages and limitations. For instance, for the method of block maxima, theoretical assumptions are less critical in practice, while it is also easier to apply. However, estimation errors can be large for relatively small block sizes. On the other hand, the method of peaks-over-threshold yields more independent exceedances than block

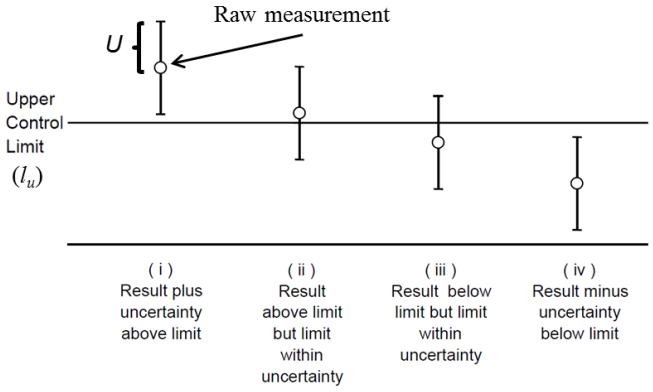


Fig. 11. Compliance of uncertainty-augmented measurements with a predetermined upper operating limit.

maxima, along with tighter confidence intervals. Its main drawback is that an independence assumption is critical, which may not hold in practice, and also the choice of an appropriate threshold is somewhat ambiguous in practice resulting in a less easier implementation.

Furthermore, a common characteristic is that, in both cases, the detection of extreme events is based on the raw data without accounting for their underlying uncertainty. In addition, given our major requirement for providing timely notifications of abnormal behavior, the selected extreme events detection method must have a small computational complexity, without sacrificing the detection accuracy. The simplest approach to satisfy both requirements, that is, to exploit the inherent data uncertainty while being computationally efficient, is obtained by modifying an alternative widely used method, the so-called *compliance with operating limits* (COL).

Without loss of generality, in the following, we restrict ourselves in the case of an upper operating limit, however, the same remarks are straightforward when compliance with a lower operating limit is required. More specifically, let l_u denote an upper operating limit dictated by a manufacturer or a specification standard. In addition, let $\tilde{y} = y \pm U$ be a measurement augmented by its associated expanded uncertainty interval. In contrast to the typical COL method, for which only two cases exist when checking for compliance between the raw measurement y and the upper limit l_u, as shown in Figure 11 there two additional cases for its uncertainty-aware counterpart, hereafter denoted as U-COL. Specifically, the four possible cases of U-COL are as follows: (i) both the measurement and the expanded uncertainty interval are above the upper limit l_u; (ii) the measurement is larger than l_u and the expanded uncertainty interval contains l_u; (iii) the measurement is lower than l_u and the expanded uncertainty interval contains l_u; and (iv) both the measurement and the expanded uncertainty interval are below l_u.

Among them, only case (i) triggers clearly an alerting notification for the occurrence of an extreme event, while (iv) is the only one which is in compliance with the specifications. On the other hand, in cases (ii) and (iii) we could not infer with absolute certainty whether an alert should appear or not. Nevertheless, in applications with profound social impacts, such as, for instance, water treatment, a system operator should classify cases (ii) and (iii) as possible divergences from normal operation, and thus draw more attention on the associated monitored variables.

The importance of accounting for the inherent data uncertainty in order to increase the efficiency of an extreme events detector is demonstrated in Figure 12. More specifically, this figure shows the extreme events detection performance of (a) the original COL (upper plot) and the uncertainty-aware U-COL (bottom plot) method for a temperature sensor by setting $l_u = 17\,^o$C. The red dots correspond to detected extreme events, while the orange dots correspond to potential extreme events. Although both methods achieve to identify the extreme peaks in the recorded temperature measurements, however, their key difference is that COL notifies for an extreme event only when we reach the peak of the curve. On the contrary, U-COL starts notifying for a potential deviation from "normal" behavior when the curve of the uncertainty-augmented measurements exceeds the predefined threshold. Indeed, this can be seen clearly in the two zoomed regions shown in

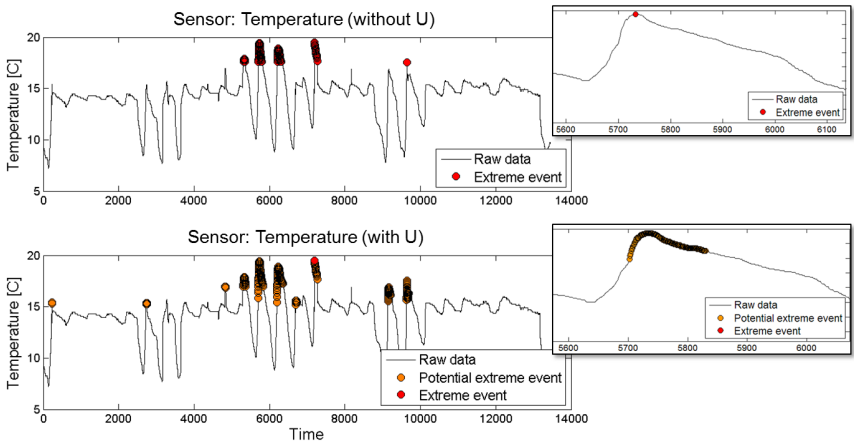


Fig. 12. Extreme events detection without and with uncertainty for a temperature sensor ($l_u = 17\,^o$C).

Figure 12. This observation reveals the increased tolerance of U-COL in detecting extreme events, when compared to its simple COL counterpart. The benefit of a system operator from using U-COL is that the U-HDMA system will start sending notifications prior to the occurrence of an event.

C. Fast and efficient monitoring of pairwise data stream correlations

Efficient discrimination between occasional and extreme events is a major issue in the design of robust data management systems. It is of great importance to ensure that a true extreme event occurs and not some coincidence, or system or network failure. On the other hand, the degree of correlation between two or more sensor data streams characterizes their interrelations and dependencies. To this end, timely identification of highly correlated streams can be further exploited as a guarantee to verify the existence of a detected extreme event. Though, the degree of "high correlation" is related to the specific application and the end-user, who has the flexibility to define how much strict this degree will be.

Extraction of pairwise correlations yields a partition of the set of available sensors into subsets of highly correlated sensors. This clustering facilitates the monitoring of the overall infrastructure by a system operator, who focuses only on a subset of sensors, where an abnormal behavior has been detected for at least one of its members. In the following, let $\mathbf{x} \in \mathbb{R}^N$, $\mathbf{y} \in \mathbb{R}^N$ be two sensor streams of length N, and $\mathbf{x}_w = (x_{t_1}, \ldots, x_{t_w}$, $\mathbf{y}_w = (y_{t_1}, \ldots, y_{t_w}$ be two time-synchronized windows of size w. The typical approach for extracting pairwise sensor stream correlations is by means of the Pearson's correlation coefficient, which is given by

$$\mathrm{corr}(\mathbf{x}_w, \mathbf{y}_w) = \frac{\sum_{i=1}^{w} x_{t_i} y_{t_i} - w \bar{x}_w \bar{y}_w}{(w-1)\sigma_{x_w}\sigma_{y_w}} \ , \tag{28}$$

where $\bar{x}_w$, $\bar{y}_w$ are the means of $\mathbf{x}_w$ and $\mathbf{y}_w$, respectively, and σ_{x_w}, σ_{y_w} denote their corresponding standard deviations.

From a computational perspective, the main limitation is that the correlation coefficient has to be recalculated for each newly acquired measurement, which increases the computational burden, especially for high-dimensional data streams or for a large number of sensors. To this end, a computationally efficient solution was proposed based on the use of the discrete Fourier transform (DFT). Working in a DFT framework, each sample x_{t_i} (similarly y_{t_i}) can be expressed as a linear combination of exponential functions

$$x_{t_k} \approx \frac{1}{\sqrt{w}} \sum_{f=0}^{K-1} X_f e^{\frac{i2\pi f k}{w}} \ , \quad k = 1, \ldots, w \ , \tag{29}$$

where X_f $(f = 0, \ldots, K-1)$ is the set of K DFT coefficients, with $K < w$. Doing so, the computation of the correlation coefficient in (28) is performed in terms of DFT coefficients.

Most importantly, the fast and efficient computation of DFTs enables the fast monitoring of synchronized sensor streams, whose correlation exceeds a predefined threshold. This is dictated by the following lemma, which gives a correspondence between the correlation coefficient and the Euclidean distance between DFTs.

