Annual Reviews in Control 43 (2017) 1–64

Contents lists available at ScienceDirect

Annual Reviews in Control

journal homepage: www.elsevier.com/locate/arcontrol

Review article

Systems & Control for the future of humanity, research agenda: Current and future roles, impact and grand challenges

Francoise Lamnabhi-Lagarrigue[a,*], Anuradha Annaswamy[b], Sebastian Engell[c], Alf Isaksson[d], Pramod Khargonekar[e], Richard M. Murray[f], Henk Nijmeijer[g], Tariq Samad[h], Dawn Tilbury[i], Paul Van den Hof[j]

[a] CNRS Laboratory of Signals and Systems, CentraleSupelec, University Paris-Sud, University Paris-Saclay, Gif-sur-Yvette, France
[b] Active-Adaptive Control Laboratory, Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, USA
[c] Process Dynamics and Operations Group, Department of Biochemical and Chemical Engineering, TU Dortmund, Dortmund, Germany
[d] ABB AB, Corporate Research, Västerås, Sweden
[e] Department of Electrical Engineering and Computer Science and Office of Research, University of California, Irvine, CA, USA
[f] Control and Dynamical Systems, California Institute of Technology, Pasadena, USA
[g] Department of Mechanical Engineering, Technische Universiteit Eindhoven, Eindhoven, The Netherlands
[h] Technological Leadership Institute, University of Minnesota, USA
[i] Mechanical Engineering Department, University of Michigan, USA
[j] Department of Electrical Engineering, Technische Universiteit Eindhoven, Eindhoven, The Netherlands

ARTICLE INFO

Article history:
Received 2 April 2017
Accepted 2 April 2017
Available online 26 April 2017

Keywords:
Systems & Control
Research challenges
Critical societal challenges

ABSTRACT

Following in the footsteps of the renowned report "Control in an Information Rich World," Report of the Panel on "Future Directions in Control, Dynamics, and Systems" chaired by Richard Murray (2002), this paper aims to demonstrate that **Systems & Control is at the heart of the Information and Communication Technologies to most application domains**. As such, Systems & Control should be acknowledged as a priority by funding agencies and supported at the levels necessary to enable technologies addressing critical societal challenges. A second intention of this paper is to present to the industrials and the young research generation, a global picture of the societal and research challenges where the discipline of Systems & Control will play a key role. Throughout, this paper demonstrates the extremely rich, current and future, cross-fertilization between five critical societal challenges and seven key research and innovation Systems & Control scientific challenges. This paper is authored by members of the IFAC Task Road Map Committee, established following the 19th IFAC World Congress in Cape Town. Other experts who authored specific parts are listed below.

© 2017 Elsevier Ltd. All rights reserved.

Contents

* Corresponding author.
E-mail address: lamnabhi@l2s.centralesupelec.fr (F. Lamnabhi-Lagarrigue).

http://dx.doi.org/10.1016/j.arcontrol.2017.04.001
1367-5788/© 2017 Elsevier Ltd. All rights reserved.

Contributions

The contributors to this paper are the following:

Ruzena Bajcsy, University of California, Berkeley, USA, **4.7**
B. Ross Barmish, University of Wisconsin-Madison, USA, **5.21**
Sergio Bittanti, Politecnico di Milano, Italy, **5.6**
Richard D. Braatz, Massachusetts Institute of Technology, Cambridge, USA, **3.4** & **5.10**
Hans Butler, ASML, Veldhoven, & Eindhoven University of Technology, Eindhoven, the Netherlands, **5.14**
Peter E. Caines, McGill University, Montréal, Canada, **4.3**
Marco Campi, University of Brescia, Italy, **4.2**
Carlos Canudas-de-Wit, CNRS GIPSA-Lab, Grenoble, France, **5.2**
Elisa Capello, Politecnico di Torino, CNR-IEIIT, Italy, **5.16**
Christos G. Cassandras, Boston University, USA, **5.17**
Tianyou Chai, Northeastern University. Shenyang, China, **5.7**
Antoine Chaillet, CentraleSupelec, Gif-sur-Yvette, France, **5.9**
Gilney Damm, University of Evry Val d'Essonne, France, **5.6**
Alexandre Dolgui, IMT Atlantique, Nantes, France, **5.13**
Francis J. Doyle III, Harvard University, Cambridge, USA, **5.11**
Marika Di Benedetto, University of L'Aquila, Italy, **4.1**
Alessandro D'Innocenzo, University of L'Aquila, Italy, **4.1**
Lars Eriksson, Linköping University, Sweden, **3.1**
Antonella Ferrara, University of Pavia, Italy, **5.2**
Yves Fregnac, CNRS UNIC, Gif-sur-Yvette, France, **5.9**
Michel Gevers, Université Catholique de Louvain, Belgium, **4.2**
Roger Goodall, Loughborough University, UK, **5.3**
Graham C. Goodwin, University of Newcastle, Australia, **5.11**
Timothy Gordon, University of Lincoln, UK, **5.2**
Giorgio Guglieri, Politecnico di Torino, Italy, **5.16**
Shinji Hara, University of Tokyo, Japan, **3.4**
Marcel Heertjes, ASML, Veldhoven, & Eindhoven University of Technology, Eindhoven, the Netherlands, **5.14**
Hakan Hjalmarsson KTH Royal Institute of Technology, Stockholm, Sweden, **4.2**
Zhong-Sheng Hou, Beijing Jiaotong University, Beijing, China, **4.2**
Jonathan How, Massachusetts Institute of Technology, Cambridge, USA, **5.16**
Iqbal Husain, NC State University, Raleigh, **5.6**
Iasson Karafyllis, National Technical University of Athens, Greece, **4.3**
Ilya Kolmanovsky, University of Michigan, USA, **5.4**
Péter Korondi, Budapest University of Technology and Economics, Hungary, **5.16**
Miroslav Krstic, University of California, San Diego, USA, **4.3**
Kwang Y. Lee, Baylor University, U.S.A., **3.2**
Kim Listmann, ABB Corporate Research Center, Ladenburg, Germany, **5.15**
Lennart Ljung, Linköping University, Sweden, **4.2**
Iven Michiel Yvonne Mareels, University of Melbourne, Australia, **3.3**
Padmanabhan K. Menon, Optimal Synthesis Inc., Los Altos, USA, **5.1**
Mariana Netto, IFSTTAR, Versailles-Satory, Versailles, France, **4.7**
Lydie Nouvelière, University of Evry Val d'Essonne, France, **5.2**
Patrick Panciatici RTE, Versailles, France, **5.6**
Elena Panteley, CNRS L2S, Gif-sur-Yvette, France, **4.3**

Ioannis Ch. Paschalidis, Boston University, USA, **4.2**
Radoslav Paulen, TU Dortmund University, Dortmund, Germany, **3.5** & **4.5**
Rik Pintelon, Vrije Universiteit Brussel, Belgium, **4.2**
Marios M. Polycarpou, University of Cyprus, Cyprus, **4.4**
Anton V. Proskurnikov, TU Delft, the Netherlands, IPME RAS & ITMO University, St. Petersburg, Russia, **5.20**
Michel Reniers, Eindhoven University of Technology, the Netherlands, **4.5**
Jean-Pierre Richard, Centrale Lille, Villeneuve d'Ascq, France, **3.1** & **5.16**
Bernard Riera, Université de Reims Champagne Ardenne, Reims, France, **5.13**
Pierre Rouchon, Ecole des Mines de Paris, France, **5.19**
Johan Schoukens, Vrije Universiteit Brussel, Belgium, **4.2**
Christian Sonntag, TU Dortmund University, Dortmund, Germany, **3.5** & **4.5**
Eduardo Sontag, Rutgers University, New-Brunswick, USA, **5.12**
Sarah Spurgeon, University College London, London, UK, **5.10**
Petr Stluka, Honeywell HBT Global Labs in Prague, Czech Republic, **5.18**
Roberto Tempo (*), CNR-IEIIT, Politecnico di Torino, Italy, **5.16** & **5.20**
Kalevi Tervo, ABB Marine & Ports, Helsinki, Finland, **5.5**
Haydn Thompson, Haydn Consulting Ltd., Sheffield, UK, **4.5**
Peter Wellstead (*), National University of Ireland, Co. Kildare, Ireland, **3.4**
Hong Yue, University of Strathclyde, **5.6**

Advices from Panganamala Ramana Kumar, Professor and College of Engineering Chair in Computer Engineering, and from Alkis Konstantellos, *Complex Systems, Advanced Computing and Control*, European Commission (retired), Brussels, are gratefully acknowledged. The authors are also very grateful to Janan Zaytoon (IFAC President) and Francis J. Doyle III (IFAC TB Chair) for their strong support of the IFAC Road Map Task Force.

(*) Regrettably and sadly, Professor Peter Wellstead and Professor Roberto Tempo who greatly contributed to this Control Research Agenda, passed away suddenly and unexpectedly on 24 June 2016 and 14 January 2017, respectively.

1. Introduction

Control technologies are everywhere – aircraft and spacecraft, chemical process plants, manufacturing, homes and buildings, automobiles and trains, GPS, cellular telephones and – these and other complex systems testify to the ubiquity of Systems & Control technology. Some artefacts of modern times would simply not be possible without Systems & Control. And for many others, substantial, even revolutionary, advances in performance, safety, reliability and affordability have been achieved as a result of the ingenuity and effort of Systems & Control researchers and engineers.

As we move deeper into the 21st century, new, complex and multifaceted challenges face humanity. These challenges in turn demand inter- and cross-disciplinary research and development. While Systems & Control, a foundational discipline for analysis and synthesis of complex systems, is uniquely positioned to participate in this endeavour, new investments, topics and ways of educating, working and exchanging are needed to address and solve these challenges. This article elaborates on these themes, presents a research agenda for the field and offers recommendations for government agencies and the research community.

1.1. Systems & Control: a rich history

Starting as a collection of empirical rules for designing servomechanisms, the field of Systems & Control has matured into a rigorous, erudite discipline in engineering science. The field has played an important enabling role in almost all major technological evolutions – from steam engines to high-speed trains and driver-assisted automobiles, high performance aircraft, rockets to spacecraft, wired telephony to cellular telephones, cameras to neuroimaging, agile manufacturing to robotics, medical devices to telesurgery, and so on. The foundations of Systems & Control have enabled improved performance, speed, efficiency, reliability and stability, as well as reduced energy consumption, costs and emissions,, with tangible benefits most apparent in applications associated with aerospace, automotive, process and manufacturing industries.

These benefits have been realized through the use of Systems & Control which typically consists of: (i) modelling and analysis of the underlying physical phenomena along with selection of sensors and actuators, (ii) development of control strategies that enable intended behaviour in an optimal fashion while satisfying constraints and minimizing resources consumed, (iii) validation and verification of the control performance using simulation studies of a suite of models with increasing fidelity, and (iv) implementation. Major challenges in all applications arise from size and scale, complex dynamics, inherent uncertainties, distributed and diverse processes and phenomena, and safety and reliability requirements. Improvements are made possible through modelling, analysis and design tools that capture the trade-offs between fidelity and tractability, advances in control methods that enable coordination and management of these systems and development of architectures and algorithms that ensure robustness, optimality, adaptability and stability.

A fundamental characteristic of the field of Systems & Control research is its rich intersection of engineering and mathematics – while the problem formulations stem from engineering systems, the tools and techniques employed are grounded in various subfields of mathematics and computer science. It is this abstract feature that enables the broad applicability of the field, in addition to tackling complex multi-domain systems seamlessly and developing the underlying control technologies.

The performance and resilience of nearly every technical system nowadays result from design of both the physical hardware and it's the management and control algorithms, usually implemented as embedded systems that allow the system to function efficiently despite the influence of disturbances, degradation, etc. Control engineers are therefore key players on development teams for these high-tech systems. Deeply involved with the operation and constraints of the components as well as the overall system, they not only take responsibility for the control elements but also provide a dynamic systems' view during the specification and development process.

The evolution, grand challenges, theoretical advances, maturation and success stories that the field has witnessed have been documented over the years in numerous textbooks spanning many disciplines, several surveys, overviews, position papers, reports and encyclopaedias, most notable of which are Aström and Kumar (2014), Murray (2003), Samad and Annaswamy (2014), HYCON2 Report (2011) and Baillieul and Samad (2017). The survey in Aström and Kumar (2014) provides a comprehensive historical account of the development of the field of control, as well as a perspective on current directions and future opportunities, and reflections on the dynamic interaction between theory and practice. The report (Murray, 2003) looks towards the future, focusing on research directions in Systems & Control, circa 2002. The on-line publication (Samad and Annaswamy, 2014) provides an overview, survey success stories and research challenges facing the field in both its 2011 and 2014 editions. The 2011 European report (HYCON2 Report 2011) introduces new key research challenges. A comprehensive collection of all methods and applications of Systems & Control is reported in the encyclopaedia (Baillieul and Samad, 2017). All these publications testify to the rigorous, analytically grounded, rich, vibrant and active field encompassed by Systems & Control.

1.2. A glimpse into future and changing paradigms

The present witnesses Systems & Control at its most active and promising juncture. New directions in Systems & Control are emerging from its relevance and potential to address societal grand challenges in areas of health and medicine; energy and climate; sustainability and development; productivity, inequality and economic growth. Large-scale infrastructures are envisioned in energy (smart grids), transportation (connected V-to-X systems), manufacturing (smart, connected and lean), urban planning (smart cities), and financial services (smart services), to name a few. Systems & Control is also critical to emerging paradigms in engineered systems ranging from nano to micro to large to planet scale. New methodologies and advances on the theoretical front allow Systems & Control to play a pivotal role in such engineered systems. There is great promise for Systems & Control in the *Internet of Things* (IoT), the network of physical objects or "things" embedded with electronics, software, sensors, and network connectivity, and *Industry 4.0,* the next transformative change in manufacturing that leverages pervasive sensing, distributed control, and robust, seamless connectivity, data analytics, etc. (http://www.mckinsey.com/business-functions/operations/our-insights/manufacturings-next-act, 2017). A case in point is the evolving field of Cyber-Physical Systems (Johansson et al., 2014).

New technologies, e.g. flexible manufacturing and autonomous driving, lead to complexity levels that cannot be mastered by systematic design using mathematical models and simulation tools alone. The technologies must be tested in large-scale demonstrators to gain experience, detect design or implementation flaws and demonstrate functionality for end users and the public. Systems & Control research is already being implemented and evaluated on large testbeds or "living labs" in the areas of transportation, microgrids, manufacturing, robotics, etc. Scientists should get even more involved in such demonstrators, and the experience gained here will trigger research on new fundamental questions in Systems & Control.

While Systems & Control has always evolved through collaborations with other fields of communication, computing, sensing and actuation, a widening appreciation of the need for Systems & Control has become increasingly apparent as more large-scale, interdependent, multi timescale, safety-critical and mission-critical systems are developed or envisioned. Dynamics, feedback, stability and optimality are increasingly recognized as pervasive and central properties of complex systems. Likewise, the Systems & Control community is also being perceived differently: in addition to providing rigor to system analysis, they are now viewed as crucial for the design, development, overall planning and operations of complex systems. The time is therefore apt for us to pause and articulate this role in all its breadth and depth. Hence this article.

A three-way perspective of this new role – top-down, bottom-up and convergence-based – is provided here. From a top-down applications perspective, societal grand challenge problems are inherently beyond the scope of any single discipline of knowledge. Indeed, they require creative collaborations among many disciplines. For example, progress in healthy aging will require partnerships among biologists, chemists, physicists, mathematicians, and engineers. Thus, Systems & Control researchers can and will contribute to cancer research but it will be in cooperative teams with experts from other fields. Similarly, transition to low carbon economy will require collaborations among engineers, physicists, chemists, biologists, social scientists, and humanists. Again, Systems & Control researchers can and will play significant roles in collaborative teams of researchers.

At the same time, from a bottoms-up curiosity-driven perspective, significant progress in closely allied disciplines offers new frontiers for creative research. For example, Systems & Control research has strong affinity with fields such as communications, signal processing, machine learning, computational science, applied mathematics, etc. Advances in algorithms, devices (semiconductor, optical, sensors, actuators), computing capabilities such as distributed and cloud computing, and availability of large amounts of data are changing these fields and offering the Systems & Control community new opportunities to forge productive collaborations that build on such transformative developments.