Lemma 1. *([83])* Let $\hat{\mathbf{x}}_w$, $\hat{\mathbf{y}}_w$ be the normalization to mean zero and variance one of $\mathbf{x}_w$ and $\mathbf{y}_w$, respectively. In addition, let $\hat{\mathbf{X}}_w = \mathcal{F}\{\hat{\mathbf{x}}_w\}$, $\hat{\mathbf{Y}}_w = \mathcal{F}\{\hat{\mathbf{y}}_w\}$ be their corresponding DFTs. Then,

$$\mathrm{corr}(\mathbf{x}_w, \mathbf{y}_w) \geq \epsilon \Rightarrow d_M\left(\hat{\mathbf{X}}_w, \hat{\mathbf{Y}}_w\right) \leq \sqrt{2w(1-\epsilon)}\ . \tag{30}$$

In (30), ϵ is a predefined threshold and $d_M\left(\hat{\mathbf{X}}_w, \hat{\mathbf{Y}}_w\right)$ is the Euclidean distance between the corresponding *truncated* DFTs, which are obtained by keeping the first $M \leq \frac{w}{2}$ DFT coefficients with the largest amplitudes. The validity of this approach is based on the *compactness* of DFT representations, that is, the concentration of the main portion of the energy for a given sensor stream in the first few high-amplitude DFT coefficients.

The above lemma implies that by focusing on those sensor pairs, whose associated truncated DFTs are "close" enough, we get a set of *likely correlated* sensor pairs. Notice that this constitutes a superset of the correlated sensors, without false negatives. Furthermore, in our U-HDMA system we are interested in identifying and tracking highly correlated sensor pairs in an online fashion by also incorporating the estimated data uncertainty. Aiming at improving the computational performance of the DFT-based approach, while maintaining its accuracy, in our U-HDMA system the problem of extracting highly correlated pairs of sensors is translated into a problem of identifying highly *similar* sensors, where the similarity is measured by an appropriately designed function.

Let $\mathbf{x}$ be the reference sensor stream and $(\mathbf{y}_1, \ldots, \mathbf{y}_C)$ the set of candidate streams. At the core of our fast and robust "correlation extractor" is an efficient *peak similarity* function. Given two windowed, yet time-synchronized, data streams $\mathbf{x}_w$, $\mathbf{y}_{i,w}$ $(i = 1, \ldots, C)$ the corresponding expanded uncertainties U_{x_w}, $U_{y_{i,w}}$ are estimated first. Then, the uncertainty-augmented windows are formed: $\mathbf{x}_w^U = \mathbf{x}_w + U_{x_w}$ (or $\mathbf{x}_w^U = \mathbf{x}_w - U_{x_w}$), $\mathbf{y}_{i,w}^U = \mathbf{y}_{i,w} + U_{y_{i,w}}$ (or $\mathbf{y}_{i,w}^U = \mathbf{y}_{i,w} - U_{y_{i,w}}$). After their normalization to mean zero and variance one, $\hat{\mathbf{x}}_w^U$ and $\hat{\mathbf{y}}_{i,w}^U$, respectively, the M-sized ($M \ll \frac{w}{2}$) truncated DFTs are computed, $\hat{\mathbf{X}}_w^U = \mathcal{F}\{\hat{\mathbf{x}}_w^U\}$, $\hat{\mathbf{Y}}_{i,w}^U = \mathcal{F}\{\hat{\mathbf{y}}_{i,w}^U\}$. Finally, our *uncertainty-aware peak similarity* function is defined as

$$p_{sim,U}(\mathbf{x}_w, \mathbf{y}_{i,w}) = \frac{1}{M}\sum_{j=1}^{M}\left[1 - \frac{\left|\hat{\mathbf{X}}_{w;j}^U - \hat{\mathbf{Y}}_{i,w;j}^U\right|}{2 \cdot \max\left(|\hat{\mathbf{X}}_{w;j}^U|, |\hat{\mathbf{Y}}_{i,w;j}^U|\right)}\right]\ , \tag{31}$$

where $\hat{\mathbf{X}}_{w;j}^U$ denotes the j-th element of $\hat{\mathbf{X}}_w^U$ (similarly for $\hat{\mathbf{Y}}_{i,w}^U$).

Our U-HDMA system reports as "highly-similar" those sensor pairs for which $p_{sim,U}(\mathbf{x}_w, \mathbf{y}_{c,w}) > \epsilon_U$. In order to account for the potential loss of information caused by the truncation of the set of DFT coefficients, as well as for the incorporation of the underlying uncertainties, special attention should be given on the selection of the threshold ϵ_U. An extensive experimental evaluation on real measurements acquired by various distinct sensors, we observed that by selecting an "elastic" enough threshold ϵ_U, the subset of sensor streams $\mathbf{y}_c$, $c \in \{1, \ldots, C\}$, with the highest peak similarity values with $\mathbf{x}$ will also contain the highly correlated streams with $\mathbf{x}$ (that is, those with correlation coefficient above ϵ as stated in Lemma 1). Our evaluation showed that $p_{sim,U}$ achieves at least the performance of corr by setting $\epsilon_U = \epsilon + \epsilon_{\text{offset}}$, where ϵ_{offset} is a small positive number. Although our experimentation showed that it suffices to set $\epsilon_{\text{offset}} < 0.05$, however, an automatic and adaptive rule to select an optimal threshold ϵ_U needs a more thorough investigation.

To illustrate the computational efficiency of $p_{sim,U}$, its performance is compared against the typical correlation coefficient and two state-of-the-art methods, namely, BRAID [84] and StatStream [76]. BRAID can handle data streams of semi-finite length, incrementally, quickly, and can estimate lag correlations with little error. On the other hand, StatStream resembles more the design principles of $p_{sim,U}$ by finding high correlations among sensor pairs based on DFTs and a three-level time interval hierarchy. Figure 13 compares the execution times of $p_{sim,U}$ with the above three alternatives, as a function of the window size. The results reveal a significant improvement in execution time achieved by $p_{sim,U}$, which is more prominent for higher window lengths. Most importantly, we observe that the execution time of $p_{sim,U}$ remains almost constant over the whole range of

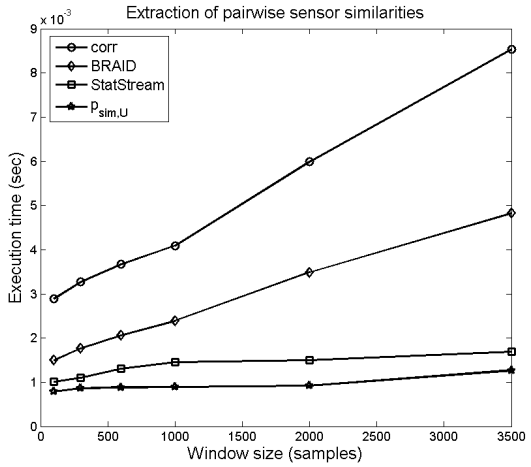


Fig. 13. Comparison of execution times, as a function of the window size, between a) uncertainty-aware peak similarity ($p_{sim,U}$), b) StatStream, c) BRAID, and d) correlation coefficient (corr).

selected lengths, in contrast to the naive (corr) and BRAID methods, whose execution times increase rapidly for increasing window length.

The BRAID algorithm, for which we set the correlation lag to zero, is characterized by gradual increase for increasing window size, since it employs all the values in the observed time interval. On the other hand, StatStream is based on a simple hash function of the mean of each sensor window. Keeping the integer part of the means, the data windows are mapped to appropriate cells in a grid structure. Doing so, only the correlations between neighboring cells are computed. The increased execution time of StatStream, compared to $p_{sim,U}$, is due to the hash function, which involves more computations for the mapping. It is expected though that the performance of StatStream could be enhanced by designing a more efficient hash function.

VI. A Use Case Scenario: The HYDROBIONETS Platform for Sea Water Desalination

The proposed framework, comprised of the techniques described in the Sections above, has been applied in an iCPS designed for the microbiological monitoring of water quality in industrial plants. The specific use case considers desalination plants and focuses on the procedure of reverse osmosis, which is a widely adopted technique across Europe and worldwide. Desalination by the means of reverse osmosis relies on *osmotic membranes*, that allow water to pass through at much higher rates than the dissolved salt.

The use of such membranes during sea water desalination suffers from the phenomenon of biofouling, which is related to the accumulation of unwanted bacterial matter on their surface. Biofouling is considered a very complex phenomenon that can be affected negatively by several variables, such as organic matter, pH, and temperature of feed water. As such, the combination of existing quantitative indices (e.g. temperature, conductivity, and pH) with novel sensing technology, capable of monitoring the presence and growth of bacteria in different locations of a desalination plant is considered critical for the early detection of biofouling.

This has been the primary motivation of the HYDROBIONETS project [1] for the design and development of an iCPS platform responsible for the autonomous monitoring of the biofouling phenomenon in industrial desalination plants. The resulting platform, henceforth called the *HYDROBIONETS platform*, combines a multi-tier network architecture and novel wireless biofouling sensors [85], with the existing, wired sensing infrastructure for optimizing the cleaning and maintenance of the osmotic membranes, thereby increasing their lifetime.

The main components of the HYDROBIONETS platform, illustrated in Figure 14 are the following [86]:

- The Wireless BioMEM Network (WBN), comprised of miniaturized, computationally-limited, and energy-autonomous sensor platforms that are responsible for monitoring the growth of biofouling bacteria at designated locations in the desalination plant. Each WBN node implements a sophisticated protocol stack that builds upon the scheduling mechanism described at Section IV-B, in order to effectively address the network and communication challenges often met in industrial environments;

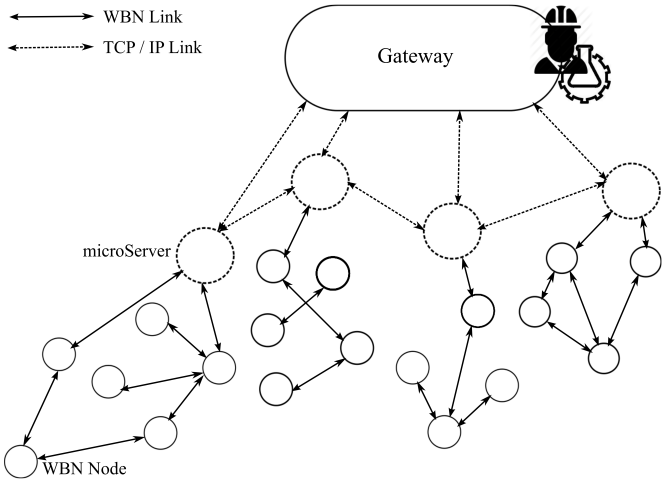


Fig. 14. The architecture of the HYDROBIONETS platform.

- The μServer devices, portable platforms with increased computational capabilities, that are responsible for the management of a cluster of WBN nodes in the industrial plant. Each μServer undertakes the network configuration, management and adaptation mechanisms for its appointed cluster of WBN nodes, whilst implementing functionalities that are cannot be computationally supported by the WBN nodes.
- The Gateway is considered the end-point between the WBN network and the existing infrastructure. It is the component of the platform from where the system administrator can interact with the biofouling sensor nodes, whilst allowing interconnectivity with the existing infrastructure of the industrial plant.

From the perspective of signal and data processing, the WBN nodes combine the characteristics of Tier 1 and Tier 2 of the proposed data-driven framework for iCPS architectures. As such, the WBN nodes located at the front-end sensing, and in coordination with their assigned μServer, are also capable of communicating with their peer components. Subsequently, the WBN nodes can directly handle the front-end signal modelling of the biofouling sensor data (ref. Section III). The extraction in-network correlation (ref. Section IV) for the growth of bacteria at different locations can exploit both the direct links between WBN nodes, as well as the exchange of information between different μServers. Finally, analysis based on the U-HDMA frame (ref. Section V) for both biofouling data, as well as relevant sensing indices, such as pH, temperature, and conductivity, is undertaken by the Gateway, which collects all information conveyed over the HYDROBIONETS platform.

A. Experimental Studies

The HYDROBIONETS platform has been deployed at a desalination pilot plant, located at La Tordera, Spain, and owned by Acciona Agua, which is a worldwide industrial leader in water treatment. Snapshots of the industrial plant are presented in Figure 15.

Driven by the small dimensions of the industrial plant, the HYDROBIONETS platform is comprised of ten WBN nodes, assigned to one μServer, and the Gateway. The WBN nodes are placed accordingly to the specifications of the end-user, at three different locations in the plant, namely: (a) the intake of the sea water, (b) the phase of pretreatment, and (c) the phase of reverse osmosis.

The WBN protocol stack has been implemented in Contiki OS [87], [88], featuring at the Routing Layer the IETF standard for low power and lossy networks. The resulting stack has been deployed on CM5000-SMA or XM1000 motes [89] (Figure 16(a)), integrated with the biofouling sensor over serial interface. Each biofouling sensor exploits an array of capacitive micro-electrodes, which merge into the treated water. The concentration bacteria in the treated water changes permittivity of the micro-electrodes, therefore modifying the magnitude of the capacitance. Subsequently, measuring the impedance of the sensor is considered sufficient for characterizing the phenomenon of biofouling [85]. A snapshot of the employed sensors mounted at the desalination plant is shown in Figure 16(b).

(a) (b) (c)

Fig. 15. La Tordera's water desalination pilot plant.

(a) (b)

Fig. 16. The technology employed for the realization of the WBN network at the pilot plant: (a) the CS5000-SMA mote [89] employed for the realization of the protocol stack on the WBN node, (b) the casing and the biofouling sensor cell of the WBN node.

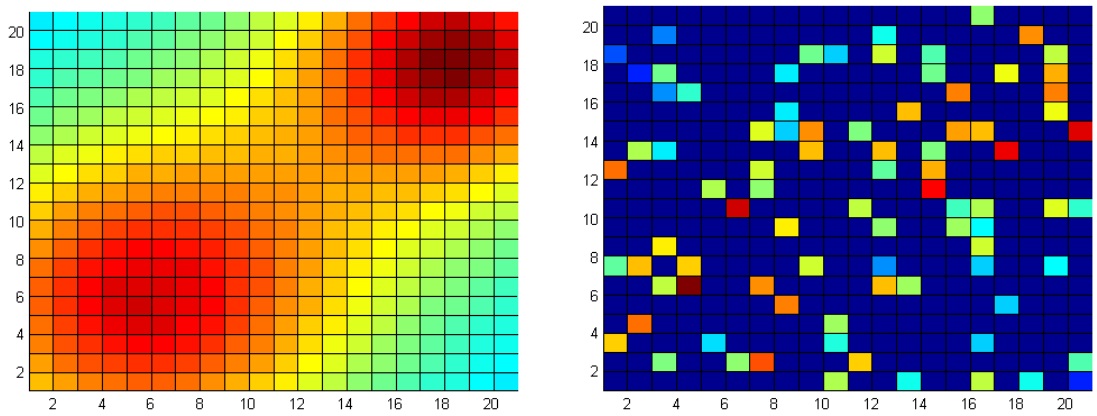
While the qualitative characterization of the biofouling phenomenon based on impedance values remains in progress, data collected both by the wireless biofouling sensors as well as existing wired sensing infrastructure are employed for evaluating the efficacy of our proposed data-driven framework. In the remaining of this Sections the respective experimental results are presented and accompanying discussions are made.

B. In-network distributed processing

An illustration of the distributed parameter estimation and field reconstruction based on iterative consensus, as described in Section IV-C, is presented in this section. In particular, the spatial field $s(x, y, t)$ to be estimated at time t and coordinates (x, y) is the superposed temperature originated by M distinct heat sources, each one emitting with a different time-varying power. The spatial diffusion of the heat coming from each individual source is modeled as a Cauchy bell. Doing so, the temperature at spatial location (x, y) and time t is given by

$$s(x, y, t) = \sum_{i=1}^{M} \frac{p_i(t)}{1 + ((x - x_i)^2 + (y - y_i)^2)/\beta_i} \ ,$$

where each source i is located at coordinates (x_i, y_i) and emits a time-varying heat power $p_i(t)$ with spread β_i. Note that this expression shows the value of the field $s(t)$ at each location (x, y) as a spatial distortion of the state vector $\mathbf{p}(t) = [p_1(t) \dots p_M(t)]$ at (x, y), following a linear observation model. Figure 17(a) shows a snapshot of this field at a single time instant over a 20 m $\times$ 20 m grid with with $M = 2$ heat sources located at coordinates $(20, 20)$ and $(60, 60)$ and emitting heat powers of $p_1 = 50$ and $p_2 = 20$, with spreads $\beta_1 = 20$ and $\beta_2 = 30$, respectively. In this example, no assumption is made regarding the evolution of the power $\mathbf{p}$ across time.



(a) Temperature field to be estimated at a single time instant

(b) Noisy observations of the nodes at a single time instant

Fig. 17. Temperature field at a given time instant formed by two heat sources located in a 20 m $\times$ 20 m square area, each emitting a heat power whose propagation is modeled as a Cauchy bell.