A third and new perspective, the convergence paradigm (Convergence, National Academy Press 2014; Convergence: The Future of Health 2016; Sharp & Leshner, 2014), offers an exciting vision of grand challenge driven multidisciplinary research. The intellectual merit of the convergence paradigm lies in its call for deep integration of knowledge bases, tools and techniques for discovery, and most importantly, "modes of thinking" among experts from physical sciences, biological sciences, computing and engineering and social sciences to create new pathways for creation of knowledge. These pathways would differ dramatically from the traditional models for knowledge creation and thereby produce new paradigms for scientific and engineering research. The Systems & Control field is uniquely positioned for critical involvement in the creation and evolution of these pathways. A case in point is the evolving field of Cyber-Physical Systems which requires convergence of controls, communications, networks and computing with domain experts from biology, chemistry, aerospace or mechanical engineering with applications in transportation, energy, water, healthcare and manufacturing. Roles of humans in technical systems are becoming more complex, beyond users and consumers, as active agents, operators, decision-makers and enablers of efficient and resilient infrastructures. Thus, in addition to engineering, physical and computational sciences, social, behavioural and economic sciences add new dimensions to the convergence paradigm, as exemplified by Cyber-Physical & Human Systems (CPHS). Social networks and their emerging properties at a systems level pose fascinating challenges. Their intersection with finance, markets, and economics at large, and the resulting socio-technical-economic systems beg to be analysed, designed, understood and leveraged. This article addresses the emerging roles of Systems & Control and its impact on the future of humanity.

These new conceptual, intellectual and organizational frameworks can be critically useful in enabling Systems & Control researchers to engage in collaborative team research. This is true at the individual researcher level as well as at the community (of researchers) level. Such practices and engagements can complement curiosity-driven disciplinary research that will advance the knowledge base in Systems & Control. To be successful, the community must be open to new types of publications. It also needs to review and reconsider traditional modes of research and researcher evaluations for career progress. Finally, the community must engage in new, more fluid forms of organizational structures that go beyond traditional departmental structures. It is very encouraging that such changes are already beginning to happen.

1.3. Organization of this paper

Section 2 provide a brief review of the advances in the field accomplished over the past few decades. The main body of this

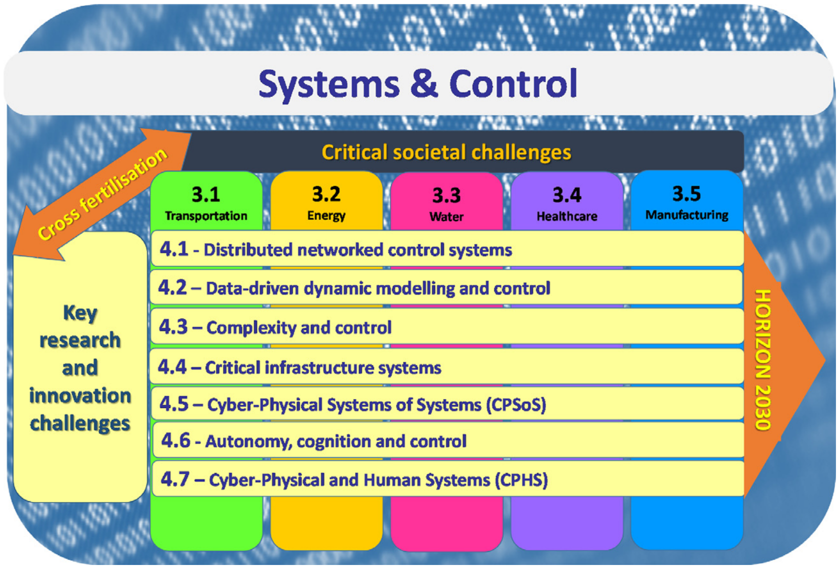

Fig. 1. Systems & Control matrix: cross fertilization of five critical societal challenges by seven key research and innovation challenges.

Research Agenda is described in Section 3 and is grouped under the five critical societal challenges of Transportation, Energy, Water, Healthcare and Manufacturing (see the matrix in Fig. 1). Seven grand challenges precipitated by these sectors are described in Section 4, while Section 5 illustrates twenty-two examples of the matrix of these challenges where Systems & Control will have a strong impact in the coming years. In each of these examples, i) the current state, ii) the needs for the future and the obstacles, and iii) the ways Systems & Control can contribute to facing these challenges, are described. Section 6 offers some operational recommendations to provide the means to develop this extremely important scientific and technological discipline whose critical role in Information and Communication Technologies (ICT) is essential to meet the policies in the future. Finally, Section 7 includes testimonials from industry members regarding the role and the societal impact of Systems & Control.

2. Advances in Systems & Control in the past fifteen years

In April 2000, the Panel on Future Directions in Control and Dynamical Systems was formed by the US Air Force Office of Scientific Research (AFOSR) to provide a renewed vision of future challenges and opportunities in the field, along with recommendations to government agencies, universities and research organizations for how to ensure continued progress in areas of importance to the industrial and defence base. The Panel prepared a report which was completed in April 2002 and published by SIAM in 2003 (Murray, 2003). The intent of the report was to raise the overall visibility of research in Systems & Control, highlight its importance in applications of national interest, and indicate some of the key trends which are important for continued vitality of the field.

The panel made five major recommendations:

- Substantially increase research aimed at the integration of control, computer science, communications, and networking. This includes principles, methods and tools for modelling and control of high level, networked, distributed systems, and rigorous techniques for reliable, embedded, real-time software.
- Substantially increase research in control at higher levels of decision making, moving toward enterprise-scale systems. This includes work in dynamic resource allocation in the presence of uncertainty, learning and adaptation, and artificial intelligence for dynamic systems.
- Explore high-risk, long-range applications of control to new domains such as nanotechnology, quantum mechanics, electromagnetics, biology and environmental science. Dual investigator, interdisciplinary funding was suggested as a particularly useful mechanism in this context.
- Maintain support for theory and interaction with mathematics, broadly interpreted. The strength of the field relies on its close contact with rigorous mathematics, and this was felt to be increasingly important in the future.
- Invest in new approaches to education and outreach for the dissemination of control concepts and tools to non-traditional audiences. The community should do a better job of educating a broader range of scientists and engineers on the principles of feedback and the use of control to alter the dynamics of systems and manage uncertainty.

The field of controls has moved forward in many of directions that are largely consistent with some of the recommendations from the Panel's report and exemplified by the IEEE CSS online publication on the Impact of Control Technology (Samad and Annaswamy, 2014). Highlights of these advances include:

- In the areas of control integration, computer science, communications and networking, there have been substantial developments resulting in focused research efforts, including establishment of the IFAC technical area in networked control systems in 2005 and creation of national initiatives in the US in Cyber-Physical Systems (2008) and robotics (2011). There has been less effort in the area of control at higher levels of decision-making and enterprise-scale systems, although some of the work in formal methods is related to this as well as new initiatives in smart infrastructures (electric grid, transportation networks, etc.). Optimization-based control (including model predictive control) is now one of the dominant techniques for control across a range of applications. These techniques use knowledge of the current system state, online dynamic models of system behaviour, and constraints on inputs and states to provide high performance control systems that react to changes in the system and environment. Whereas a decade ago the main application of model predictive control was in chemical process control, advances in computing and communications have enabled the application of similar techniques to a wide range of systems, from manufacturing systems and supply chains, to aircraft engines and smart buildings. A major advance occurred with the development of explicit model predictive control, which allows offline computation of feedback control laws that consider nonlinear dynamics and constraints for systems of modest state dimension.

- Another advancement involved the use of layering and abstraction in control systems. Layering is an essential architectural feature of many complex systems, nicely represented in the multi-layer protocols used in network-based communications. Layering and abstraction enable modularity and are a key approach to implementing product families that exploit this modularity. Multi-layer architectures that include low-level feedback controllers, higher-level trajectory generation (including horizon optimization receding horizon) and supervisory control that accounts for complex temporal specifications, are now becoming commonplace in many applications such as autonomous vehicles and smart grids.
- The recent establishment of major investments in Cyber-Physical Systems (Kim & Kumar, 2012) focuses on designing systems that combine information and physics. It represents a better coupling of research in computer science, controls, communications and networking, with applications in areas such as aerospace, automotive, chemical processes, civil infrastructure, energy, healthcare, manufacturing, transportation, entertainment and consumer appliances. Work in this area builds not only on work from the 1980s in discrete event systems, but makes use of advances in computer science in the intervening decades (real-time systems, embedded systems and formal methods), leading to new approaches for analysis, design and synthesis which are now being applied in robotics, aerospace systems and many other areas.
- In high-risk, long-range applications of control, there continue to be forays by researchers into new application areas including quantum control systems, synthetic biology and geoengineering. In each of these areas, control theory researchers must learn the language of the application area, while at the same time, researchers in the specific domains must learn about existing tools in control theory, and together articulate new challenges for control and tractable approaches that are driven by the specifics of those disciplines. As an example, the type of stochastic behaviours in biological systems where Poisson processes are a natural representation, require different tools than those provided by stochastic control theory in which additive white Gaussian noise (AWGN) models have been developed, motivated primarily by characteristics of electrical systems.
- Finally, in education and outreach, there have been efforts over the past decade to develop courses and textbooks that reach a broader audience. For example, two recent textbooks are specifically oriented at non-engineers (Åström & Murray, 2008; Albertos & Mareels, 2010), and massive online open courses (MOOCs) have been recently offered to very large numbers of participants.

In addition to the above directions that build on the recommendations of the Panel's 2003 report, several new themes and roles are emerging for Systems & Control theory and technology:

- In computer science, the sub-disciplines of cloud computing, "big data," and machine learning have made exceptional advances in the last decade. However, these have not been major forces within the controls community. Advances in these areas have facilitated work on enterprise-level systems, but more importantly have allowed a focus on data as a primary object of study. One area in which data can play a major role in control is building management systems, where learning both the usage patterns and the thermal dynamics can be exploited.
- In connection with large-scale, networked data collection and computing, another important set of issues are those of security and privacy. Security of supervisory control and data acquisition (SCADA) systems gained international prominence with the release of Stuxnet in 2010. This has generated greater awareness of the important role that control systems must play in guaranteeing security, going beyond cybersecurity and incorporating the physics of the system into the analysis (so-called cyber-physical security).
- Another topic that was only minimally addressed in the 2003 report is interaction with humans. While the report mentioned "human on the loop" in military applications, the more integrated role of human interaction in robotics, healthcare, and other fields where control systems directly interact with individuals and groups remains an open area of research.
- Finally, energy and sustainability have arisen as major topics of global importance where control theory and technology will clearly play an important role. Examples of distributed optimization and networked control are emerging not only in smart grids and smart cities, but also in smart buildings, wind turbines, transportation systems and global emissions planning.

This paper explores these emerging areas of research by examining them from societal and application perspectives.

3. The essential role of Systems & Control in meeting critical societal challenges

This section analyses the five (vertical) critical challenges shown in Fig. 1 and demonstrates the essential (present and future) role of Systems & Control in meeting these challenges.

3.1. Transportation

Modern society is, to a large extent, built on transportation of both people and goods, and it is amazing how well the infrastructure functions. Large amounts of food and other goods are made available, waste is transported away, and masses of people commute to and from work both by private and public transportation. For instance, sugar peas from Kenya, tenderloin from Argentina, potatoes from Sweden and strawberries from Spain can be produced and cost effectively transported, ending up on one's dinner table anywhere in the world! In the 200 years since mechanized transportation was introduced, the capacity, speed, efficiency and geographical coverage of transport systems have improved dramatically. The goal of moving passengers and freight faster, in greater quantities, safely, efficiently and cleanly, remains the core motivation to improving transportation technology. The integration of mechanical engineering with electronics, along with the decisive contribution of Systems & Control engineering, was the key to the fantastic advances made by motor vehicles in the last decades concerning emissions, fuel consumption, safety, diagnostics and comfort. and will, of course, be the key to future progress too.

Transportation is fundamental to our society as we know it, supporting increased mobility demands for passengers and freight. But its negative effects include growing levels of motorization and congestion. In addition, the transportation sector is linked to environmental problems, impactingclimate change, air quality, noise, water quality, soil quality, biodiversity and land take. In 2014, the World Health Organization estimated that outdoor air pollution was responsible for 3.7 million annual premature deaths. Exposure to vehicle emissions or other mobile sources is noted as a major contributor to this global disease burden. The transport system has a major impact on our lives and environment, e.g.:

- **Air and water quality**. Vehicles and vessels produce pollution in the form of gas and particulate matters that affect air and water quality causing damage to human health. Toxic air pollutants are associated with cancer, cardiovascular, respiratory and neurological diseases. Acid precipitation detrimentally impacts the built environment, reduces agricultural crop yields and causes forest decline.
- **Noise.** Noise pollution represents the general effect of irregular and chaotic sounds. It is traumatizing for the hearing organ and may affect quality of life by its unpleasant and disturbing character. Increasing noise levels have a negative impact on the urban environment, reflected in falling land values and loss of productive land uses.
- **Climate change**. The activities of the transport industry release several million tons of gases each year into the atmosphere. To add insult to injury, we are consuming finite resources to fuel the transports.
- **Safety.** Albeit with a decreasing trend, world-wide road fatalities were estimated at 1.3 million in 2014. Although significantly lower in total number of fatalities, accidents caused by air, rail and marine transport are typically larger, resulting in more media attention. They also typically lead to much higher losses in terms of material damage. Human error accounts for a significant percentage of these accidents which is why further automation is expected.

By 2030, the number of motor vehicles on the world's roads will roughly double from 2010—from 1.4 billion to about 2.8 billion cars, trucks, motorcycles and other vehicles. Without efforts to improve the energy efficiency of and minimize the pollutant emissions from those vehicles, there will be costly consequences of such rapid increase for local air quality and public health and for the global climate. These challenges are being met with innovative powertrains, advanced after-treatment devices that are all relying on modern controls. Historically the controls community has had a high involvement in the development of clean vehicles. A fundamental contribution in this direction was the introduction of three-way catalysts for gasoline engines that rely on feedback control. This technology is still state of the art, but has naturally evolved and been refined over the years from a single sensor system to multi sensor feedback. This is a formidable example of revolution followed by evolution. Furthermore, the accurate control of injection systems with precise control of the injection process attenuates the knocking noise in diesel engines.

Moreover, the pending depletion of fossil fuels, increased oil prices and the need for reducing carbon dioxide emissions, dramatically conflict with an increasing demand for mobility, posing new and complex challenges to the transportation community. This is further stressed when considering the increasing demands for transportation in developing countries. To meet these demands, there are many on-going efforts to make vehicles function as efficiently and cleanly as possible.

Rail transport is arguably today's most environmentally friendly means of transport, at least when it is electrical and the electrical power comes from renewable generation. In the growing megacities of the world, rail is a crucial part of the infrastructure to handle the fast-growing transport needs. Additional high-speed trains and underground networks are to be expected.

Environmental concerns have motivated legislative action by governments around the world to reduce emissions. Global commitments to CO2 reduction drive the development for improved fuel economy. Customers demand performance and efficiency. All these objectives must be delivered at low cost with high reliability. Combinations of demands from legislators and customers are driving technology, and as a result, modern vehicle powertrains must satisfy challenging and often conflicting requirements. Fig. 2 shows the drive for future CO2 reductions as Corporate Average Fuel Economy (CAFE) standards.

While uncertainties regarding the future of oil production exist, there are indications that the end of the dominance of the internal combustion engine is approaching. In the meantime, we must note that is difficult to outperform something that has evolved, been optimized, and refined over more than 100 years. In the near future, most vehicles will still be equipped with combustion engines, but changes are coming with the increase of electric vehicles and the introduction of fuel cell vehicles on the markets. As oil production is expected to peak and gradually decline, energy prices are expected to continue their upward trend, triggering the most important technological transition in transportation since the introduction of the automobile. Yet, energy prices are prone to significant fluctuations. In this environment, we have the following promising technologies:

- **Automated/Intelligent transport systems.** Development of information and communication technologies (ICT) has the potential to improve the speed, efficiency, safety and reliability of movement by relying upon complete or partial automation of the vehicle, transshipment (for freight) and control. There is an evolution from mechatronic systems to cyber-physical systems.
- **Alternative fuels (including hybrid electric).** Fossil fuels (diesel and gasoline) have been the main energy carrier for transportation, but alternatives are appearing: natural gas and bio fuels, as well as hydrogen and electrified vehicles, are moving in. There is also a clear trend towards electrification in all transport, save for aircraft which will require other alternatives to fossil fuels.