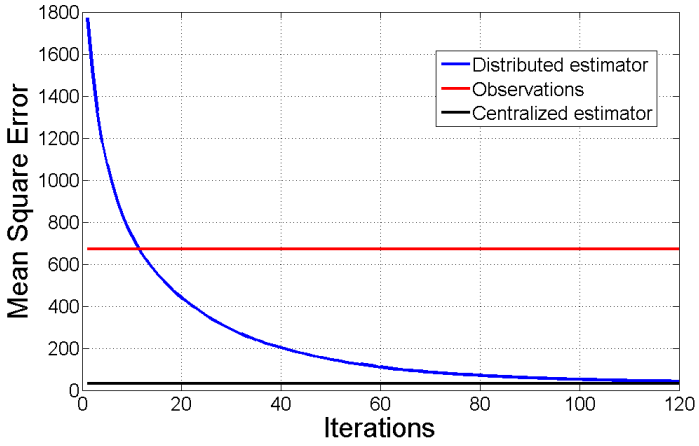
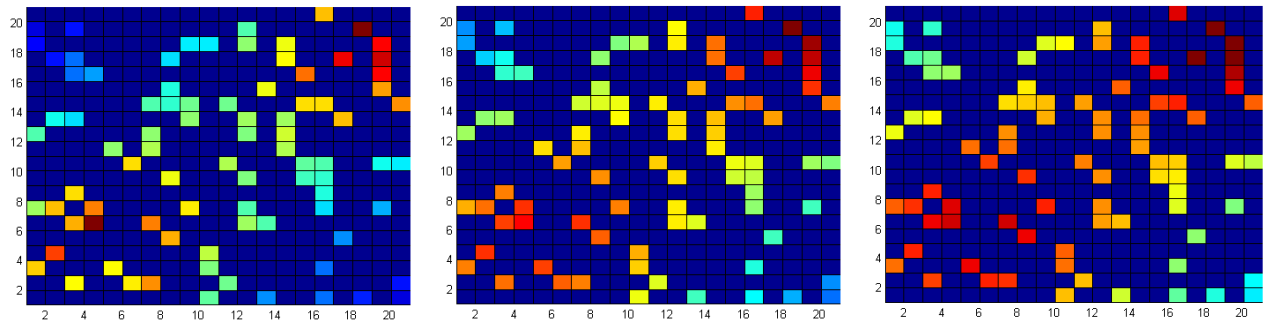


Fig. 18. Evolution of the mean squared error (MSE) of the distributed estimation of the field across the iterations of the consensus process.

To monitor the temperature field, a set of $S = 100$ nodes are deployed randomly throughout the area. This is also in agreement with the non-uniform deployment adopted by the HYDROBIONETS platform, whilst the importance for monitoring temperature fields stems from the fact that temperature affects the operating conditions of the sensor network, as well as the industrial process.

Following our observation model (10), data obtained by the nodes are corrupted by additive zero-mean Gaussian noise, that is, the measurement of node i at time t is given by $y_i(t) = s_i(t) + w_i(t)$, where $s_i(t) = s(x = x_i, y = y_i, t)$. A snapshot of the measurements of all nodes at the same time instant is shown in Figure 17(b) for noise variance $\sigma^2 = 30$. The communication range area of each node has a radius of 8 m. Since we have no knowledge about the evolution of the field, we cannot make any inference from the previous estimations. For this purpose, the optimal parameter estimator is computed at each time instant separately.

Based on the acquired noisy observations, the nodes start an iterative process to exchange information with the neighbors inside their range, as explained in Section IV-C. The final aim is to compute the maximum likelihood estimator of the field. The evolution of the mean squared error (MSE) of the distributed estimator (blue curve) along the iterations is shown in Figure 18. More specifically, the MSE between the original and estimated temperature values is computed for each sensor in each iteration, which is then averaged over



(a) Iteration 40 (b) Iteration 80 (c) Iteration 120

Fig. 19. Successive screenshots of the iterative temperature field reconstruction until a consensus is reached ($S = 100$, $\sigma^2 = 30$).

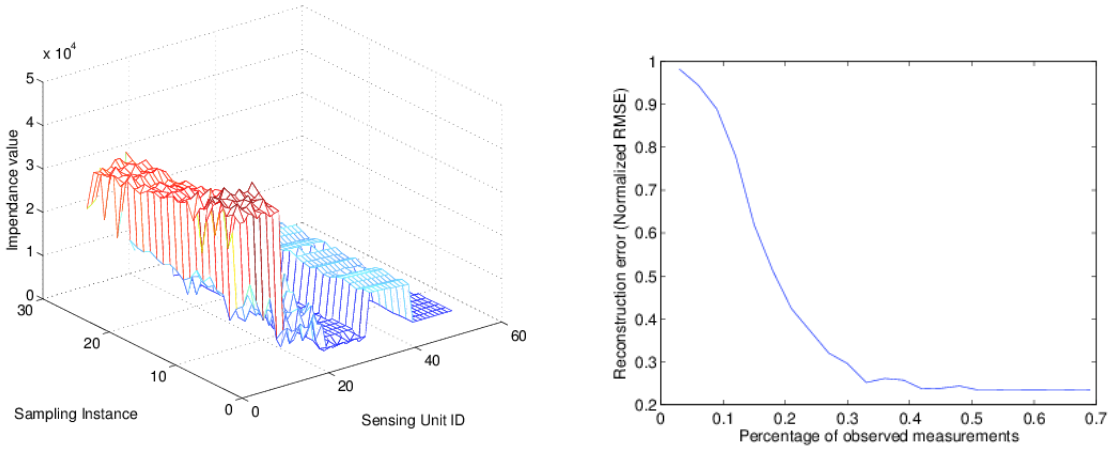
all the $S = 100$ sensors. In this plot, the red curve corresponds to the average MSE between the initial sensor measurements and the real values of the temperature field at the location of each sensor, whereas the black curve corresponds to the average MSE between the centralized (MLE) estimator and the real values of the temperature field. A comparison among the three curves reveals that as the nodes reach a consensus, the distributed estimation error decreases and approximates closely the error of the centralized estimator. Figure 19 visualizes the evolution of the distributed estimation process over the two-dimensional temperature field, and across the iterations, by employing the noisy observations and the distributed exchange of information among the nodes.

C. iCPS data recovery via MC

To validate the merits of introducing intelligence into the signal acquisition and processing, we present a few indicative results from the testbed developed within the HYDROBIONETS platform. We consider a collection of five sensor nodes, each one capable of measuring impedance in one out of ten frequency channels, leading to a total of 50 sensing units. Each measurement is communicated via the μServer and the Gateway and stored locally in a database. To generate the corresponding measurements matrix, one must select the temporal resolution of the sampling process, which dictates the number of rows of this matrix. For instance, assuming one measurement every hour, a measurements matrix corresponding to a single day will have 24 rows.

Figure 20 presents an indicative collection of impedance measurements acquired by a set of biofilm sensors over a period of three days with a temporal sampling rate equal to one measurements every three hours (180 min resolution), along with the recovery performance of the measurements matrix from a subset of its entries. In Figure 20(a), the existence of spatio-temporal correlation can be observed in the data, while Figure 20(b) demonstrates the relationship between the recovery error and the sampling rate. Specifically, we observe that 25% of the measurements are sufficient in order to reduce the recovery error to less than 25% of the original signal's magnitude. Furthermore, it can be seen that by increasing the sampling rate beyond 40% has no effect on the reconstruction quality. This phenomenon is attributed to the noise that corrupts the data acquisition process and increases artificially the rank of the matrix, setting a lower bound on the recovery error.

Whereas missing entries are typically attributed to lost packets and node failures, this situation can also arise by increasing the temporal resolution of the sampling process. This can be easily observed in Figures 21(a), 21(c), where the same number of stored values as in Figure 20 is employed, however, at different temporal resolution. More specifically, while in Figure 20 each entry corresponds to a period of three hours, Figures 21(a) and 21(c) present the measurements matrix generated by considering one entry every two hours and one hour, respectively. One can see that an increase of the temporal resolution will yield the introduction of zero-valued entries due to the lack of measurements for the predefined time intervals. Figures 21(b) and 21(d) present the estimated measurements matrices for the two cases, where zeros have been replaced



(a) Measurements (b) Recovery performance

Fig. 20. Impedance measurements of biofilm sensors: (a) complete set of measurements over a period of three days at a sampling rate of one measurement every three hours; (b) reconstruction error for a given sampling rate. MC is able to achieve good performance even at sampling rates as low as 30% of the total measurements.

with approximated values estimated via MC. While it is not possible to provide a quantitative performance evaluation, since these values where never recorded in practice, from a visualization perspective it is much easier for a system operator to monitor the iCPS infrastructure conditions and perform the necessary actions.