Relating to these trends is a conjecture that all future vehicles will be connected. One interesting component is plug-in electric vehicles that can interact with smart grids. Communication and information are enablers, and learning to robustly and safely exploit them will be critical. To succeed, algorithms and planning tools that cover all aspects from basic driver support to fully autonomous vehicles, must be developed. We already have, for example, jet engines, as well as many marine vessels, connected via satellite to land-based service centres for monitoring or control.

Coupled to the information revolution are aspects of safety and integrity, i.e. cyber security. To design systems that are safe against computer viruses, burglary and terrorism that can harm individuals and systems, cyber security-based thinking must be incorporated in the design process and the awareness of engineers.

Ongoing engineering efforts are devoted to identifying and shaving losses, giving rise to advanced powertrains and, in the case of marine vessels, entire power systems that are more complex but support system optimization. Models provide a cornerstone for this work, since they encapsulate knowledge and establish the foundation for many decisions and designs. To move forward, models covering all domains of transportation systems, from low-level combustion chemistry and pollutant formation to high-level transport systems models, will be required. The challenge facing researchers in this domain is the diversity and complexity of the subsystems that make up vehicles, vessels and transport systems. There is a need for engineers to optimize these systems which operate at different (or even vastly different) time scales. Our community can contribute by developing tools to systematically analyse and optimize these intrinsically complex systems, and support the development of robust and accepted solutions.

Aerodynamic losses are the main source of energy wastage for vehicles (cars, trains) traveling at speeds higher than 50 km/h. According to existing ecological estimates, reducing these losses by 25% will decrease car pollution by more than 107 tons of CO2 per year. Today, vehicle designers have achieved near-optimal static

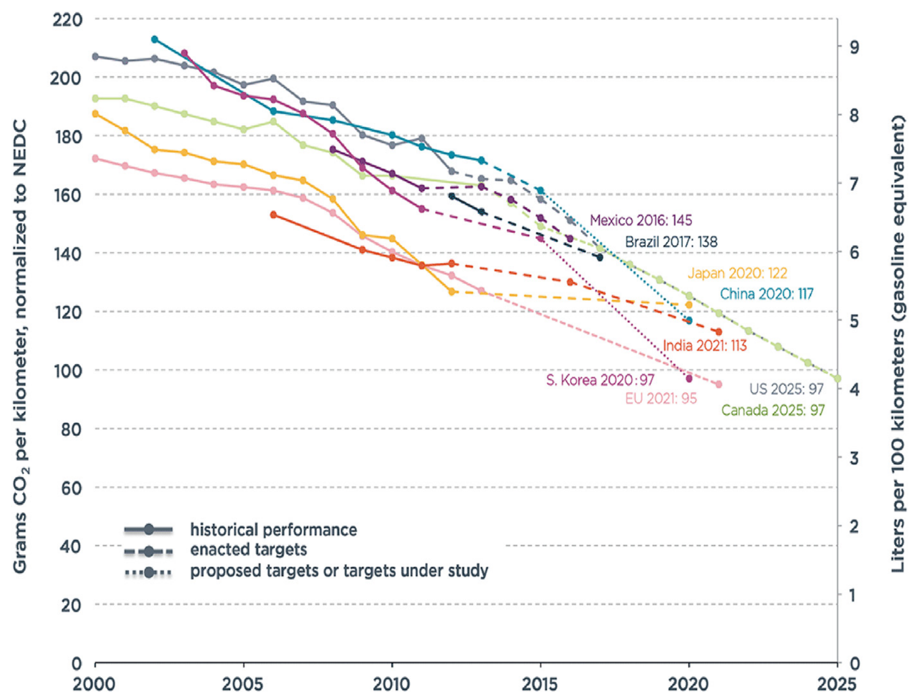

Fig. 2. Collection of historical data and future CAFE standards that in effect is a driving force behind the development of fuel-efficient vehicles, with lower CO2 footprint. Source: www.theicct.org.

solutions, and further improvements based on vehicle shape optimization requires numerous trials for minimal enhancement. In this context, active flow control strategies constitute the next challenge. Moreover, the same issues also concern aeronautics, where aerodynamic lift can be improved by an appropriate feedback control of the boundary layer (i.e. reducing the flow separation). This represents an interdisciplinary research effort joining micro- and nano-technologies (designing flow sensors and blowing actuators to be placed on the car surface or aircraft wings), flow mechanics (properly dimensioning and placing these sensors and actuators) and control (closing the loop in a safe manner). The main issue is that the underlying physics in aerodynamic studies (i.e. the Navier–Stokes equations) are known to be highly nonlinear and contain distributed parameter effects corresponding to the various natural diffusion phenomena which can be seen as time delays or transport equations, for instance. Various flow control strategies have been found efficient in drastically reducing flow separation. However, in general they do not lead to optimal solutions where separation is fully avoided for a minimal energy cost.

Efficient and clean transportation solutions need to be developed. We firmly believe that the control system society, with key competences in modelling, system analysis, control and optimization, will play an important role in the innovation and engineering of clean and efficient transportation solutions for the future.

3.2. Energy

The 21st century is witnessing huge paradigm shifts in key sectors that affect quality of life worldwide. The energy sector provides a very compelling example of these paradigm shifts. Motivated by huge concerns of sustainability, climate change, carbon emissions and aging infrastructures, several research initiatives are underway on the energy front. One such initiative is the smart grid, a transformative, global imperative for the energy challenge facing us in the 21st century. A smart grid is an end-to-end cyber-enabled electric power system, from fuel source to generation, transmission, distribution and end use that has the potential to enable integration of intermittent renewable energy sources and help decarbonize power systems, enable energy efficiency and operate resiliently against physical and cyber-attacks. Central to the realization of these goals is control that can gather any and all information that is available about the grid, facilitate the functioning of resilient transmission and distribution networks, shape any and all loads that are responsive, allocate generation to all generators, storage units, and electric vehicles, and enable the delivery of reliable and affordable power everywhere and at all times. The increased deployment of feedback and communication implies that loops are being closed where they have never been closed before, across multiple temporal and spatial scales, thereby creating a gold mine of opportunities for control. Control systems are needed to facilitate decision making at various timescales and line scales, from devices to systems, from generation sources to consumers, from planning over multi-year and year timescales to operating frequency and voltage control in seconds and milliseconds.

The paradigm shift associated with smart grids pertains to the fundamental way in which power is delivered to the end-user. The traditional approach consists of centralized generation, located far away from the load centres that are typically urban, with redundant long-range transmission paths and radial distribution. The smart grid concept seeks to drastically change this picture – along with conventional generation, distributed generation resources will dot the energy landscape, necessitating a judicious combination of traditional fossil-fuel-based, fixed and predictable generation with renewable, variable, intermittent, stochastic and uncertain generation based on renewable resources such as solar and wind power. This in turn introduces myriads of challenges at all levels of the grid. These challenges not only necessitate new tools in the traditional topics of generation, transmission, and distribution, but introduce a whole host of new concepts, tools and technologies.

An emerging concept that has been increasingly researched of late and shows significant potential is Demand Response – the notion of adjustable demand in response to grid conditions and incentives. The idea is to determine a desired profile for demand response to complement and supplement intermittencies and variations in renewable generation and ensure reliable and affordable power delivery to the end user. Determining these profiles requires assembling a range of flexible loads and storage devices at various locations that vary in timescale and authority. Solutions to these challenges significantly intersect with Systems & Control related problem formulations.

Ensuring the desired flexibility in demand brings economics into the picture, as flexibility in demand is often facilitated through economic incentives. Approaches denoted as *transactive control* – a mechanism through which system- and component-level decisions are made through economic contracts negotiated between the components of the system, in conjunction with or in lieu of traditional controls – will need to be examined. Dynamic market mechanisms will become increasingly important with the move towards real-time decision making, as the underlying forecast models of renewables and consumer and load control behaviour become more accurate. Consumer and load control behaviour with their distributed decision making must be addressed. Real-time and closed-loop demand response may result in coupling energy resources and markets at timescales leading to significant stability and robustness questions.

The above discussions indicate that control-centric challenges and opportunities abound in the context of a smart grid. The grid can be viewed as a massively networked large-scale Cyber-Physical System of Systems, requiring decentralized, distributed, hierarchical, hybrid and adaptive systems tools. For example, distributed, real-time closed-loop architectures are needed that accommodate uncertainties in renewable generation and match supply to demand by making use of ubiquitous real-time information and optimizing global objectives into coordinated local algorithms. Scalable algorithms that are decentralized and deployable at huge distributed scales need to be developed, supported by local decisions and global coordination. A redesign of currently existing architectures that includes primary, secondary, and tertiary layers is needed to integrate renewable energy as a dispatchable source while providing optimal alternatives to expensive ancillary services.

While transmission systems have always faced formidable challenges in terms of controlling swing oscillations in large-scale systems with few measurements, significant nonlinearities and uncertainties, the introduction of renewables into the picture further exacerbates the possibility of oscillation swings following major disturbances. New opportunities are present in the form of measurements using PMUs (phase measurement units), and actuators such as FACTS (Flexible Alternating Current Transmission System). These devices and control algorithms with distributed computation and communication need to be designed so that wind farms and solar farms can be optimally located and integrated to promote an efficient transmission system. Although HVDC (high voltage, direct current) transmission systems have been in use for many years, newer voltage-source converter (VSC) technology, generally referred to as VSC–HVDC, particularly in the new Modular Multilevel Converters (MMC), opens up a much wider array of applications where cables are needed in order to connect wind farms a substantial distance from shore, to connect non-synchronous zones using underground lines (European grid to UK or to Norway for example), and to span long distances using aerial lines (from 700 km on).

Smart distribution management systems need to employ advanced actuator technology including power-electronics devices that enhance controllability and power transfer capability and have the potential to prevent cascading failures. Given that more decision-making may need to occur at the distribution level, distribution analytics that manage communication, estimation, computation, protection, optimization and control should be investigated. Design of distributed generation (DG) clusters in terms of the type of sensors, communications and control architectures that can enable efficient and reliable power flow, as well as appropriate contractual structures that facilitate these goals need to be carried out. New topological complexities that may result from system changes due to micro-grid operations and "mesh" structure must be tackled. Protection systems with preventive control strategies using on-load tap changers and smart capacitors and switches need to be increasingly deployed to ensure satisfactory voltage/VAR(Volt-Amphere Reactive) control.

The power grid provides an essential service to the country's citizens. It is therefore imperative that control architecture designs should realize, distinguish and transition between a normal and emergent state, as well as launch the corresponding sequence of corrective, restorative and healing actions. Multi-layers of protection and cyber-physical security against not only natural anomalies and failures but also cyber-attacks have to be enabled. An increasingly popular concept of resilience, the ability of a system to perform in a stable manner even when pushed very close to its limits of operation, is a highly desirable property that needs to be addressed in conjunction with smart grid cyber-physical security.

Yet another aspect of the power grid that needs attention is the design of an efficient electricity market. An electricity market represents a system of entities that are involved in the trading of electricity, an important planning component. As electricity cannot be stored in large quantities at the current cost of energy storage, and any electricity that is produced must be consumed, the electricity market is responsible for ensuring transmission of electricity in a reliable and efficient manner. Emerging challenges in smart grids are due to the introduction of new actors into the market including renewable energy generators, storage providers, and as mentioned above, demand response-compatible consumers (possibly through an aggregator). This in turn necessitates the use of Systems & Control tools that allow these actors to efficiently exchange information and make decisions and enable power delivery that is reliable and affordable.

- Power systems are stabilised through very robust and conservative control strategies, heavily based on redundancy and natural inertia. Nevertheless, even in classical grids stability remains a difficult problem. Timescales range from tens of milliseconds for transient stability (keeping one power plant connected to the grid) to a few seconds when considering the synchronisation of the full system. With the high penetration of renewables and the consequent increase of power electronics, such timescales will be much faster, and at the same time the natural inertia will be greatly reduced. The future grid will be much harder to stabilise and will impose the use of completely new control schemes.
- Many current practices in electricity markets may be viewed as suboptimal solutions to a stochastic, multi-stage, dynamic programming problem. With increasing penetration of renewables and the corresponding increasing intermittency and uncertainty in the underlying market operations, the central question is the realization of market mechanisms that can provide optimal solutions despite strongly stochastic and temporal variations. The challenge is maximizing efficiency while guaranteeing reliability even in the presence of possible load loss and varying generation without falling back on conservative decisions.
- Fast reserves, used to enable frequency regulation, are typically procured in hour-ahead or day-ahead markets. Growing penetration of renewables, however, makes such a practice highly inefficient and expensive. The question is whether judicious use of Demand Response, including flexible building loads, batteries in electric vehicles and other storage solutions can mitigate

these drawbacks. New dynamic market mechanisms need to be designed to provide efficient market price signals and maintain energy balance in real-time despite intermittent and uncertain renewable generation.

- Given the significant impact that increased uncertainties stemming from renewables can have on market transactions, accurate forecast modelling is a crucial ingredient in determining resource dispatch. With the trend in more accurate forecast models for entities such as the weather and demand over decreasing horizons, market models at multiple timescales that incorporate varying forecast models and their modelling errors need to be developed.
- Innovations in electricity markets entail additional, frequent and judicious information exchange between various stakeholders in the grid. This in turn introduces new challenges in the cyber-physical domain pertaining to computational, communication and information systems. New safety-critical components may be necessary in these markets, thereby raising issues of bandwidth, reliability and cyber-security.

Yet another important component of smart grids is power electronics. As more renewables penetrate the power grid, sensors and actuators are required to integrate the resulting energy into the grid in an efficient manner, which brings one to the realm of advanced power electronics. Active and reactive power injection need to be monitored, modulated and delivered at the right time and right place. The use of smart inverters, synchroconverters, families of microgrids, energy clusters, mixed AC-DC solutions and control algorithms comprise important research investigations that the controls community can launch and complete successfully.

In summary, the smart grid is an ideal poster child for controls, with a need and opportunity for closing several loops at various levels, locations and time-scales, to deliver reliable, affordable and sustainable power to all consumers. It could be argued that by introducing billions of active endpoints through deployment of sensors, actuators, and communication devices, at generators, transmission lines, substations, renewable energy sites, distribution feeders, microgrids, buildings, homes, smart appliances, power electronic devices, storage units, and electric vehicles, and appropriate closed-loop strategies, one can realize the ultimate smart grid vision.

3.3. Water

Despite earth having an ample amount of fresh water to support life, the 2030 Water Resources Group in its report “Charting Our Water Future: Economic Frameworks to Inform Decision-Making,” presents the sobering message that the world is managing itself towards economic water scarcity. Similarly, in the United Nation World Water Development Reports (The United Nations World Water Development Reports 2015), the document “Water for a Sustainable World” starts with describing our world as it can be, where every living organism equitably shares amply available high quality water resources. But in the other report of The United Nations World Water Development Reports (2015), “Facing the challenges: case studies and indicators,” it is painted a rather grim picture of the present fresh water reality.

By and large there is a strong spatial correlation between water stress and population density. Areas under water stress include the Western United States, Northern Mexico, most of India, the North China Plain, countries around the Mediterranean, Middle Asia, the Middle East, Eastern Australia (Murray–Darling basin), as well as the narrow strip west of the Andes. Climate change adds complexity, as extreme weather patterns strongly affect where and when water stress occurs. For example, countries such as Brazil, the UK and the Southeast Asia region experience non-trivial water stress issues. Climate change points to developing more resilient engineering solutions.

The main driving force behind water stress is indeed population growth accelerated by the present universal expectation of rising standards of living. The outlook for water sustainability is therefore grim, as the world population is expected to grow from the present 7 billion to 8.5 billion by 2030, and may continue to grow till 2100 (the UN Medium prediction forecasts 11 billion people on earth by 2100). The World Wildlife Fund captures the human dimension of the present water crisis (World Wildlife Fund 2017): “As a result, some 1.1 billion people lack adequate access to water, and a total of 2.7 billion find water scarce for at least a month of the year.” The physical dimension is appreciated by realising that 95% of the world’s fresh water resources are under significant stress (i.e., river basins experience over-extraction, and entire regions suffer from lowering water tables). Of all water extracted for human use, an average 70% goes towards food production (unchanged since the first World Water Report in 2003). Under a steady business-as-usual scenario, the world will require nearly twice as much water by 2050 compared to today. This estimate is critically influenced by how much meat the average human diet will contain. Presently, the World Bank estimates that the world extracts around 4500 cubic km of water from the hydro cycle every year. Without creating significant new supply, the world’s sustainable renewable fresh water extraction is estimated at 5000 cubic km per year. The 2030 Water Resources Group estimates the economically available and renewable fresh water extraction at only 4200 cubic km; therefore, the world has already entered into a non-sustainable water usage pattern. Clearly, doubling water use, as implied by the steady-state scenario, is not feasible.

The obvious engineering answers are to create more supply through desalination and improved recycling, and/or to be more efficient in water use. The former solutions require significant investment in new infrastructure and are typically energy-intensive. In contrast, the sun does all the heavy lifting in the hydro cycle, providing free transport and free cleaning. Improving water use efficiency requires a significant investment in better water management. It is a very reasonable approach because the present water use efficiency is low – typically less than 50%. It is our thesis that much can be done to improve the water efficiency using Systems & Control engineering ideas. The simple mantra “measure – model – manage” can be used to great effect.

Other approaches, not discussed here, include food engineering, i.e. genetically modifying crops to use less water; and intensifying food production per unit of water and unit of energy through new precision farming techniques. Precision farming is also a very fertile area for applying control engineering ideas.

Measure – Model – Manage water networks

On a world scale, measuring water (flow, volume, pressure, quality) at the natural scale of water catchment basements and at time intervals that support decision making at all levels, is not trivial. Presently, from a systems engineering point of view, the water sector is information-poor. Even today there are jurisdictions where water use measurements are lacking, in line with the opinion that water is or should be a free good. Furthermore, the time and spatial scales on which water needs to be managed are enormous. On the one hand, catchment basements are continent size, whereas a single end user may require as little as 50 litres of clean water in a day. Similarly, timescales vary from a century for climate change, infrastructure developments and ground water movements, to water demand variations that take place on a timescale of days, and water flow dynamics that may exhibit timescales in the order of seconds. Keeping track of fresh water on a global scale is an immense task, and much progress has been made since the call to action in the first World Water Report 2003. Traditionally much work has been done, and is done, to build models that make sense of very sparse data sets with the purpose of planning water infrastructure and water availability over long periods of time (100 years). Risk management strategies are key here. Present climate change conditions are causing real havoc with the uncertainty estimates. What actually constitutes a "once in a hundred-year event" is in all probability significantly altered due to climate change. Indeed, climate change is characterised by greater uncertainty in weather, with indications that extreme weather events will be more likely. But how much more likely, remains elusive because we lack data. More resilient infrastructure, or infrastructure that can respond to extremes, is called for.

As indicated, irrigation water accounts for nearly 70% of all water extractions, and 70% of this, or nearly half of all the world's water is spent in broad acreage irrigation using gravity-based water transport. Such systems have been in operation since the Sumerians settled Mesopotamia, about 5000 years ago. Efficiency (amount of water used for irrigation over amount of water extracted from the environment) in most of these systems is low: less than 50% of water released from a dam reaches its intended destination, with some modern irrigation installations performing better and others, far worse. More water than what is consumed for all non-irrigation purposes combined is somehow "wasted" in the irrigation process. Do observe that "wasted" is a difficult word in this context, as the unused water is returned to the environment. How useful or wasteful this is, depends on many factors. The real problem here is that the extracted water was not used for its intended purpose, i.e. it was not managed.

Overall water use efficiency in irrigation can be spit in two components: from the dam to the farm gate via the bulk distribution channels and from the farm gate to the plant roots. Both efficiency factors are typically around 70%.

A first control objective is to improve the conveyance efficiency of irrigation systems from the dam to the farm gate. This can be achieved through better understanding channel dynamics, transients and the storage capacity of the irrigation channels themselves. Intuitively, it is feasible to implement much tighter water flow schedules using automated or controlled systems based on real-time in-channel flow measurements and automated flow actuation as compared to the current practice of manually-operated, under-measured and under-actuated systems. Automated conveyance can reach a very high efficiency, typically leaving only evaporation and seepage as uncontrollable losses.

The on-farm water efficiency can be greatly improved by timing irrigation events in response to crop needs. Proper timing of the appropriate quantity of irrigation water depends on crop behaviour, and this varies over the growing season and with climatic/weather-affected evapotranspiration conditions. Periodic irrigations, and most manually scheduled irrigations lead to over irrigation, as these regimes are not well-aligned with crop need. Delivering water nearly "on-demand" and according to crop needs, generally reduces the amount of water required (less risk for the farmer, less need for over-irrigation) and invariably improves water (and crop) productivity significantly. These productivity gains must pay for the modernisation.

Improving on-farm water-use efficiency brings with it significant side benefits: less runoff, implies less pollution of rivers, estuaries and bay areas; the better the water efficiency, the less fertiliser is necessary, which leads to a significant cost saving; and improved water efficiency reduces soil degradation (over irrigation creates soil salinity problems).

Two aspects of the irrigation water service, precision timing and conveyance efficiency, can be pursued as important control objectives in an automated network of water distribution channels. A further quality of service objective that control and automation must pursue is maintaining precise water levels at given points along the channels. Indeed, in an open-channel gravity-fed water system, the potential energy, the water level, determines the amount of land or length of channel that can be supplied with water.

Measuring and controlling water flow in open channels, and even in pipelines, is costly. At present, most irrigation systems have very little measurement infrastructure, and few are serviced by remote monitoring. Most of our work, and that of our collaborators, has utilised measurement and actuator infrastructure developed in a collaboration between the University of Melbourne and Rubicon Water Pty Ltd over the last 15 or so years. It consists of in-channel water flow actuators equipped with co-located water flow and water level measurements. These are all linked through a wireless internet over which measurements are communicated, control objectives are updated and general system operational and maintenance information is exchanged. Each local flow actuator can act as a stand-alone control agent or work collaboratively with its neighbours to implement an overarching management strategy prescribed in a hierarchical manner across the entire system, where the collaboration is software defined by the communication graph. To illustrate, the Goulburn-Murray Water irrigation district (nearly 68,000 square km of land), conveys water in about 6000 km of major irrigation canals to service 15,000 on-farm water outlets. When it is fully automated in late 2018, it will contain well over 10,000 in-channel actuators. The entire control system will track approximately one million variables. The actuators and sensors are designed to act every 10 min, over a period of more than 25 years, but allow for a more intensive duty cycle when necessary.

Using the flow and water level measurements along the channels, appropriately augmented with GPS-derived channel information, several mathematical models can be constructed for the overall system: in support of local day-to-day channel control objectives (precise water orders, with precision timing minimal losses at the correct water levels); or to enable a water market (buying and selling, holding water reserves in dams and large channel segments); or to support management objectives such as infrastructure planning and maintenance. Decentralised and distributed control strategies that approximate model predictive control are well suited to the task.

3.4. Healthcare

Understanding disease mechanisms and developing successful treatments are enormous societal challenges where Systems & Control can play a critical role. This Section illustrates few pressing medical issues with current difficulties and complexities.

Unsolved diseases in an aging world. Improvements in sanitation, water quality and environmental control, combined with advances in medical science, mean that we are all living longer. Statistics on longevity show that since the late 19th century, global life expectancy has increased by approximately 0.28 of a year annually. While this rate of increase in life expectancy shows no sign of slowing, the point at which diseases of age (such as: Alzheimer's and Parkinson's disease) become significant, has not seen a corresponding increase. Consequently, ever-growing numbers of our elderly are falling victim to neurodegeneration. For example, the risk of dementia exceeds 30% for a person over 90 years of age, and other neurodegenerative disorders show a similar trend. This, in itself, would not be a critical problem if effective treatments were available. Unfortunately, this is not the case; preventative treatments or cures do not yet exist for most neurodegenerative conditions. Moreover, the complexity of these conditions is such that progress towards cures and treatments is uncertain and glacially slow.

The drugs don't work anymore. The difficulties in understanding complex conditions, like diseases of age, mean that drug in-

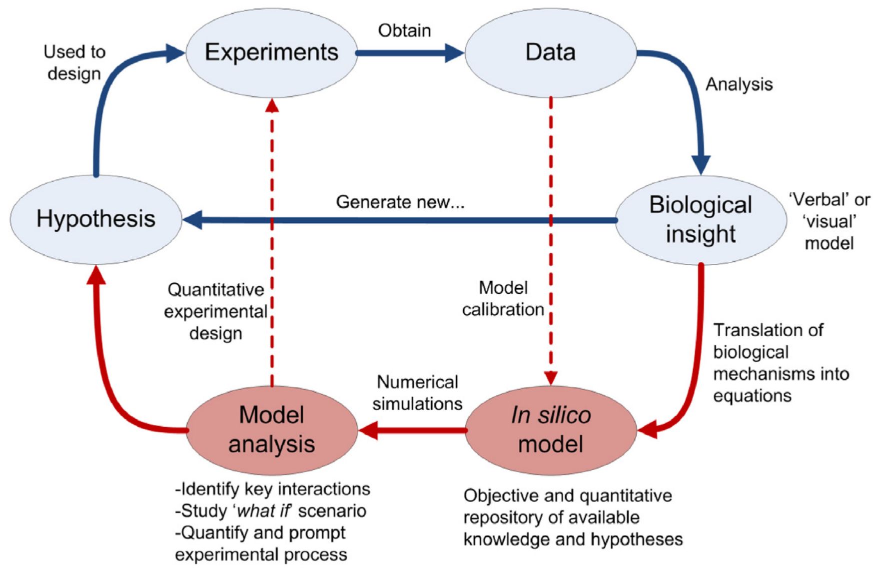

Fig. 3. Conventional disease research model (upper part of figure) and the alternative Control and Systems science disease analysis paradigm (lower part of figure).

dustries are losing the ability to develop effective treatments for unsolved diseases while remaining competitive. Over the 50 years prior to 2010, the cost of drug development increased exponentially, while the rate at which new drugs appeared was static. Productivity and profitability of drugs companies are low, and there is talk in the industry of *a pharmaceutical 'ice age' and potential extinction*. The problem is that all the tractable diseases have been resolved. What remains are diseases too complex for the traditional disease research and drug development model; this model is failing. With the current research methodology, success depends crucially on the individual brilliance (or luck) of the researcher, combined with experimental iterations that are costly, time consuming and frequently inconclusive. The faltering nature of drug development, shrinking product pipelines and inability to solve diseases of the aging brain reveal a failing business model – a situation that calls for disruptive innovation – a new paradigm for disease research.

Disease modelling and analysis – a new frontier for control and systems science. The traditional disease research/drug development model is illustrated in the upper part of Fig. 3. In this model, data from experiments with disease treatments are studied by experts, who use their personal know-how to generate hypotheses for a further experiment. Experiments would initially comprise of laboratory tests ('in-vitro'). If these experiments are promising, then a new cycle of experiments begins with animals ('in-vivo') and so on – culminating with experiments with human subjects. Each of these stages can take many years, and may fail at any stage due to flawed hypotheses, erroneous subjective judgements, poor data, or during the translation of results from 'in-vitro', to 'in-vivo' animal trials, then human trials, and most serious of all, in commercial use on the general population.

The lower half of Fig. 3 is the Systems & Control science disease analysis paradigm. In the right lower block, knowledge of the biological mechanism is translated into equations to form a mathematical model of the disease processes (**'in-silico' disease model** to biologists). This model is calibrated using experimental data and information from biological databases. The model is validated by simulation, modified and improved until it corresponds with known biology and physiology. In this way, the mathematical model becomes the objective repository of quantitative and structural knowledge of the disease. The validated model is then used to provide a **disease model analysis** tool in which key (dynamic) interactions and disease mechanisms can be identified by analysis of the model's internal state behaviour. This itself is a huge advance on tradition because only rarely can the dynamical changes in disease states be measured 'in vivo' with accuracy and sufficient bandwidth. New theories for disease mechanisms from the model analysis phase (left lower block) are first thoroughly tested and refined in 'what if' simulations, before using then to develop focussed testable hypothesis that can feed into the traditional 'in-vitro', 'in-vivo' experimental cycle for verification.

Analysing Parkinson's disease – a case study in control and systems analysis. Parkinson's is the second most widespread neurodegenerative disease. There are no cures or preventative strategies, and the available treatments address only the visible symptoms of tremor and movement disorders. It is a long-term condition that progresses dynamically at a variable rate through the nervous system, eventually affecting the entire brain. The tremors are the first visible indication of Parkinson's and occur sometime (possibly years) after its inception. There are numerous risk factors for Parkinson's, any combination of which can be causal and responsible for its manifestation and progression. A big unsolved problem in Parkinson's is 'what is the pathogenic mechanism?' Parkinson's resemblance to a multi-factorial failure in a physical system makes it an obvious candidate for a Systems & Control approach to answering this question. The disease can be modelled and analysed 'in-silico' at three levels (Fig. 4). Level 1: All Parkinson's risk factors damage the brain energy availability such that a mathematical model of brain energy metabolism forms a unifying framework for systematic analysis of the risk factor interrelation**.** Level 2: Conceptual models of the molecular players allow a mathematical model of the key pathways for Parkinson pathogenesis. Level 3: Simulation of Level 1 and Level 2 models motivates abstraction of the pathway model in the form of a feedback motif for the pathogenic mechanism.

The family of models in Fig. 4 provides a systematic basis for a more efficient, faster and focussed disease research. The models encode existing information, and analysis generates new knowledge. In particular, analysis of the core feedback mechanism relating the key molecular players, models the pathogenic mechanism as a bifurcation process in the nonlinear dynamical of a neurochemical interaction.

The Parkinson's pathogenesis is an example from a growing group of projects that address 'in-silico' modelling and analysis of

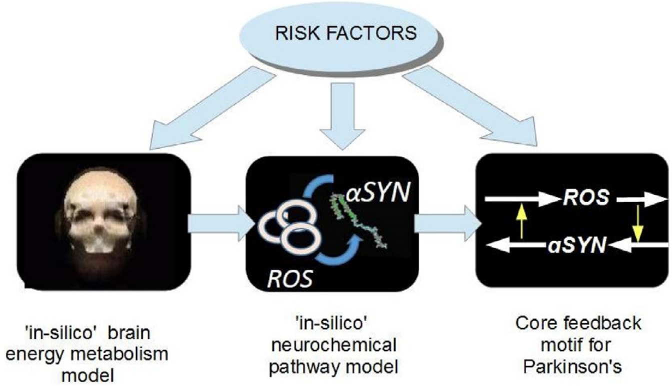

Fig. 4. Working through three levels of 'in-silico' modelling for metabolism and disease mechanisms, reveals how pathogenesis is modelled by feedback interaction between reactive oxygen species (ROS) and the protein alphasynuclein (αSYN).

complex diseases. There is a gap throughout this rich and developing area, where Systems & Control sciences have a unique opportunity (Wellstead, 2017) to unravel the complex dynamics of disease structures and mechanisms. Parkinson's is an important example of how disease dynamics can be modelled and analysed, but there are many similar diseases of age which beg attention – most notably Alzheimer's disease. The potential social, human and economic rewards from control theory and dynamical systems analysis applied to disease are enormous – the Systems & Control science community owes it to themselves to participate.

3.5. Manufacturing

Integrated large production complexes in the chemical and petrochemical industries are major consumers of energy and raw materials, as well as major sources of employment and income. They produce virtually all the raw materials for convenience products in modern industrial society. The ecological and economic viability of their production depends crucially on the careful management of the ensemble of different units which in many cases are simultaneously producers and consumers of intermediates and carriers of energy (see Fig. 5). These sites host large numbers of autonomously operated production plants, often owned by different companies, with complex energy and material stream interconnections to ensure operational excellence and competitiveness of the production. The plants belong to competing value chains within one company or among different owners. Complex networks of carriers of energy and of various chemicals are operated to optimize the use of energy, materials and intermediates and byproducts.