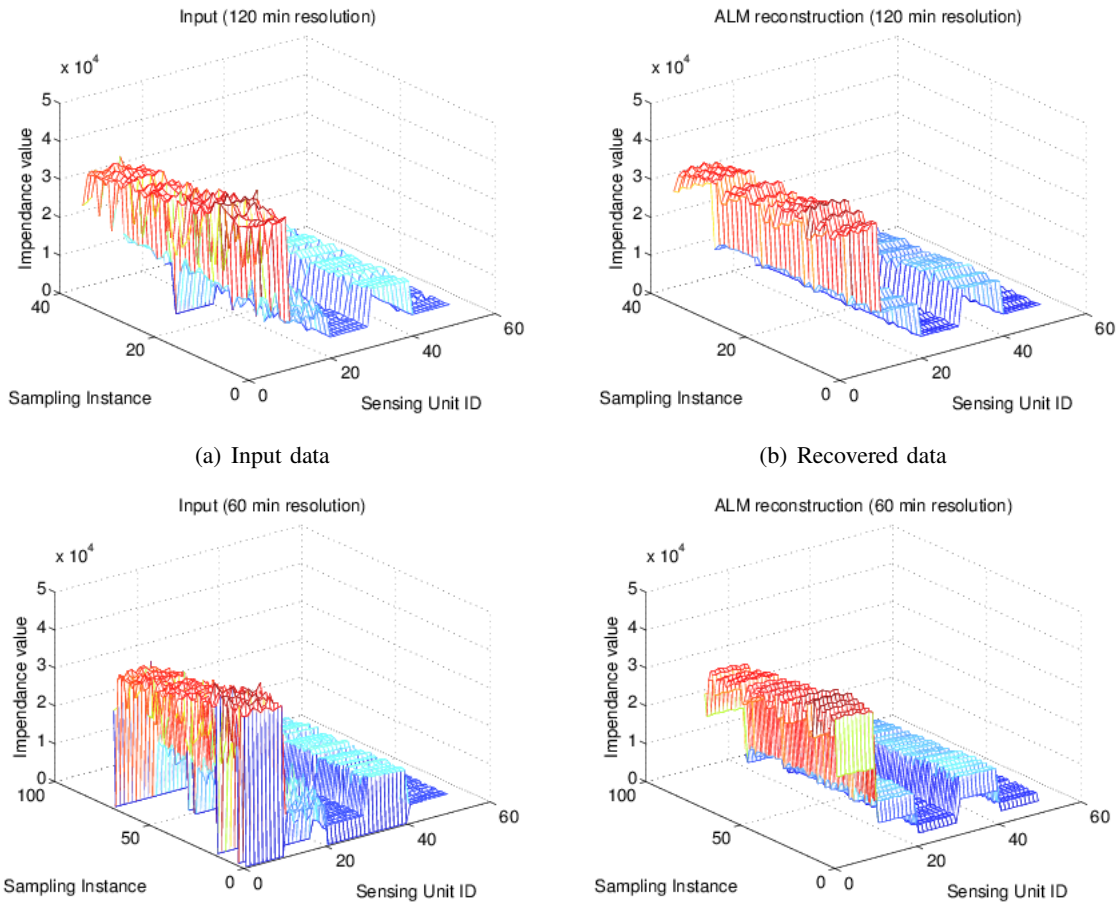
D. High-level sensor data analysis and alerting

In the following, the performance of the U-HDMA system introduced in Section V, is evaluated on a real dataset in terms of managing the underlying data uncertainty and providing early warnings. In particular, the high-level analysis and early warning is performed for a dataset provided by Acciona Agua, which consists of 22 sensors of several types (pressure, temperature, conductivity, turbidity, pH, flow, and redox) deployed in La Tordera's desalination plant. The corresponding measurements cover a period of 1 month ($2 - -29$ April 2013) at a sampling rate of one measurement every three minutes. Full sensor specifications, such as sensor precision, sensitivity, and resolution, along with the corresponding measurements are provided for each individual sensor.

The inherent uncertainty of the acquired sensor data is estimated over sliding windows. If not stated explicitly otherwise, in the subsequent experimental evaluation the window size is set equal to 80 samples, which corresponds to a time interval of approximately 4 hours, while the step size is fixed at 1 sample corresponding to a time-step of about 3 minutes. The expanded uncertainty is computed by fixing the coverage factor in (27) at $k = 1.96$, which is equivalent to a 95% confidence level.

The performance of the spreadsheet-based approach (ref. [90]) for estimating the underlying uncertainty in several distinct sensor streams is illustrated first. Figure 22 shows the estimated expanded uncertainty for four randomly chosen sensors in our dataset. An additional potential, which is revealed by this figure, is the use of the estimated uncertainty as an indicator of abnormal behavior. Indeed, the time instants where the uncertainty presents a peak coincides with the time windows where the corresponding sensor measurements vary significantly compared to the previously recorded values. However, a more thorough study is required towards the design of an efficient extreme event detector based on expanded uncertainties.

As a final illustration, we evaluate the performance of the uncertainty-aware extreme event detector, U-COL, introduced in Section V-B. To this end, Figure 23 shows the identified instants (red dashed lines) for which an alerting notification is sent by the U-HDMA system for the conductivity and temperature sensors shown in Figure 22. More specifically, the typical COL and the uncertainty-aware U-COL methods are employed for early warning about an abnormal behavior in the acquired sensor data using the following rule: *"an alerting notification is sent by the system if N_c consecutive measurements exceed a predefined operational*



(a) Input data

(b) Recovered data

(c) Input data

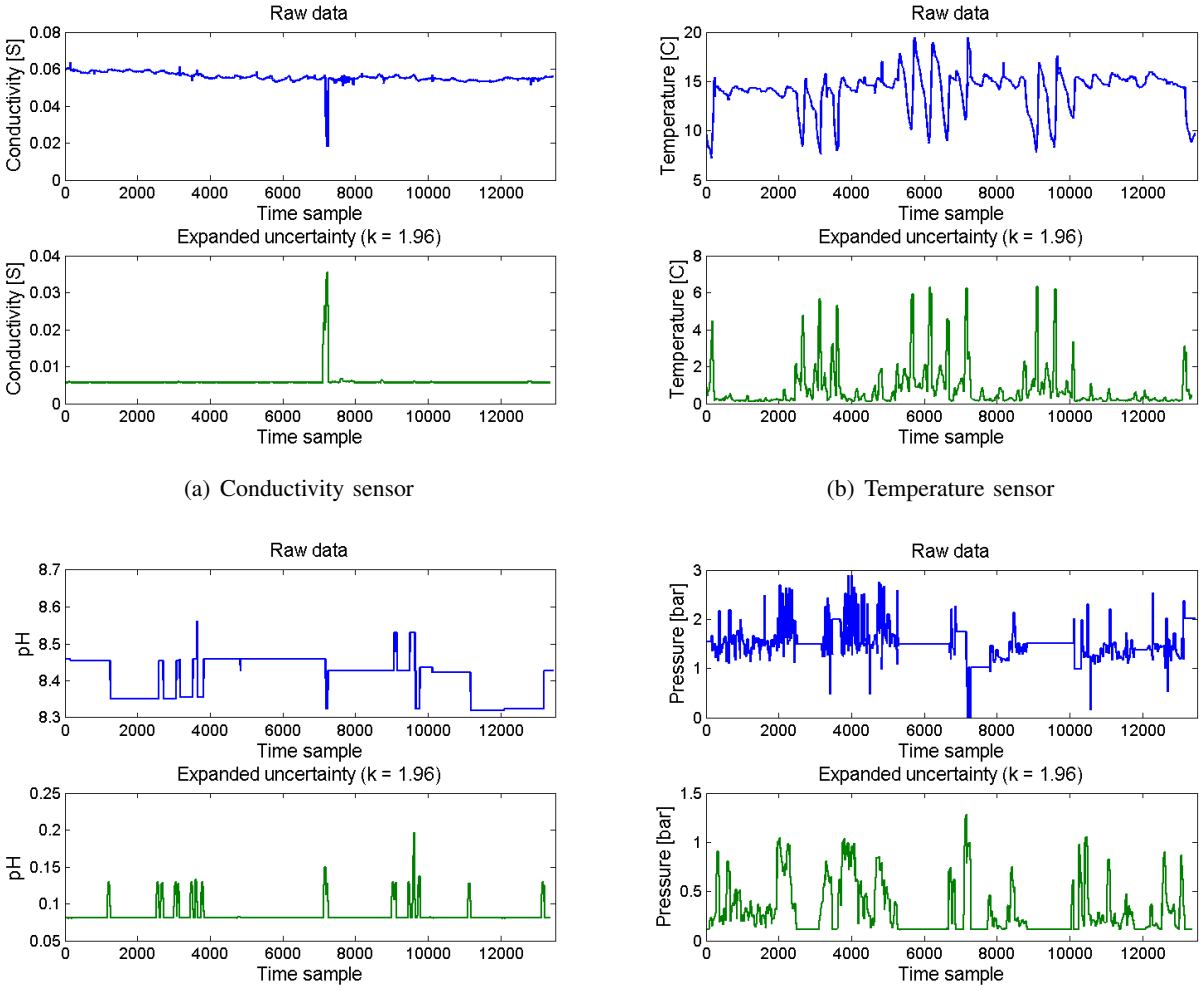
(d) Recovered data

Fig. 21. Illustration of recovery performance in the case of zero-valued measurements due to artificial increase of sampling rate. (a) and (c) correspond to the same measurement period as in Figure 20(a), albeit at higher sampling rates, leading to an increase in the number of missing entries; (b) and (d) present the performance of MC in removing artificially introduced zero-entries, and the output data which is both consistent with the measurements, as well as less abrupt due to the zeros.

upper limit". Notice that in the case of U-COL, the measurements are augmented with their corresponding estimated expanded uncertainty. In this example, we set $N_c = 20$, that is, the system operator is notified for a potential alert when the 20 most recent sensor measurements satisfy the above alerting rule. Besides, for demonstration purposes, for both sensors the upper limit is set to $\max\{\text{sensor data}\} - 0.01 \cdot \max\{\text{sensor data}\}$, that is, 10% below the maximum recorded value. Clearly, accounting for the inherent data uncertainty improves the early warning performance, as it can be seen in Figure 23 for both sensors. Indeed, in both cases, U-COL is able to detect the occurrence of abnormal behavior in the sensor data, even if the recorded raw measurements do not strictly exceed the corresponding operational upper limit.