Production flexibility is limited by many different constraints on individual units which must not be violated in order to prevent, e.g., accelerated equipment degradation or plant trips. Each unit operates most efficiently, in terms of economics and energy and resource consumption, under specific conditions which are often incompatible with the global state of the production system due to the interconnections and limited resources. The primary goal of site-wide management is to achieve an optimal global performance. The degree of freedom to achieve this goal lies in the ability to vary production intensity to compensate for utilities' changing availability and market prices. Essentially each plant operates autonomously, as an independent agent that tries to reach its production objectives as part of the value chain. Therefore, this is an area with an enormous potential for the use of distributed management and control approaches leading to better coordination and hence better economic and ecological performance.

In the process industries, one finds a tight integration of physical plants with computer-based management tools such as Enterprise Resource Planning (ERP), Manufacturing Execution System (MES), Supervision, Control & Data Acquisition (SCADA), Distributed Control System (DCS), Human-Machine Interfaces (HMIs) and Programmable Logic Controllers (PLCs) that constitute a hierarchical system, the so-called automation pyramid. These systems interact with the site on different levels and timescales and, thus, employ conceptually different viewpoints on the site and its parts. The models used by these systems differ in temporal granularity (some models are partially or completely steady state), degree of abstraction of the occurring phenomena (static models are commonly used for simplification of some higher-level tasks despite the dynamic nature of the modelled processes), and reliability as the model parameters might be uncertain or changing with equipment aging or replacement and dynamic reconfiguration of the site.

The process industries have for many years pioneered the application of advanced control strategies. Model Predictive Control (MPC) penetrated as an industrial standard for control of large chemical plants (reactors, distillation units) mostly in petrochemical industry. This is due to MPC's ability to effectively handle operation of multi-input, multi-output, constrained dynamic systems that represent the chemical production site core. Industries report the major success points of MPC lie in increased throughput, improvement of process stability, reduction of energy consumption, increased yield of more valuable products, reduction of quality giveaway, reduction of down times and better use of raw materials.

Site management in the process industry is a complex optimization problem that spans multiple timescales and layers of decision-making. It starts in the upper layer with the planning of production and supply chain optimization that allocate levels of production to respond to market conditions and satisfy customer contracts. This information is passed to the lower layer where production scheduling takes place and where RTO (real-time optimization) adjusts the set points for operation of specific plants. MPC or standard control technique is then used for meeting the demand as set by scheduling and RTO by passing the control signals to the actuators (pumps, valves, boilers, etc.). The models used along this hierarchy vary largely in the depth of abstraction of the physical plants, the employed formalism (as they are used for different purposes) and nature (first principles models are preferred for modelling of physical plants but are difficult to employ for modelling at higher levels), execution timescales, etc. For example, scheduling batch production exploits static models which do not incorporate any information about the uncertainties present at the plant (status of equipment, quality of raw materials) or the experience and actions of plant personnel. This affects the feasibility of the decisions made at the upper level, as many effects of dynamics and uncertainties are not considered, and restricts a penetration of automated solutions for fully-autonomous production planning and execution.

In the discrete-part manufacturing industry (Fig. 6), automatic control systems facilitate production of airplanes, automobiles and many consumer goods. Raw materials and sub-components enter a factory, are transferred among processing stations and transformed into finished products. Traditionally, these controllers are programmed at a low level, in a distributed fashion, without any global view of the entire system. This leads to sub-optimal production strategies that can be vulnerable to disruptions and even to cyber-attacks. The collection of large amounts of plant floor data into cloud-based systems opens the promise of not just system-level visibility but also enterprise-level control and optimization. Improved scheduling and control techniques can lead to lower prices for consumers while providing higher profits to industries.

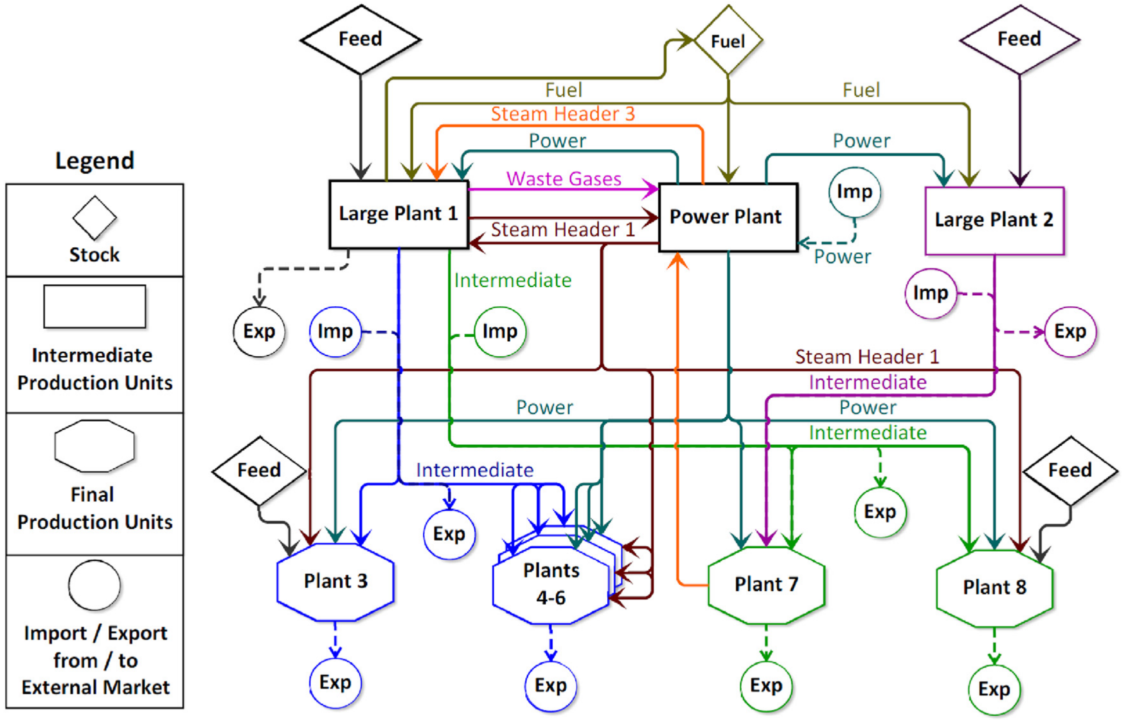

Fig. 5. Schematic representation of an integrated chemical production complex.

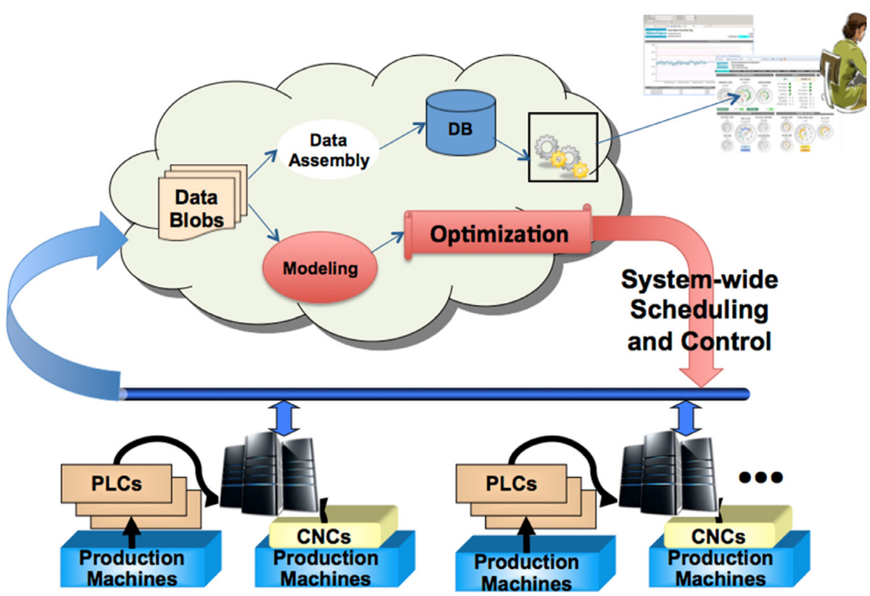

Fig. 6. Data from the plant floor - either from programmable logic controllers (PLCs) or computer numerical controllers (CNCs) - is pushed to the cloud, where engineers use dashboards to monitor production, and automatic modelling and optimization methods are used for system-wide scheduling and control to improve the overall enterprise-wide operations.

Significant opportunities for automatic control exist at every level of these complex manufacturing and production processes. At the lowest levels, small gains from applying modern control approaches to replace classical PID controllers can result in big wins for large operations. Better models of the combined cyber and physical system components can improve monitoring and control strategies. Unscheduled downtime is one of the largest avoidable costs in any production operation, so better diagnostics and optimized predictive maintenance can pay large dividends. The greatest challenges for control technologies require a global view of a complex operation, from supply chain through production operations and even out to the products' consumers or end users. This global view will be composed of myriad dynamic models seamlessly stitched together, connecting through time and space. The appropriate interfaces between individual models and controllers need to be defined, and the mathematical techniques describing their connections and evolution must be developed. Once these challenges have been overcome, the opportunity for large-scale op-

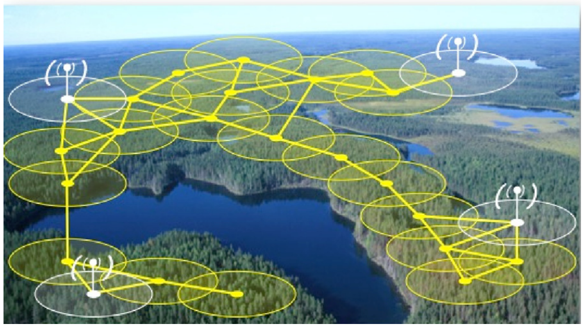

Fig. 7. Example of networked control systems.

timization with uncertainty will allow for improved productivity and efficiency of production operations.

4. Key research and innovation challenges

In this section, the main research and innovation challenges have been structured into seven domains, each of them having their own models and/or data and methodologies (horizontal lines of Fig. 1 matrix).

4.1. Distributed networked control systems

Given the importance of distributed sensing and control for a wealth of application domains, such as autonomous vehicles, traffic control, manufacturing plants, smart cities and healthcare, the effects of practical limitations of the interconnections such as bandwidth, noise, delay and distortion, must be characterized and understood. In the traditional view of control, communication among controlling and controlled entities occurs instantaneously with signal degradation caused only by stochastic noise. In reality, delays and signal degradation are common, especially when the distance between elements of a communication system is considerable. Recently, three new approaches to the design of distributed networked control systems have been pursued. The first assumes that the communication protocol is given and aims at designing the control algorithm so that the networked closed-loop is robust to the specific protocol-induced non-idealities (e.g. delays, signal degradation, packet drops, etc.). The second, conversely, assumes that the control algorithm is given and aims at designing and configuring the communication protocol so that the networked closed-loop guarantees a specified control performance. The two approaches above are based on the separation-of-concerns principle and may be combined in the third approach of co-designing the communication protocol and the control algorithm: this clearly improves optimality of the networked closed-loop performance at the price of an increase in the dimension of the design space.

Networked control (Fig. 7) provides an interesting example of a domain where interactions between research communities can reap large benefits. Wireless as well as mixed communication, where wireless connectivity is combined with wired communication, pose significant challenges, such as mitigating the effects of the unreliability and stochastic time-varying characteristics of wireless connections on control performance. In this context, it is important to consider issues such as channel characterization, to make sure that channel communication quality meets the demands of the control algorithm and power consumption, to maximize the life of wireless nodes. In general, these issues are currently considered using heuristics and ad-hoc methods.

How to design control for networked systems over wireless communication channels?

Much of the current research on networked control systems focuses on fixed structures for communication and interaction. But networks of mobile vehicles or smart energy grids with time-varying sets of producers and consumers, give rise to system structures which change over time. Control methods for networked systems must cope with communication links that may be established for only a limited period of time. For example, in a connected vehicle scenario, delayed information on the routes taken by drivers can actually create an instability that is difficult to correct, whereby drivers acting on information about congestion decide to take the same alternative routes at the same time. Ensuring stability or performance for systems in which communication or coupling between subsystems is restricted in this way, goes well beyond present investigations of time delays or (single) packet-losses in communication. Combining control techniques and improved communication can indeed prevent this kind of instability. Another important aspect is that distributed systems must be able to continue smooth operations when a few nodes are removed or added. This result can be obtained only if algorithms are devised with the ability to self-reconfigure. These topics, which are known and well-studied in (static) distributed computing environments, emerge in a much more complex fashion in distributed control systems, due to the dynamic phenomena that characterize all physical systems particularly feedback-controlled systems.

How to design control algorithms and network architectures for secure operations resilient to failures and malicious intrusions?

Among the advantages of networked control systems, one of the most notable is the resilience to failures that is made possible by the redundancy of communication nodes and links. With this positive characteristic, though, comes several challenges that require additional analysis and understanding to enable reliable application of networked control systems, e.g. in safety-critical systems. While a failure of a single node can typically be better tolerated in a networked system than in a centralized system, the robustness with respect to domino effects should be formally assessed. Misbehaviours of agents can result from accidental faults or malicious intrusions of potential adversaries: the latter are generally more difficult to detect since the attacker may apply stealthy strategies exploiting some knowledge of network configuration and plant dynamics (e.g. “reply” and “zero-dynamics” attacks). As witnessed in recent real-life events (e.g. the Stuxnet attack), cyber-terrorist intrusions can have devastating effects on infrastructures such as power distribution networks, and security is indeed a critical aspect to consider in networked control systems where (wireless) nodes are more vulnerable to attacks from potential malicious adversaries that can propagate false information or dangerous control commands throughout networks. It is therefore crucial that a networked control system be endowed with the capability to monitor the correct functioning of its nodes, detect and isolate failures and malicious attacks, and reconfigure the system to mitigate their effect on the performance of the closed-loop system. Currently, resilient control algorithms and architectures are being developed to cope with attacks and networking non-idealities and vulnerabilities that have been ignored in existing implementations for too long.

How to design control for networked systems with structures that change over time and with communication links established in a limited period of time?

In summary, co-design of control algorithms and networking protocols for monitoring and re-configuration of networked control systems is of paramount importance in developing efficient, reliable, predictable, safe and secure control architectures for large-scale networked industrial automation. Despite the impressive opportunities offered by newly developed (wireless) communication technologies for control purposes (lower configuration, commissioning and maintenance costs; easier installation; broader sensing/actuation capabilities; compositionality; runtime adaptation and reconfiguration), tough scientific challenges arise:

- **Complexity:** system designers and programmers need suitable abstractions to resolve the inherent complexities arising from the interaction of wireless devices, communication protocols and complex dynamical systems. To this aim, theoretical results developed for classical modelling frameworks in Systems & Control theory need to be extended to incorporate established modelling frameworks in telecommunications and computer science. The outcome of this integration, as evidenced by the recent and rich scientific literature on hybrid systems theory, generally leads to mathematical frameworks where analysis of and design for even basic properties (e.g. stability) is computationally intractable. Efficient methods for model-reduction with guaranteed precision are critically needed and are currently the subject of intensive research both in the Systems & Control and computer science scientific communities.
- **Reliability:** development of effective and computationally practical methods to guarantee robust and predictable behaviour despite (wireless) networking non-idealities is essential. To this aim, classical results on robust and fault-tolerant control must be specialized to dynamical models that can tightly approximate specific protocol-induced non-idealities. Moreover, robustness metrics exploited in Systems & Control theory (e.g. $\mathcal{H}_\infty$ norm) need to be integrated with reliability metrics in communication theory (e.g. SNR, outage probability).
- **Security:** it is evident that (wireless) networked technology is intrinsically vulnerable, and security mechanisms for control loops are needed. Such security methods have been developed separately by the Systems & Control and telecommunications scientific communities. The co-design approach to integrating networking and control gives birth to a new scenario, where the exploitation of *physical-layer* methods (e.g. classical Fault Detection and Isolation methods, where knowledge of the physical systems dynamics is leveraged to detect anomalies) and *cyber-layer* methods (e.g. Intrusion Detection Systems and cryptography) in a cooperative fashion has potential to tremendously improve the level of security of networked control systems.

Developing new algorithms, protocols and procedures to enable next-generation networked control systems to tackle the challenges described above, can only be achieved through tight collaboration between scientists in Systems & Control theory, telecommunications and information theory, and computer science.

4.2. Data-driven modelling, machine learning and control

4.2.1. Data-driven dynamic modelling in a model–intensive future

To understand, predict and control the environment, scientists and engineers intensively utilise mathematical models as this allows for the simulation of the world on computers. Model-based design and optimization is the dominant paradigm in the systematic design, operation and maintenance of complex engineering systems. This is reflected in a wide range of technology domains, from manufacturing and automotive systems to interacting robotics, from smart energy grids to personalized health systems. Future engineering developments require addressing dynamic systems that

- are progressively more complex, interactive and distributed in nature,
- operate, to a large extent, autonomously,
- accommodate changing environments and objectives, and
- have "learning" abilities,
- thereby maintaining a verifiable high performance through active actuation and control.

In all these future developments, models will continue to serve as a basis of accumulated knowledge to facilitate optimal design and operational strategies. While construction of underlying dynamic models relies to some extent on structural, physical and other first-principles relations, it will need to be complemented by data-driven approaches to estimate dynamic behaviour and changing characteristics of the environment. Specific examples where data-driven modelling may be needed include:

- On-line accuracy and validity assessment of the model in relation to its intended use (goal-oriented models). Systems must be able to assess the validity of the models they use. This validity assessment must be an integral part of model development.
- Adaptation of the model to account for changing dynamics (parameters) and topology/interconnection structures, and characterization of the model environment (disturbances);
- Active learning, by probing the system/environment to generate sensor information that is suitable for model adaptation and to satisfy demands for accuracy, autonomy and robustness.

These three requirements call for a paramount role for data-driven modelling, which must be integrated in virtually all future complex engineering systems.

When applied to the modelling and control of dynamic systems, data-driven modelling is known as system identification. System identification is a well-established field and discipline, for which fundamental principles were developed starting in the 1960s and 1970s. It has also delivered an established framework for handling large scale systems which includes standard tools for model estimation, model validation, accuracy analysis and model error and uncertainty quantification, as well as experiment design in different structural configurations in open and closed loop operation.

In order to maintain verifiable high performance, future engineering systems will need to be equipped with on-line capabilities for active model learning and adaptation, and for model accuracy assessment.

The new requirements for complex engineering systems deliver pressing challenges for the field of system identification to address.

- **The curse of complexity:** Complex systems can hardly be identified in all their parts from manageable data records. Attempting a full description would in fact inevitably result in highly uncertain and non-robust models. On the other hand, a model should not be better than required for its intended use; users prefer simple models that perform adequately to overly complex models that describe systems in excessive detail. For example, lowering the desired bit error rate of a receiver in a communication system requires a more accurate channel model. Developing task-oriented methods to better capitalize on the available information is a major challenge presently facing the field of system identification.
- **Error quantification:** One of the stiffest challenges in system identification is verifying that the model captures the system properties that are relevant to the problem under consideration. While new techniques are emerging to steer models to better describe impacts to the final performance, the problem of error quantification by and large remains an open issue.
- **Structural constraints and variations:** Many systems are highly structured and have specific behaviour properties (e.g. monotonicity). These characteristics should be respected by the model which leads to including specific constraints in the identification process. However, this is not an easy task. For example, effective identification in structured dynamic networks is, to a large extent, still an open issue. It is natural to decompose many systems into hierarchical models with different levels of abstraction, where the top level often represents some emerging behaviour of the system. For example, biological systems have levels ranging from molecular processes to functional behaviours. A very interesting problem is how to ensure that the emerging behaviour is captured by the model beyond the low-level characteristics. There is also a wide range of systems where structure varies in time. For example, internet topology changes continuously as routers and users are added and removed. These changes must be captured when they appear, by responsive identification methods.
- **Plug-and-play models:** Object-oriented modelling is a very powerful paradigm for physical modelling. In this paradigm, libraries can be constructed using components that are easy to connect, thus speeding up model building dramatically. Identifying large-scale systems is also eased by a modular approach that can also capture the interaction between connected subsystems. Here, an important issue is assessing how model errors in individual modules propagate through the entire model. In fact, examples can be constructed where seemingly innocent errors in individual components accrue to unacceptable levels when the errors are correlated.
- **The cost of modelling:** Reportedly, modelling costs represents the major burden in most advanced engineering projects, in some cases exceeding 50% of the total development cost. The most obvious cost is the number of man-hours of expert engineers. System identification is, at the present stage, very far from being automatic: numerous knobs must be tuned and decisions made during the identification procedure which require expertise and human intervention. Significant efforts are underway to develop robust procedures and algorithms that are entirely data-driven to guide non-expert users to reliable solutions. A second important cost is that of experimentation. This cost depends upon the length of the experiment (i.e. the amount of data required to obtain quality estimates) but also on the performance degradation that often results during experimentation when – as is often the case – the experiment requires applying excitation signals that degrade the quality of the product. The full analysis and exploitation of these concepts is one of the major challenges facing the system identification community.

From a research perspective, these challenges are not tied to a particular application/technology, therefore, addressing them with a generic systems approach is particularly attractive.

4.2.2. Machine learning and systems & control

A major opportunity for progress in Systems & Control derives from exciting developments in the field of machine learning. Historically, there have been close connections between machine learning, artificial intelligence and control systems. Goals and visions of learning and adaptive control are closely related. For example, many researchers have worked on utilising artificial neural networks for adaptive control. The backpropagation algorithm for training multi-layer neural networks traces its origins to some techniques in linear systems theory. Control systems are critical for creating intelligent, autonomous machines such as self-driving cars and autonomous robots.

Recent years have seen very impressive and exciting developments in the fields of machine learning and artificial intelligence (AI). Stunning progress has been made on real-world performance of face recognition, object recognition, speech recognition, language translation and myriad other problems. Google's AlphaGo program defeated the world champion in the board game of Go, and IBM Watson handily beat the Jeopardy champion a few years ago.

Within machine learning, the subfield of deep learning has been inspired by the vision of using brain simulations and understanding to improve design, training and performance of learning algorithms. Taking advantage of graphics processors and parallel computing, new and significantly more effective techniques for feature learning and training multi-layer neural network models have been developed. Deep learning is particularly well-suited to perception-oriented tasks. But a particularly interesting recent example is the direct application of deep learning to self-driving cars. Reinforcement learning, a key technique in machine learning, has strong connections with control theory in general and system identification in particular. Recently, ideas from reinforcement learning have been combined with deep learning architectures leading to significant progress in AI.

These advances are creating a fertile ground for deeper collaborations between Systems & Control researchers and the machine learning and AI community. Fundamental requirements from a control viewpoint are likely to pose new questions in machine learning and AI. Conversely, methodological advances in machine learning and AI potentially offer new powerful tools for identification and control. Such collaborations would occur through creative combinations of techniques from these related domains and/or they could be inspired by major application needs such as design and verification of intelligent autonomous systems. Another key direction involves utilising streaming data from distributed sensors over communications networks to achieve control over large-scale distributed systems such as energy systems, traffic systems, etc. The best-case scenario would leverage the unique strengths of these complementary areas to provide analytical assurance on performance, stability and the generalized applicability, higher level reasoning, perception and intelligence.

4.3. Complexity and control

4.3.1. Complexity in dynamical systems

Complex engineering systems are often modelled by nonlinear ordinary differential equations (ODEs) and partial differential equations (PDEs) with or without time delays. Nonlinear control theory is characterized by major discoveries and numerous achievements in solving problems of a general and difficult nature over the last four decades. Emerging after the topics of nonlinear optimal control of the 1960s, the first theoretical breakthroughs of nonlinear control theory addressed structural issues of geometric nonlinear control in the 1970s and the feedback linearization and regulator theories of the 1980s. The 1990s ushered in the era of stabilization theory and adaptive and robust nonlinear control founded on the tools of control Lyapunov functions, input-to-state stability, nonlinear small gain theorems, integrator backstepping, forwarding and nonlinear observers. The most recent decade has witnessed breakthroughs in designing controllers for infinite-dimensional nonlinear systems, including nonlinear systems with delays, the development of analysis tools for hybrid systems and the resurgence of extremum seeking. Applications of advanced nonlinear controllers have become commonplace in aerospace and automotive engineering, energy systems, process control and numerous other areas.

From its modest beginnings in the early 1990s, when the focus was on specific structures of interconnected scalar ODEs, nonlinear control design has been ever-advancing, with future challenges arising from interconnected systems of multiple PDEs and ODEs with nonlinearities, as well as with PDEs whose moving boundaries are governed by differential equations.

How does one control complex dynamical systems with modeling uncertainties, delays, sampled-data measurements, and constraints on input and states?

Employing observer and controller designs for such complex dynamical systems, and realistic scenarios such as modelling uncertainties, delays, sampled-data measurements and constraints on input and states, will keep the area of nonlinear control in step with the development of applications in engineering, biomedical and social sciences.

4.3.2. Complexity in networks

Many self-organizing systems are almost invariably complex dynamical systems since they depend upon the fact that, in certain parameter regimes, when the number of constituents reaches a critical density, coordinated behaviour ensues and patterns emerge. Similarly, evolutionary and adaptive systems are typically complex dynamical systems since they consist of numerous constituent components that individually represent relatively simple physical or biological processes, computational agents, etc., and these interact with each other according to known rules as to define a system, i.e. 'a system of systems.' The individuals may be similar or distinct, defining homogeneous or heterogeneous systems, respectively. The behaviour of isolated individuals is typically described by nonlinear dynamical systems. Exchanges among individuals determine connection graphs with complex topology that define networks of varying degrees of connectivity and an evolutionary or random/stochastic topology. In many cases the network structure is only implicitly known, concealed within massive quantities of empirical and simulation data.

Even though the nature of each constituting system and the interconnections among them differ drastically from one domain of study to another, at the level of mathematical and even philosophical abstraction, they exhibit the same characteristics, share similar requirements and may be analysed via common approaches.

Complex systems exhibit some or all the following features:

- **Self-organization**. Coherent macro-level behaviour of a complex system appears as a result of local interconnections between the micro-level components. Examples: swarming among birds, insects and fish; growth and decay of the human body; the development of the World Wide Web.
- **Adaptation/evolution.** Complex systems often exhibit the ability to evolve and adapt to changing environmental conditions. Examples: evolutionary development in biology (through mutation and selection) and rapid adaptation of individuals in crisis situations.
- **Transition.** Significant positive feedback or random forcing can lead to transitions between meta-stable states. Transitions at tipping points may exhibit a hysteresis effect, making return difficult. Transitions may be preceded by early warning signals such as increased correlation of dynamic behaviour and slowed recovery from perturbations. Examples: phase changes in materials, irreversible ecological changes, climate change and economic instabilities.
- **Fragility and resilience.** Behaviour of self-organized, adaptive dynamics may be highly robust with respect to external perturbations: the system can be pushed far from its equilibrium and still return to it when the external force is removed. Yet other systems may be quite sensitive and vulnerable to certain perturbations. Examples of resilience in complex systems: homeostasis in biological systems, robustness of the Internet to random failures and vulnerability to targeted attacks, and self-organization via social networks. For instance, in energy-transformation networks, the improper management of faults, overloads or simply adding to, or subtracting a generator from, the transportation network may result in power outages or even in large scale (continent-wide) blackouts. In neuronal networks, experimental evidence shows that inhibition/excitation unbalance may result in excessive neuronal synchronization, which, in turn, may be linked to neuro-degenerative diseases such as Parkinson and epilepsy.

How does one control complex systems exhibiting self-organization, adaptation or evolution, transition and fragility or resilience dynamics?

Complex systems have always been around. What is new and what has increased the urgency for research in this area, is mankind's success in engineering systems with complexity beyond our understanding, systems that can change our welfare and lifestyles if we can control them. However, the greater challenge here, is learning to control such complex systems. Perhaps at a first glance less technical, an example of a global economy connecting all banks over the world, producing and trading extremely intricate financial products have brought the world a serious financial crisis. Understanding such a banking network is, at the moment, far beyond our understanding. To what extent the earlier mentioned characteristics of self-organization, adaptation/evolution, transition and fragility and resilience, play a role in a network of banks, is by no means easy to answer.

The following discussion highlights some of the key features of complex systems:

- Complexity can arise in different ways, one of which is due to nonlinear individual dynamics; another is due to the large number of intricate interconnections existing in the system. The interplay between these two sources of complexity - network complexity and individual nodal dynamics and its influence on collective network dynamics is one of the key issues in the analysis of complex systems. In addition, both an increase in the number of system components and changes in the structure of interconnections can allow for many different types of system behaviour.
- Many complex systems – electrical, social, biological and neuronal – often present a natural several-timescales partition which may be induced by a multi-layered structure of the system (e.g., as in electrical networks), hierarchical network structure, dynamics of system component interconnections (as in neuronal systems), and possibly, dynamics of interconnections.
- Most complex systems are 'open' – they never function in isolation but in interaction with an environment. Therefore, analysis of such systems should also consider uncertainties and perturbations induced by interactions with the environment, which can be dynamic or heterogeneous. Perturbations in network topology, computation or communication may propagate errors throughout the network that can degrade performance or, worse, result in positive feedback loops which tend to amplify the effect of the errors, thereby destabilizing the system.
- Robustness of complex networks with respect to misbehaviour of agents, which may be either due to accidental faults or adversarial intrusion of potential adversaries. Among the fundamental issues in this direction are the influence of structure modularity and delayed coupling on the robustness properties of the overall network and its collective behaviour.
- One more source of complexity concerns **spatial networks.** While most of the work on complex systems heavily exploits the structural and topological properties of the network interconnections, the spatial aspect has received less attention. However, constraints imposed by spatial/geographical embedding can critically affect behaviour of a networked system.

The high complexity of these systems implies that many relevant questions cannot be addressed by considering an isolated component of the system or even using the knowledge available within a single discipline. Instead, these questions are inherently multidisciplinary, requiring understanding of the interplay between different disciplines. Issues surrounding such coupled, interdependent networks are now becoming subjects of research, but because of their inter-disciplinarity, this research is still in its infancy.

Addressing these challenges necessitates, besides the quest for understanding complexity, developing methods to influence or manage or otherwise gain control of such complex systems. The Systems & Control researcher is needed in the inter-disciplinary complexity arena!

4.4. Critical infrastructure systems

Critical Infrastructure Systems is a key class of applications where safety and security are becoming extremely important. These systems provide the lifeline that physically ties communities and facilitates quality of life and economic growth. Examples include energy and power systems, water systems, telecommunication networks, transportation systems and healthcare. The safety and security of critical infrastructure systems is a crucial challenge for the years ahead. Experience has shown that critical infrastructure systems *do fail*. Failures may occur due to natural disasters (such as earthquakes, flooding, etc.), accidental failures (such as equipment failures, human error, software bugs), or because of malicious attacks which may occur directly at strategic locations in the network or remotely by compromising communication commands or automation software. Failures in critical infrastructures are becoming more frequent and more dangerous. Large segments and components of the critical infrastructures in most developed countries are old or outdated, resulting in the deterioration of their performance and condition.

When critical infrastructures fail, the consequences may be tremendous in terms of societal, health and economic aspects. For example, if a large geographical area experiences a blackout for an extended period, huge economic and societal costs may result. In November 2006, a local fault in Germany's power grid cascaded through large areas of Europe, resulting in 10 million people left in the dark in Germany, France, Austria, Italy, Belgium, and Spain. Similar cascading blackouts have taken place in the USA. Likewise, there are tremendous health hazards when something goes wrong with the water supply, especially if it is not detected and accommodated quickly. Many businesses cannot operate when the communication networks are down. In the case of faults in transportation systems, we witness the effect of traffic congestion quite often in sprawling metropolitan areas around the world.