VII. Conclusions and Future Research

In this chapter, the main architectural characteristics towards designing efficient data-driven industrial cyber-physical systems were analyzed. Furthermore, an integrated framework of signal and data processing techniques was presented, for treating different layers of information abstraction. By also accounting for the potential limitations and imperfections of the associated sensor network infrastructure employed to observe various physical parameters of the industrial environment, we focused on three major aspects, namely, i) signal processing-driven performance optimization for industrial sensor networks, ii) in-network signal processing for distributed estimation and tracking of spatio-temporal fields, and iii) high-level analysis and early warning by employing the recorded iCPS data.



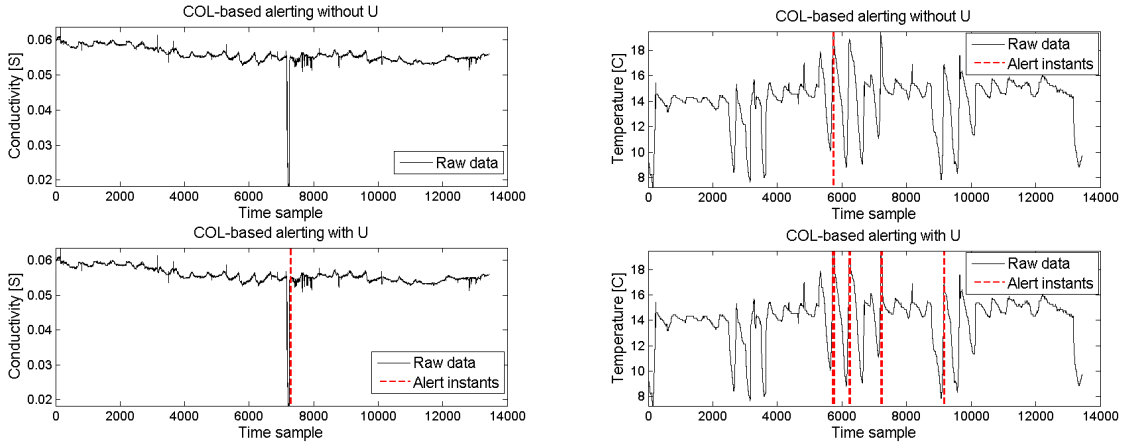
(a) Conductivity sensor

(b) Temperature sensor

(c) pH sensor

(d) Pressure sensor

Fig. 22. Raw data and estimated expanded uncertainties for four distinct electrochemical sensors: (a) conductivity; (b) temperature; (c) pH; (d) pressure (window size $= 80$, step size $= 1$, $k = 1.96$).



(a) Conductivity sensor

(b) Temperature sensor

Fig. 23. Raw data and identified alerting instants using the typical COL and the uncertainty-aware U-COL methods for: (a) conductivity sensor; (b) temperature sensor (window size $= 80$, step size $= 1$, $k = 1.96$, $N_c = 20$).

Along with a strong emphasis on providing the essential theoretical background, the effectiveness of the resulting framework was evaluated on a real iCPS. In particular, the signal and data processing techniques described herein were applied on real sensor data recorded by several distinct sensors deployed for monitoring a water desalination plants. Comparison with well-established and state-of-the-art methods for sensor data processing and distributed in-network inference revealed a superior performance of the integrated multi-tier iCPS architecture as described in this chapter.

Further extensions and enhancements, spanning the whole extent of the data processing chain in the iCPS setting presented herein, could yield additional improvements in the overall performance. Concerning the representation of industrial data based on the concepts of compressed sensing and matrix completion, as presented in Section III, innovative signal processing and learning algorithms can provide elegant solutions to issues that hinder the efficacy of iCPS. However, the majority of work in this area still relies on deliberately introducing these algorithms in various stages of data acquisition, processing and understanding. We expect that significantly more profound benefits can arise from the intelligent design and integration of these algorithms into the hardware platforms. Such designs will consider the end-to-end architectures, optimizing the overall performance of iCPS, instead of individual stages. Furthermore, by introducing domain-expert knowledge into the recovery process, we expect a surge in the number of iCPS applications due to the clear and measurable benefits that are associated with these methods.

Concerning the parameter and state estimation problems described in Section IV, linear observation and process models are assumed, along with zero-mean Gaussian noise. However, in practice, such assumptions are satisfied rarely. To this end, we have to focus on the design of distributed implementations for optimal parameter and state estimation in the general case of non-linear observation models and non-Gaussian noise. Furthermore, a thorough investigation of the performance of random and asymmetric network topologies has also to be carried out.

Finally, the efficiency of high-level data analysis and early warning methods presented in Section V can be further improved in several directions. First, statistical dependencies among distinct sources of uncertainty usually exist in practice. Although it is often a very difficult task to quantify such dependencies in practice, however, an increased accuracy in estimating the underlying sensor data uncertainty is expected by employing sensitivity coefficients (ref. (26)) which better approximate the input-output interrelation for a given sensor. A second extension is related to the performance of the similarity function employed for measuring pairwise sensor correlations. Specifically, an incremental implementation of the DFT-based peak similarity function, given by (31), as new sensor measurements are acquired, could further reduce its computational complexity, and subsequently the execution time, without sacrificing its effectiveness in identifying high correlated pairs of sensors. Moreover, the design of more sophisticated uncertainty-aware extreme event detectors, capable of simultaneously exploiting information even from heterogeneous, distinct sensors, could achieve a superior performance in terms of accurate detection of extreme events in different, yet correlated, sensor streams.

It is highly anticipated that the presented methods and the accompanying illustrations of real-life results, will act as a fertile ground for further enhancements and adaptations to distinct use cases, as well as for yielding novel directions for iCPS standardization.

ACKNOWLEDGMENTS

This work is supported by the HYDROBIONETS project (ICT-2011-7) funded by the European Commission in FP7 (GA-2011-287613) and the PEFYKA project within the KRIPIS ction of the GSRT, Greece. We are also grateful to Acciona Agua [2] for providing the premises of La Tordera's desalination plant, as well as to Ateknea Solutions [3] and CNM (Centre Nacional de Microelectrònica) [4] for assisting the collection of real biofouling data.

[2]http://www.acciona-agua.com/

[3]http://ateknea.com/

[4]http://www.imb-cnm.csic.es/

REFERENCES

[1] "The hydrobionets project: Autonomous control of large-scale water treatment plants based on self-organized wireless biomem sensor and actuator networks," http://www.hydrobionets.eu/.

[2] F.-J. Wu, Y.-F. Kao, and Y.-C. Tseng, "From wireless sensor networks towards cyber physical systems," *Pervasive and Mobile Computing*, vol. 7, no. 4, pp. 397 – 413, 2011.

[3] T. S. Rappaport *et al.*, *Wireless communications: principles and practice*. Prentice Hall PTR New Jersey, 1996, vol. 2.

[4] K. Remley, G. Koepke, C. Holloway, D. Camell, and C. Grosvenor, "Measurements in harsh rf propagation environments to support performance evaluation of wireless sensor networks," *Sensor Review*, vol. 29, no. 3, pp. 211–222, 2009.

[5] J. Ferrer Coll, "Rf channel characterization in industrial, hospital and home environments," pp. xiii, 65, 2012, qC 20120119.

[6] I.-U.-H. Minhas, "Wireless sensor network performance in high voltage and harsh industrial environments," Master's thesis, Blekinge Institute of Technology, School of Engineering, 2012.

[7] S. Ghadimi, J. Hussian, T. S. Sidhu, and S. Primak, "Effect of impulse noise on wireless relay channel," *Wireless Sensor Network*, vol. 4, no. 6, pp. 167–172, 2012.

[8] J. Hespanha, P. Naghshtabrizi, and Y. Xu, "A survey of recent results in networked control systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 138–162, Jan 2007.

[9] K. J. Astrom and P. Kumar, "Control: A perspective," *Automatica*, vol. 50, no. 1, pp. 3 – 43, 2014.

[10] W. Zhang, M. S. Branicky, and S. M. Phillips, "Stability of networked control systems," *IEEE Control Systems Magazine*, vol. 21, no. 1, pp. 84–99, Feb. 2001.