Managing and monitoring critical infrastructures systems is expected to become even more difficult in the future. Currently, these infrastructures are huge, and increasing in interconnectedness, exceeding their original capacities for size and complexity. Moreover, due to deregulation and the use of renewable energy, critical infrastructures are increasingly more heterogeneous and distributed. It is becoming more and more difficult to understand the details of how these networks work and to model interactions between various components. Monitoring and control of these infrastructures is becoming more multi-scale, hierarchical and distributed, which makes them even more vulnerable to propagating failures and targeted attacks.

How to manage and monitor critical infrastructure systems? How to model the interactions between the various components?

In situations where a fault arises in some of the components (e.g., sensors, actuators, communication links), or an unexpected event occurs in the environment, serious degradation in performance or, even worse, overall system failure may occur. Standard feedback control systems are typically not able to handle abrupt, significant changes in dynamics due to a fault or persistent erroneous sensor data, while in some cases the feedback controller may contribute to “hiding” incipient faults that develop slowly over time, until it is too late to prevent a serious system failure. The issues of fault detection, diagnosis and automatic recovery will become even more crucial in the future as engineering systems become more interconnected, distributed and interacting, while at the same time functioning under more demanding operating conditions and in more unstructured environments.

In addition to ensuring physical security against accidental faults, the security of complex, large-scale distributed systems against malicious attacks, i.e., cybersecurity, is a crucial issue for many governments and businesses. The challenge of preserving the integrity of coordination and cooperation of the components of distributed control systems against malicious interferences is paramount for control and automation systems. The relevance of such problems became clear after several attacks, most prominently the global-scale Stuxnet worm attack in 2010. Security issues have been addressed in Industrial Automation Protocols both at the LAN/WAN level and the Field Bus and Device level. Control and automation systems are often implemented on embedded devices using software tools. The drive to provide richer functionalities, increased customizability and flexible adaptability requires the ability to dynamically download software to devices. Without adequate countermeasures, this ability may expose vulnerabilities that are conducive to severe security breaches. The magnitude of this problem will worsen with the rapid increase in software content of networked control systems.

How to ensure security against accidental type of faults or against malicious attacks?

Rebuilding critical infrastructures from the ground up is out of the question. Therefore, we need to work with existing infrastructures to enhance their efficiency, reliability and sustainability utilising instrumentation and smart software. To achieve this, there is a need to develop both a system-theoretic framework for modelling the behaviour of critical infrastructure systems and algorithms for intelligent monitoring, control, fault diagnosis and security of such systems.

The goal of the framework needed to develop smart critical infrastructures, is the enhancement of their reliability, fault-tolerance and sustainability. Several features of such infrastructures provide opportunities for realizing this goal. The first of these is the increasing presence of sensors and actuators. Whether it’s electric power systems, water distributions networks, manufacturing processes, transportation systems or robotic systems, a common feature is extensive instrumentation for monitoring and control. Networked embedded systems and sensor/actuator networks are increasingly present in these infrastructures, where a large amount of sensor data is collected and processed in real-time to activate the appropriate actuators and achieve the desired control objectives. The need for advanced monitoring and control algorithms is becoming more crucial and challenging as engineering systems and infrastructures are becoming more complex, large-scale, distributed and interactive. New sensor/actuator devices are continually being developed at reduced costs and in larger quantities, while traditional monitoring and control specifications and objectives are being expanded to include new system aspects such as energy efficiency, environmental impact and security.

A second feature is the distributed or decentralized nature of controlled systems, which necessitates a hierarchical architecture for ensuring safety and security, where neighbouring fault diagnosis agents are cooperating at a local level, while transmitting their information, as needed, to a regional or global monitoring agent, responsible for integrating in real-time local information into a large-scale “picture” of the health of the overall system. Key motivations are to exploit spatial and temporal correlations between measured variables and develop the tools and design methodologies that will prevent relatively small faults or unexpected events from causing significant disruptions or complete system failures in complex distributed dynamical systems.

Existing approaches to automated fault diagnosis and fault-tolerant control deal with monolithic centralized systems under certain constraints on the model and the type of fault. For distributed large-scale systems, a new monitoring and control paradigm is needed to address distributed unknown environments, where mathematical models may not be accurate or available. Therefore, fault diagnosis and security methods must be developed, where the time history of the observed data and the inter-relations between spatially distributed sensing are exploited. It is important to note that in the case of distributed interconnected systems there is additional redundancy due to the spatial correlations between observed data. This spatial redundancy is present not only from an observation viewpoint, but also from an actuation perspective, and can be exploited using learning and cooperative schemes in a fault-adaptive framework.

Yet another feature of critical infrastructures that should be noted is the effect of interdependence between critical infrastructures. As digital instrumentation and communications become more advanced expand to a global scale, this effect takes even more of a centre stage. For example, the normal water supply and telecommunications operations require the steady supply of electric energy. The reverse is also true: generation and delivery of electric power relies on the provision of fuel and water, as well as various telecommunications for transferring data and controlling the power plants and networks. Brief power outages may result in traffic congestion, or worse, fatal accidents. The increasing interdependence between critical infrastructures raises concerns about the vulnerability of these interconnected systems and the possible effects of cascading faults between various infrastructures.

4.5. Cyber-physical system of systems

Increasingly, attention is being focused on the understanding and design of large-scale system of systems where cyber and physical components must interact and integrate to realize overall performance measures. This class of systems, denoted as Cyber-Physical System of Systems (CPSoS) introduces a variety of challenges and opportunities for the Systems & Control community (Engell et al., 2016). For optimization and control of CPSoS, which are spatially distributed processes (e.g. electric power networks, large chemical or metallurgical plants, traffic systems, water distribution systems), a key issue is how to operate a large-scale physical system to maximize overall performance measures under operational constraints, drawing from a broad range of information sources, and to ensure a robust and stable operation even in case of subsystem failures. Performance measures typically describe economic and environmental criteria, while constraints are related to dynamical restrictions of the process, limitations of equipment and resources, and policy-based and regulatory boundaries (e.g. safety codes and emission limits). Information sources include measurements from numerous sensors with different accuracies, reliabilities and availabilities, models of processes, sensors and communication channels and external factors such as market analyses and weather forecasts.

How should wide-area socio-technical systems with heterogeneous information sources be managed and controlled in operational regimes to maximize overall performance under relevant constraints?

Due to the scope and complexity of large-scale cyber-physical systems, as well as ownership or management structures, the control and management tasks in CPSoS may be impossible to accomplish in a centralized or strictly hierarchical top-down manner with one authority tightly controlling all subsystems. Instead, there is a significant distribution of authority with partial local autonomy. In order to analyse and synthesize such CPSoS to ensure reliability, efficiency and other desired performance metrics, several challenges need to be addressed, the most important of which are outlined below.

- **Distributed autonomy.** The interaction and coordination of systems with partial autonomy in systems of systems, possibly with dynamic membership, must be studied broadly. Examples of applicable methods include population dynamics and control and market-based mechanisms for the distribution of constraining resources. In distributed management and control systems, local controllers have limited information regarding the evolution of other subsystems and the actions of their controllers. The value of information for stability and performance, as well as the resiliency of distributed systems to failures of system components needs to be analysed. Methods for the systematic design and verification of mechanisms for dynamic reconfiguration of the overall system due to a disconnect or shutdown and re-integration of subsystems are required. The partial autonomy of the components from the overall system of systems perspective leads to uncertainty about the behaviour of the subsystems. Therefore, system-wide coordination must consider uncertain behaviour and must nonetheless guarantee an acceptable performance of the overall system. Stochastic optimization and risk management require further development. Additional understanding regarding the influences of management structure (centralized, hierarchical, distributed, clustered) on system performance and robustness is also necessary.

How can we design and verify methods for detection, reconfiguration, and mitigation in wide-area socio-technical systems to deal with different types of failures and recoveries?

Such features have naturally led to the development of methods in Systems & Control theory which are based upon a synthesis of dynamical game theory, stochastic control and (non-classical) information theory. The present challenge is to fully develop such a synthesis and integrate it with methods permitting computationally tractable real-time solutions.

- **Failures, faults and anomalies.** Failures take even more of a centre stage in a CPSoS, as the number of faults and their impacts increase enormously with expansion in scale and the number of cyber-physical interactions. Hence there is a strong need for mechanisms that detect abnormal states and for fail-safe measures and fault tolerance at the systems level. Advanced monitoring of the system state and triggering of preventive maintenance based on monitoring results can significantly reduce the number of unexpected faults, maintenance costs and downtimes. Faults may propagate through multiple layers of the management and automation hierarchy. Many CPSoS often experience cascading effects of component failures. These abnormal events must therefore be handled across the layers.
- **Adaptation and Plug-and-Play.** CPSoS are operated and continuously improved over long periods of time. Performance targets and system structures change, and the overall system must adapt to these changes while maintaining stability and performance. New functionalities or improved performance must be realized with limited changes to portions of the overall system. Components are modified and added, and the scope of the system may be extended or reduced. So engineering, to a large extent, must be performed at runtime. Additions and modifications of system components are greatly facilitated by plug-and-play capabilities of components that are equipped with their own management and control systems (“decentralized intelligence”).
- **Cybersecurity**. Cybersecurity is a very important element in CPSoS. A specific challenge is recognizing obstructive signal injections or takeovers of components in order to cause malfunctions, suboptimal performance, shutdowns or accidents, e.g. power outages. Detecting such attacks requires considering both the behaviour of the physical elements and the computerized monitoring, control and management systems. To detect unsecure states, suitable isolation procedures and soft (partial) shut-down strategies must be designed.
- **Emergent behaviour and self-organization.** Due to distributed autonomy and dynamic interactions, CPSoS can realize self-organization and exhibit structure formation and system-wide instability, in short, emergent behaviour. Predicting these system-wide phenomena is an open challenge at the moment. Distributed management and control methods must be designed such that the overall system does not show undesired emerging behaviour. Inputs from the field of dynamic structure or pattern formation in large systems with uncertain elements should be combined with classical stability analysis and assume-guarantee reasoning. Methods must be developed such that sufficient resiliency is built into the system so that local variations, faults, and problems can be absorbed by the system or be confined to the affected subsystem and its neighbours, and no cascades or waves of disturbances are triggered in the overall system. In fact, emergent behaviour can be regarded as the result of successful adaptation. But the challenge is to understand this as a process in a dynamic recursive setting. Rather than terminate, evolutionary processes respond to continually changing circumstances by increasing the sophistication, and typically the complexity, of a system’s control mechanisms.

How can we leverage the structural and physical properties of the wide-area socio-technical systems to develop effective solutions to cyber-security problems?

- **Interactions with humans.** The human role becomes significantly more complex in a CPSoS, as it is no longer limited to that of user or operator, but becomes that of empowered agent. Human interventions in decision-makingintroduce an additional nonlinearity and uncertainty in the system. Important research issues include human capacity of attention and how to provide motivation for sufficient attention and consistent decision-making. It must be investigated how the capabilities of humans and machines in real-time monitoring and decision making can be combined optimally. Future research on the monitoring of the actions of the users and anticipating their behaviours and modelling their situation awareness is needed. Social phenomena (e.g. the dynamics of user groups) may also be considered.

How to co-design control algorithms and communication protocols?

The efficient design and operation of such systems requires new design support methodologies and software tools in the following areas:

How can we do model-based design of complex real-time systems with heterogeneous and diverse modeling tools and representations for various component subsystems?

- **Validation and verification in CPSoS.** A typical example of a CPSoS is a distributed embedded system with several CPUs and communication buses acting on multiple physical systems. Given the combined presence of continuous dynamics and discrete switching, CPSoS design verification is nontrivial and cannot be carried out through exhaustive testing. Validation and verification (V&V) requires a combination of tools from Systems & Control science, such as those used for stability analysis, with tools from computer science, such as static analysis, model-checking, formal verification methods, and other tools from real-time analysis. Different aspects of control systems are often modelled using different formalisms, especially if both continuous and discrete dynamics are present. This calls for the integration of multiple formalisms into one environment, seamlessly reusing models and integrating results obtained by different tools. Besides the transformation of models from one formalism into another, abstraction of models with guaranteed preservation of certain properties is crucial, as it is neither feasible nor reasonable to try to consider all details of a system in each stage and on each level of the design process.

  New approaches are needed for V&V during the continuous evolution of a CPSoS, ensuring correctness by design during its evolution, and for verification, especially, on the system of systems level. New algorithms and tools should enable the automatic analysis of CPSoS which are large-scale, dynamically varying and evolving. This includes formal languages and verification techniques for heterogeneous distributed hybrid systems including communication systems, theory for successive refinements and abstractions of continuous and discrete systems so that validation and verification at different levels of abstraction are correlated, and the joint use of assume-guarantee reasoning and simulation-based (Monte Carlo) and exhaustive (model checking) verification techniques.

- **System Integration in CPSoS.** One of the most difficult and important problems to solve in the near future is integrating mathematically-sound control design and analysis methods into industrial environments where heterogeneous software and formalism dominate, and where design data, documentation and models abound on all design levels. These must be integrated such that inconsistencies and errors are detected early. Such integration can lead to an enormous increase of the efficiency in the design phase and in later operations. While the lack of such integration today leads to a large waste of resources and lack of robustness to errors, we can expect that as systems design projects become more and more complex, the lack of such integration mechanisms may lead to an inability to successfully complete complex projects at all.

4.6. Autonomy, cognition and control

In consumer products, there is an increasing trend towards higher degrees of autonomy in complex technical systems such that they can perform complex tasks (“missions“) and react autonomously to environmental influences rather than being guided by humans. Not incidentally, the term “mission” comes from the military sector, where research has recently focused on unmanned (semi-) autonomous vehicles (UAV).

Future autonomous systems such as robots in search and rescue operations, mining, deep sea exploration, etc. will need cognitive capabilities including perception, reasoning, learning and others. Cognitive control systems in automobiles, aircraft, houses and engineering applications will enhance safety and performance. An example is an energy management system in a hybrid electric vehicle: by continuously perceiving the driving style of the pilot and learning to model this behaviour, it may leverage this knowledge to optimize fuel consumption. Cognitive control can also be used in social and group environments where cooperative execution of complex tasks is required of agents that can be humans, machines or both.

Autonomous and cognitive systems:

- exhibit goal-oriented behaviour in sensing, thought and action and flexibly change goals and behaviour depending on context and experience;
- act in unstructured environments without human intervention and robustly responds to dynamic changes;
- interact with humans and other cognitive systems to jointly solve tasks.

Today’s autonomous systems operate well in static, predictable environments, but cannot cope with uncertainty and dynamic changes induced by interaction with complex and intelligent systems, such as humans. Designing a control system for this type of task requires acquiring knowledge and understanding of the interacting players via perception, reasoning and learning. Autonomous systems, e.g., parking assistants, home robots for the elderly and (almost) unmanned production systems, consist of a hierarchy of feedback control loops in which large streams of data, often from image sensing, are processed. Today, collecting and storing large amounts of data is no longer a challenge, but extracting meaningful information from these data requires the adequate combination of methods that are based on explicit mathematical models and hence can be analysed rigorously. Advances in computing power and algorithms currently enable systems with a high degree of autonomy to be conceived, but the route from lab implementation to safe and certified operation in the real world is very long. For deployment of autonomous systems, their behaviour must be validated under all conceivable conditions.

How to design autonomous systems with cognitive capabilities so as to maximize synergy with humans for benefits to society?

Systems & Control theory provides well-understood methods to establish stability, i.e., guaranteed convergence to a desired state, of complex systems. But this is inadequate for autonomous systems where continuous dynamics interact with algorithms that make discrete decisions. Testing of such systems can only cover a very small portion of the possible situations. A current trend is ex-post verification which is out of reach for systems of significant size since it can only occur for a very small part of the possible situations. Therefore, a major challenge for the next decade is to design such systems with in-built verifiable monitoring and control mechanisms. To address such a challenge, interdisciplinary research is required on the edge of Systems & Control theory, information science, computer science and statistics, which will enable the verification and validation of highly autonomous systems.