[11] *Wireless Hart*, The HART Communication Foundation Std.

[12] *ISA100, Wireless Systems for Automation*, International Society for Automation Std.

[13] "Ieee standard for local and metropolitan area networks–part 15.4: Low-rate wireless personal area networks (lr-wpans)," *IEEE Std 802.15.4-2011 (Revision of IEEE Std 802.15.4-2006)*, pp. 1–314, Sept 2011.

[14] X. Cao, P. Cheng, J. Chen, and Y. Sun, "An online optimization approach for control and communication codesign in networked cyber-physical systems," *Industrial Informatics, IEEE Transactions on*, vol. 9, no. 1, pp. 439–450, Feb 2013.

[15] R. Alur, A. D'Innocenzo, K. Johansson, G. Pappas, and G. Weiss, "Compositional modeling and analysis of multi-hop control networks," *Automatic Control, IEEE Transactions on*, vol. 56, no. 10, pp. 2345–2357, Oct 2011.

[16] V. Gupta, B. Hassibi, and R. M. Murray, "Optimal {LQG} control across packet-dropping links," *Systems & Control Letters*, vol. 56, no. 6, pp. 439 – 446, 2007.

[17] S. Deshmukh, B. Natarajan, and A. Pahwa, "State estimation in spatially distributed cyber-physical systems: Bounds on critical measurement drop rates," in *Distributed Computing in Sensor Systems (DCOSS), 2013 IEEE International Conference on*, May 2013, pp. 157–164.

[18] V. Gupta, A. Dana, J. Hespanha, R. Murray, and B. Hassibi, "Data transmission over networks for estimation and control," *Automatic Control, IEEE Transactions on*, vol. 54, no. 8, pp. 1807–1819, Aug 2009.

[19] M. Pajic, S. Sundaram, G. Pappas, and R. Mangharam, "The wireless control network: A new approach for control over networks," *Automatic Control, IEEE Transactions on*, vol. 56, no. 10, pp. 2305–2318, Oct 2011.

[20] M. Pajic, S. Sundaram, J. Le Ny, G. J. Pappas, and R. Mangharam, "Closing the loop: A simple distributed method for control over wireless networks," in *Proceedings of the 11th International Conference on Information Processing in Sensor Networks*, ser. IPSN '12. New York, NY, USA: ACM, 2012, pp. 25–36.

[21] K.-D. Kim and P. Kumar, "Cyber physical systems: A perspective at the centennial," *Proceedings of the IEEE*, vol. 100, no. Special Centennial Issue, pp. 1287–1308, May 2012.

[22] L. Sha, S. Gopalakrishnan, X. Liu, and Q. Wang, "Cyber-physical systems: A new frontier," in *Sensor Networks, Ubiquitous and Trustworthy Computing, 2008. SUTC '08. IEEE International Conference on*, June 2008, pp. 1–9.

[23] D. Donoho, "Compressed sensing," *Information Theory, IEEE Transactions on*, vol. 52, no. 4, pp. 1289–1306, 2006.

[24] E. Candes, Y. Eldar, D. Needell, and P. Randall, "Compressed sensing with coherent and redundant dictionaries," *Applied and Computational Harmonic Analysis*, vol. 31, no. 1, pp. 59–73, 2011.

[25] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A simple proof of the restricted isometry property for random matrices," *Constructive Approximation*, vol. 28, no. 3, pp. 253–263, 2008.

[26] J. Tropp and A. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *Information Theory, IEEE Transactions on*, vol. 53, no. 12, pp. 4655–4666, 2007.

[27] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 267–288, 1996.

[28] C. Johnson, "Matrix completion problems: a survey," in *Proceedings of Symposia in Applied Mathematics*, vol. 40, 1990, pp. 171–198.

[29] E. Candès and B. Recht, "Exact matrix completion via convex optimization," *Foundations of Computational mathematics*, vol. 9, no. 6, pp. 717–772, 2009.

[30] E. Candès and Y. Plan, "Matrix completion with noise," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 925–936, 2010.

[31] B. Recht, M. Fazel, and P. Parrilo, "Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization," *SIAM review*, vol. 52, no. 3, pp. 471–501, 2010.

[32] J. Cai, E. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM Journal on Optimization*, vol. 20, no. 4, pp. 1956–1982, 2010.

[33] Z. Lin, M. Chen, and Y. Ma, "The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices," *arXiv preprint arXiv:1009.5055*, 2010.

[34] R. Keshavan, A. Montanari, and S. Oh, "Matrix completion from a few entries," *Information Theory, IEEE Transactions on*, vol. 56, no. 6, pp. 2980–2998, 2010.

[35] S. Pudlewski, A. Prasanna, and T. Melodia, "Compressed-sensing-enabled video streaming for wireless multimedia sensor networks," *Mobile Computing, IEEE Transactions on*, vol. 11, no. 6, pp. 1060–1072, June 2012.

[36] A. Griffin and P. Tsakalides, "Compressed sensing of audio signals using multiple sensors," *Reconstruction*, vol. 3, no. 4, p. 5, 2007.

[37] X. Yu, H. Zhao, L. Zhang, S. Wu, B. Krishnamachari, and V. O. Li, "Cooperative sensing and compression in vehicular sensor networks for urban monitoring," in *Communications (ICC), 2010 IEEE International Conference on*. IEEE, 2010, pp. 1–5.

[38] H. Mamaghanian, N. Khaled, D. Atienza, and P. Vandergheynst, "Compressed sensing for real-time energy-efficient ecg compression on wireless body sensor nodes," *Biomedical Engineering, IEEE Transactions on*, vol. 58, no. 9, pp. 2456–2466, Sept 2011.

[39] J. Cheng, H. Jiang, X. Ma, L. Liu, L. Qian, C. Tian, and W. Liu, "Efficient data collection with sampling in wsns: Making use of matrix completion techniques," in *Global Telecommunications Conference (GLOBECOM 2010), 2010 IEEE*. IEEE, 2010, pp. 1–5.

[40] A. Majumdar and R. K. Ward, "Increasing energy efficiency in sensor networks: blue noise sampling and non-convex matrix completion," *International Journal of Sensor Networks*, vol. 9, no. 3, pp. 158–169, 2011.

[41] S. Li, L. D. Xu, and X. Wang, "Compressed sensing signal and data acquisition in wireless sensor networks and internet of things," *Industrial Informatics, IEEE Transactions on*, vol. 9, no. 4, pp. 2177–2186, Nov 2013.

[42] F. Fazel, M. Fazel, and M. Stojanovic, "Random access sensor networks: Field reconstruction from incomplete data," in *Information Theory and Applications Workshop (ITA), 2012*. IEEE, 2012, pp. 300–305.

[43] G. Tsagkatakis and P. Tsakalides, "Dictionary based reconstruction and classification of randomly sampled sensor network data," in *Sensor Array and Multichannel Signal Processing Workshop (SAM), 2012 IEEE 7th*. IEEE, 2012, pp. 117–120.

[44] A. Fragkiadakis, I. Askoxylakis, and E. Tragos, "Joint compressed-sensing and matrix-completion for efficient data collection in wsns," in *Computer Aided Modeling and Design of Communication Links and Networks (CAMAD), 2013 IEEE 18th International Workshop on*. IEEE, 2013, pp. 84–88.

[45] J. Haupt, W. U. Bajwa, M. Rabbat, and R. Nowak, "Compressed sensing for networked data," *Signal Processing Magazine, IEEE*, vol. 25, no. 2, pp. 92–101, 2008.

[46] H. Hu and Z. Yang, "Spatial correlation-based distributed compressed sensing in wireless sensor networks," in *Wireless Communications Networking and Mobile Computing (WiCOM), 2010 6th International Conference on*. IEEE, 2010, pp. 1–4.

[47] M. Sartipi and R. Fletcher, "Energy-efficient data acquisition in wireless sensor networks using compressed sensing," in *Data Compression Conference (DCC), 2011*. IEEE, 2011, pp. 223–232.

[48] Q. Ling and Z. Tian, "Decentralized sparse signal recovery for compressive sleeping wireless sensor networks," *Signal Processing, IEEE Transactions on*, vol. 58, no. 7, pp. 3816–3827, 2010.

[49] F. Chen, A. P. Chandrakasan, and V. M. Stojanovic, "Design and analysis of a hardware-efficient compressed sensing architecture for data compression in wireless sensors," *Solid-State Circuits, IEEE Journal of*, vol. 47, no. 3, pp. 744–756, 2012.

[50] L. Xiang, J. Luo, and A. Vasilakos, "Compressed data aggregation for energy efficient wireless sensor networks," in *Sensor, Mesh and Ad Hoc Communications and Networks (SECON), 2011 8th Annual IEEE Communications Society Conference on*. IEEE, 2011, pp. 46–54.

[51] C. Luo, F. Wu, J. Sun, and C. W. Chen, "Compressive data gathering for large-scale wireless sensor networks," in *Proceedings of the 15th annual international conference on Mobile computing and networking*. ACM, 2009, pp. 145–156.

[52] J. Cheng, Q. Ye, H. Jiang, D. Wang, and C. Wang, "Stcdg: An efficient data gathering algorithm based on matrix completion for wireless sensor networks," *Wireless Communications, IEEE Transactions on*, vol. 12, no. 2, pp. 850–861, 2013.

[53] G. Quer, R. Masiero, D. Munaretto, M. Rossi, J. Widmer, and M. Zorzi, "On the interplay between routing and signal representation for compressive sensing in wireless sensor networks," in *Information Theory and Applications Workshop, 2009*. IEEE, 2009, pp. 206–215.

[54] C. Caione, D. Brunelli, and L. Benini, "Distributed compressive sampling for lifetime optimization in dense wireless sensor networks," *Industrial Informatics, IEEE Transactions on*, vol. 8, no. 1, pp. 30–40, 2012.

[55] N. Nguyen, D. L. Jones, and S. Krishnamurthy, "Netcompress: Coupling network coding and compressed sensing for efficient data communication in wireless sensor networks," in *Signal Processing Systems (SIPS), 2010 IEEE Workshop on*. IEEE, 2010, pp. 356–361.

[56] A. Y. Alfakih, A. Khandani, and H. Wolkowicz, "Solving euclidean distance matrix completion problems via semidefinite programming," *Computational optimization and applications*, vol. 12, no. 1-3, pp. 13–30, 1999.

[57] A. Javanmard and A. Montanari, "Localization from incomplete noisy distance measurements," *Foundations of Computational Mathematics*, vol. 13, no. 3, pp. 297–345, 2013.

[58] V. N. Ekambaram and K. Ramchandran, "Non-line-of-sight localization using low-rank+ sparse matrix decomposition," in *Statistical Signal Processing Workshop (SSP), 2012 IEEE*. IEEE, 2012, pp. 317–320.

[59] R. Rangarajan, R. Raich, and A. O. Hero, "Euclidean matrix completion problems in tracking and geo-localization," in *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*. IEEE, 2008, pp. 5324–5327.

[60] D. Milioris, G. Tzagkarakis, A. Papakonstantinou, M. Papadopouli, and P. Tsakalides, "Low-dimensional signal-strength fingerprint-based positioning in wireless lans," *Ad hoc networks*, vol. 12, pp. 100–114, 2014.

[61] S. Nikitaki, G. Tsagkatakis, and P. Tsakalides, "Efficient recalibration via dynamic matrix completion," in *Machine Learning for Signal Processing (MLSP), 2013 IEEE International Workshop on*. IEEE, 2013, pp. 1–6.

[62] M. Mardani, G. Mateos, and G. B. Giannakis, "Unveiling anomalies in large-scale networks via sparsity and low rank," in *Signals, Systems and Computers (ASILOMAR), 2011 Conference Record of the Forty Fifth Asilomar Conference on*. IEEE, 2011, pp. 403–407.

[63] R. Paffenroth, P. Du Toit, R. Nong, L. Scharf, A. P. Jayasumana, and V. Bandara, "Space-time signal processing for distributed pattern detection in sensor networks," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 7, no. 1, pp. 38–49, 2013.

[64] B. Zhang, X. Cheng, N. Zhang, Y. Cui, Y. Li, and Q. Liang, "Sparse target counting and localization in sensor networks based on compressive sensing," in *INFOCOM, 2011 Proceedings IEEE*. IEEE, 2011, pp. 2255–2263.

[65] Q. Ling, Y. Xu, W. Yin, and Z. Wen, "Decentralized low-rank matrix completion," in *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*. IEEE, 2012, pp. 2925–2928.

[66] S. M. Kay, *Fundamentals of Statistical Signal Processing. Estimation Theory*. Prentice Hall, 1993.

[67] D. Simon, *Optimal State Estimation*. Wiley-Interscience, 2006.

[68] A. Tahbaz-Salehi and A. Jadbabaie, "Consensus over ergodic stationary graph processes," *Automatic Control, IEEE Transactions on*, vol. 55, no. 1, pp. 225–230, Jan 2010.

[69] C. Asensio-Marco and B. Beferull-Lozano, "Link scheduling in sensor networks for asymmetric average consensus," in *Signal Processing Advances in Wireless Communications (SPAWC), 2012 IEEE 13th International Workshop on*, June 2012, pp. 319–323.

[70] R. Olfati-Saber, "Distributed kalman filter with embedded consensus filters," in *Decision and Control, 2005 and 2005 European Control Conference. CDC-ECC '05. 44th IEEE Conference on*, Dec 2005, pp. 8179–8184.

[71] C. Aggarwal, *Managing and Mining Uncertain Data*. Springer, 2009.

[72] T. Thanh, P. Liping, D. Yanlei, M. Andrew, and L. Anna, "CLARO: Modeling and processing uncertain data streams," *The VLDB Journal*, vol. 21, no. 5, pp. 651–676, 2012.

[73] M.-Y. Yeh, K.-L. Wu, P. Yu, and M.-S. Chen, "PROUD: a probabilistic approach to processing similarity queries over uncertain data streams," in *Proceedings of the 12th International Conference on Extending Database Technology: Advances in Database Technology*. Saint-Petersburg, RU: ACM New York, 2009, pp. 684–695.

[74] M. Datar, A. Gionis, P. Indyk, and R. Motwani, "Maintaining stream statistics over sliding windows," in *Proceedings of the 13th annual ACM-SIAM Symposium on Discrete Algorithms*. San Francisco, CA: SIAM, 2002, pp. 635–644.

[75] J. Gehrke, F. Korn, and D. Srivastava, "On computing correlated aggregates over continual data streams," in *Proceedings of the 2001 ACM SIGMOD International Conference on Management of Data*. Santa Barbara, CA: ACM New York, 2001, pp. 13–24.

[76] Y. Zhu and D. Shasha, "StatStream: Statistical monitoring of thousands of data streams in real time," in *Proceedings of the 28th International Conference on Very Large Data Bases*. Hong Kong, China: VLDB Endowment, 2002, pp. 358–369.

[77] K. Ishikawa and J. Loftus, *Introduction to Quality Control*. Tokyo: 3A Corporation, 1990.

[78] I. Farrance and R. Frenkel, "Uncertainty of measurement: A review of the rules for calculating uncertainty components through functional relationships," *Clinical Biochemist Reviews*, vol. 33, no. 2, pp. 49–75, 2012.

[79] B. Taylor and C. Kuyatt, "Guidelines for evaluating and expressing the uncertainty of nist measurement results," *NIST Technical Note 1297*, 1994.

[80] E. J. Gumbel, *Statistics of Extremes*. Courier Dover Publications, 2004.

[81] J. Pickands, "Statistical inference using extreme order statistics," *The Annals of Statistics*, vol. 3, no. 1, pp. 119–131, 1975.

[82] L. de Haan and A. Ferreira, *Extreme Value Theory: An Introduction*. Springer, 2006.

[83] D. Rafiei and A. Mendelzon, "Similarity-based queries for time series data," in *Proceedings of ACM SIGMOD International Conference on Management of Data*. Tucson, Arizona: ACM New York, 1997, pp. 13–25.

[84] Y. Sakurai, S. Papadimitriou, and C. Faloutsos, "BRAID: Stream mining through group lag correlations," in *Proceedings of ACM SIGMOD International Conference on Management of Data*. Baltimore, Maryland: ACM New York, pp. 599–610.

[85] D. 8.2, "Hydrobionets: Demonstration activities of low-scale wbn test-bed," http://www.hydrobionets.eu/index.php/deliverables, Ateknea Solutions, Tech. Rep., 2014.

[86] D. 4.2, "Hydrobionets: Network protocol design," http://www.hydrobionets.eu/index.php/deliverables, Kungliga Tekniska Hoegskolan (KTH), Tech. Rep., 2013.

[87] A. Dunkels, B. Gronvall, and T. Voigt, "Contiki - a lightweight and flexible operating system for tiny networked sensors," in *Local Computer Networks, 2004. 29th Annual IEEE International Conference on*, Nov 2004, pp. 455–462.

[88] *The Contiki Operating System: Version 2.6*, 2012. [Online]. Available: http://www.contiki-os.org/start.html

[89] "Advanticsys wireless sensor modules," http://www.advanticsys.com/wiki/, 2013.

[90] A. Seliniotaki, G. Tzagkarakis, V. Christofides, and P. Tsakalides, "Stream correlation monitoring for uncertainty-aware data processing systems," in *Proceedings of the 5th International Conference on Information, Intelligence, Systems and Applications (IISA '14)*, Chania, Greece, July 7–9 2014.