The rapid development of technology for sensing and actuating devices, data storage and communication creates new opportunities for data-based control. In particular, increasingly complex data sets and models require methods that can handle the data in a scalable and robust way. Applications include energy grids, transportation networks as well as health care systems. System identification is the term for data-driven modelling used in the control community. The dominant approach, based on statistical analysis of parametric models, has been very successful as a basis for feedback control. However, the need to handle large data sets in high dimensions is currently inspiring new methods using non-parametric models closely related to Machine Learning. More research is needed to analyse how these models can be used efficiently in real-time applications.

How to control autonomous systems acting in unstructured environments without human intervention and robustly respond to dynamic changes?

Systems of Systems (SoS), by their very nature are large, distributed and extremely complex, presenting a myriad of operational challenges. To cope with these challenges there is a need for improved *situational awareness.* Gaining an overview of an entire SoS is inherently complicated by the presence of decentralized management and control. Introducing cognitive features to aid both operators and users of complex CPSoS is a key requirement to reduce the complexity management burden from increased interconnectivity and the data deluge presented by rising levels of data acquisition. This requires research in numerous areas to allow vertical integration from the sensor level to supporting algorithms for information extraction, decision support, automated and self-learning control, dynamic reconfiguration features and consideration of the sociotechnical interactions with operators and users. The following key subtopics have been identified as being necessary to support a move to Cognitive CPSoS:

- **Situational awareness in large distributed systems with decentralized management and control.** Operating a SoS efficiently and robustly requires abilities to detect changes in demands and operational conditions (from both the equipment and other factors) and address anomalies and failures within the system. This can only be achieved via the introduction of much greater levels of data acquisition throughout the CPSoS and the use of this data for optimization, decision support and control. A key enabler is the introduction of novel, easy to install, low cost, sensor technologies and monitoring concepts. If wireless monitoring is to be used, ultra-low power electronics and energy harvesting technologies will be required to avoid the need for, and associated maintenance costs of, battery change. Increased data gathering will also require robust wired and wireless communication protocols that can deal with efficient transmission of individual data values from a multitude of sensors to streaming data at high rates, e.g. for vibration and video monitoring.
- **Handling large amounts of data in real time to monitor system performance and to detect faults and degradation.** Future challenges include: physical system integration of highly complex data acquisition systems, management of the data deluge from the plethora of installed sensors and the fusion of this with other information sources. This will require analysis of large amounts of data in real-time to monitor system performance and to detect faults or degradation. Here there is a need for visualization tools to manage the complexity of the data produced allowing managers to understand the “real-world in real-time”, manage risk and make informed decisions on controlling and optimizing the system.

How to design and maintain complex high-performing control systems in a fully model-based fashion?

- **Learning good operation patterns from past examples and auto-reconfiguration and adaptation.** There is a great opportunity to aid system operators by incorporating learning capabilities within decision support tools to identify good operational patterns from past examples. Additionally, to deal with the complexity of managing system faults, which is a major burden for operators of large systems, auto-reconfiguration and adaptation features can be built into the systems.
- **Analysis of user behaviour and detection of needs and anomalies.** Finally, humans are an integral element in socio-technical systems. Therefore, systems of systems need to be resilient to the effects of the natural unpredictable behaviour of humans. There is thus a need to continuously analyse user be-

haviour and its impact upon the system to ensure that this does not result in system disruption.
The result of combining real-world, real-time information for decision support with autonomous control and learning features will be *Cognitive Cyber-Physical Systems of Systems* that will support both users and operators, providing situational awareness and automated features to manage complexity that will allow them to meet the challenges of the future.

Much can be expected from the development of technologies for increased autonomy from military and niche applications to broader markets, e.g., driver assistance, personal assistance, monitoring of elderly people, rehabilitation, intelligent heating, etc. The basis of autonomy is the processing of data from the environment and thus feedback – a classical feedback loop is the simplest implementation of the principle of autonomy of a technical system. Autonomous systems pose enormous challenges to design and testing methodologies in order to assert safe and functionally correct behaviour of such very complex real-time data processing systems in all conceivable situations.

4.7. Cyber-physical and human systems

In recent years, the increasing number of sophisticated machines that cooperate closely with humans (e.g. in healthcare applications) or that require human operators have led to debates on how to best consider human factors in Control. This trend has also been visible through the increasing representation of human factors in control conferences and events over the last decade.

While many control loops in technical devices of all kinds work autonomously without human intervention (apart from the initial tuning), large systems such as airplanes, power plants and urban traffic control systems are controlled by human operators who interact with the automatic control systems.

In human operator controlled systems, one type of interaction has operators (e.g., pilots) controlling the systems but with their orders “translated” into the actuation (movement of the wing flaps in an airplane, opening of valves in a power plant) by a subordinate control system. In this case, the dynamics of the controlled system must be shaped in such a way that its behaviour is considered as satisfactory by the human operator. Therefore, the scientific approach is typically guided by human factor analysis. A more advanced approach has the controlled system working autonomously while human operators supervise its behaviour and intervene in case it is not judged appropriate or safe, by modifying set-points or by switching (partly) to manual control. Early detection of faults or changes in system behaviour is a crucial task in this case. A control system may also compute proposals for optimal system operation, e.g. an optimal adaptation to a change in power demand in an electrical power plant, and offer these proposals to the operators who can accept or modify them and trigger their execution.

In human cyber-physical shared control systems, there is no superiority of one or the other: humans and the cyber-physical system (robots) share the tasks and goals, they communicate and coordinate their joint action.

Typical control system designs usually rely on three main steps: modelling, decision and control. When a human being is a part of such a control system all three steps become much more challenging to tackle, as illustrated below.

Modelling includes a large scope of needs, and thus different natures of modelling have been explored in the last decades. In shared control, models of the human capabilities are equally important as the robotic system. There are two types of models for the human: physical capability models and cognitive models.

- Physical capability models rely on understanding the kinematic and dynamic capabilities/limitations of the individual. Consider an example of jointly human-robot lifting a heavy object. By understanding the kinematic capabilities of the human, the robot can adjust its configuration to the human to accomplish the task. Another example of this symbiotic collaboration occurs between the human driver and semi-automated car. The design of lane-keeping control systems that share the control of the steering wheel with the driver is now possible.
- Psychologically-grounded cognitive models are needed for other kinds of tasks involving tactical and strategic thinking. Examples include body motion modelling: human instrumentation with sensors in the day-to-day life for detection of fatigue-related indicators after a cerebrovascular accident (VCA); models of the creativity processes; modelling the human-decision making; and models of the motion of groups of human beings. On another level, brain models continue to represent a huge challenge for healthcare applications.

Decision systems have grown in importance with the increasing of systems complexity, as illustrated below:

- **With the deployment of automated systems, the switch from automatic to manual mode and vice-versa is a key issue** that has been widely discussed lately by the industrial and academic communities. This includes the anticipation of fault detection and consequently the need for switching off the system when fault is detected in advance. Also, the decision of what to do in situations when fault cannot be detected in advance (e.g. fault in an automated vehicle leading to an emergency stop) is a key recent research topic. In aviation, the importance of modelling the pilot-aircraft coupling to prevent adverse interactions and determine when to return control of the aircraft to the pilot, has recently been underlined. This has high potential in the prevention of loss-of-control events (Hess, 2002).

How to optimally conjugate automated systems with the interplay of humans?

- **Also, the modelling of the human can play a very important role in the decision system design.** In transportation, an example is found on the choice of the driving assistance system to be activated in each risky situation - a mental driver model is essential to accomplish the choice - in driving, the control processes in the human mind can be described at different levels, with a widely-adopted scheme that proposed a description of the driving task in three levels, as follows. The *operational level* refers to the moment-to-moment vehicle handling where only an *intervening* assistance system can recover a safe state of the vehicle, if the situation is such that it demands the driver to use his mind in this level. For example, if the vehicle slides due to an icy road, the driver’s mind is functioning in the operational level and an intervening system (ESP-like system) is absolutely necessary to recover a safe state. The *tactical level* includes selection of speed and following distance, decision to overtake, etc., in which the driver has some time to plan. Intervening or warning systems can be enough in the tactical level depending on the situation. Finally, the *strategic level* includes tasks such as route planning and navigation for which the driver has time to plan his actions far in advance. Warning systems here can hugely increase accident prevention. As opposed to the first example, drivers will receive an alert notifying them of upcoming road hazards and will have enough time to progressively reduce their speeds before entering the risky zone. These three human cognitive models correspond to three (temporal) types of driving assistance systems that involve different challenges in terms of design and evaluation. A decision system is normally very complex, and many other parameters can be involved as follows: the movement of the surrounding vehicles, the awareness of the driver, the weather, etc.

How to integrate knowledge from social and behavioral sciences with cyber-physical systems to create a new field of Cyber-Physical and Human Systems?

In control, we identify two main design classes: human-centered or machine-centered control designs. In the first, the goal is to help a human. In the second, the goal is to properly operate/control a machine. The human-centered design requires high-level understanding and modelling of the human being. The machine-centered design often involves a H-M interface design since, in most cases, a human operates/supervises the machine(s) and process(es) as pointed out above. Many other fundamental research topics need to be studied and considered when designing a control system for the human being. For example, situational awareness – keeping the human operator "aware" when operating a machine. In the automotive domain, the "complacency" phenomenon may appear in case of an overly assistive system. This parallels human physiology where a constant sensory signal (e.g. the pressure of a seat on your back), is rapidly ignored as there is no novel information. Thus, the system must be designed to engage the driver with informative feedback but avoid saturation. Another element is the level of dependence of an individual on a machine **(e.g.** some people can't live without an insulin pump or a dialysis machine). Similarly, the level of intrusiveness of a machine with respect to the human, is a crucial point in the driver assistance systems design, that is also closely related with the situation awareness: the willingness of an individual with respect to the level of help given by the machine. Disabled people frequently wish to determine the level of assistance themselves. A final element involves the adequate design of controlled systems – meeting human needs while allowing humans to effectively interact with them – is a pressing issue in all kinds of assistance systems, especially for elderly and disabled persons. Learning is a process governed by feedback as well, so the design of computer-based teaching systems could also benefit much from a system-theoretic understanding of the dynamics of human learning and how it can be stimulated.

How to support humans with autonomous systems in a cooperative fashion?

Advanced control offers huge benefits for a more economic, safer and ecologically more benign operation of potentially unsafe technical systems, but these benefits are often not fully achieved because human operators do not accept the automatic system and switch off important parts of its functions. The more complex and sophisticated control systems become, the more difficult it is for operators to understand the computed strategies and this may lead to more instead of less manual interventions. There is little understanding of the information needs of the operators and of the way in which they perceive and process the information provided.

The interplay of humans and highly automated systems with complex dynamics and complex controllers or the design of complex control systems to assist humans require deep study with multidisciplinary efforts. Researchers from cognition, ergonomics and Systems & Control are examples. Furthermore, at a higher level, Cyber-physical systems are strongly impacting the society in many arenas: healthcare, mobility and urban planning, for example. Market demands are changing, and impacts to the environment are significant. Going further, deeper ethical and philosophical discussions on which are the frontiers of technology and human constitute a fundamental topic. What are the limits on what only the technology can do? What can (should) only a human being do? We have discussed the Human-CPS systems and their benefits, but which cases concern only the machine and which cases concern only the human have not been considered.

Currently, levels of CPS-Human cooperation include:

- Human-machine symbiosis (e.g. smart prosthetics)
- Humans as "operators" of complex engineering systems (e.g. aircraft pilots, car drivers, and process plant operators, robotic surgery)
- Humans as agents in multi-agent teams (e.g. road automation)
- Humans as elements in controlled systems (e.g. comfort control in homes and buildings)

Future applications will to a much higher extent intertwine human operation with technical equipment. For the example of manufacturing systems, today's paradigm of strictly separating the workspace of robots and human operators will be partially replaced by cooperation of humans and robots to increase efficiency of complex production steps. In future traffic systems, autonomously operating vehicles will interact with human-operated vehicles. **Future control systems must therefore reach a new level of robustness with respect to the non-determinism of human behaviour. Control techniques must be able to timely respond to new situations and unpredicted events arising from human behaviour,** i.e. new controllers must be designed for system dynamics which include various types of probabilistic components and stochastic distributions. Achieving high-control performance while guaranteeing safety (e.g., avoiding collision of vehicles) will be a major challenge in this field.

5. Examples of high-impact Systems & Control applications in the coming decades

In this section, twenty-two examples of Systems & Control applications are sketched. For each of them it is shown how the Systems & Control can contribute to facing the corresponding challenges. Therefore these examples illustrate the cross fertilizing and bi-directional interplay (Fig. 1) between the FIVE critical societal challenges described in Section 3 (Transportation, Energy, Water, Healthcare and Manufacturing) with the SEVEN key research and innovation challenges described in the Section 4 (Cyber-Physical Systems of Systems; Distributed networked control systems; Autonomy, cognition and control; Data-driven dynamic modelling and control; Cyber-Physical & Human Systems; Complexity and control in networks; Critical infrastructure systems). Each sub-section of Section 5 links to at least one of these critical societal challenges.

5.1. Road and air traffic management

5.1.1. Road traffic management

Rush hour traffic congestion is common in most metropolitan areas. The most obvious impacts of traffic congestion for citizens are increased travel times, fuel consumption, emissions and noise.

These effects are amplified when infrastructure is not operated at maximum capacity during congestion, implying that fewer vehicles than possible can proceed. The situation is worsening due to the continuous increase of transportation demand while space is lacking to build new infrastructure. New technologies in sensor networks and floating car data allows information about the traffic to be measured in many ways and at many points on the roads. This information can be used to develop innovative adaptive traffic management policies, to provide real-time high-value information for users and traffic operators. These include traffic-responsive ramp metering and varying speed limits, traffic prediction indicators (i.e. traveling time, optimal routing, incident indications, etc.).

Traffic problems are typically addressed at the level of a single vehicle or subsystem (e.g., in a specific arterial corridor or a part of an urban road). Current control and resource optimization strategies are inefficient when considering traffic at the local network level. Today's fragmented and uncoordinated approach presents a significant obstacle to improving urban mobility and energy efficiency. Future on-line traffic management systems will use integration of and cooperation between intelligent vehicles and roadside infrastructure via in-car navigation, telecommunication and information systems to provide a balanced utilization of the transportation system that considers various objectives and constraints such as travel times, reliability, delays, emissions, CO2 reduction and fuel consumption.

Control challenges associated with this objective are:

- **Mathematical models**: New traffic models need to be developed to account for the diversity of elements in the traffic system, such as vehicle classes, and multiple transportation modes. These will lead to design control strategies for traffic regulation specific for each vehicle class and transportation mode. The multi-modal models to be used as a basis for the controller design need to be simple, in order to keep the controller complexity to a minimum, as required by online implementation. The multi-modal models can be also used for the validation of the control algorithms, as well as to monitor and control the overall pollution of the traffic network. Indeed, it is well-recognized that vehicles are distinct in terms of the amount and type of the emissions that they generate. Then, a model accounting for multiple vehicle classes (e.g. distinguishing between cars and trucks) and transportation modes (i.e. separately considering buses, trams, etc.) results in being more appropriate than a conventional single mode model in representing a traffic system as far as air and noise pollution is concerned. The multi-modal models used for validation and emission evaluation are necessarily more complex that those used for the controller design. Moreover, since emissions strictly depend on the age, type and size of each vehicle, a combination of micro and macro models will be necessary to capture these aspects with sufficient accuracy. More refined models will also require more efficient and advanced computing facilities, such as parallel and/or cloud computing.
- **Coordinated control among subsystems**: Control methodologies and architectures for operating the road network as a whole in metropolitan areas will be developed. It is indeed of paramount importance to ensure the expeditious movement of traffic on urban roads and a smooth interface with surrounding traffic networks. The control objective will be to proactively prevent traffic congestion by efficiently distributing the flow of traffic in the metropolitan area, avoiding tailbacks spreading to neighbouring traffic systems. Circulatory traffic signals, variable message signs and suitably synchronized traffic lights will be used to implement strategies designed relying on consensus arguments, game theory approaches, supervisory and distributed control methodologies. The balance of local control actions and their coordination must consider global criteria regarding the average traffic in the entire metropolitan areas and, in particular, the necessity of maximizing the capacity offered by the infrastructures.
- **Model-based travel time forecasting**: This will require development of advanced online prediction algorithms based on modelling of users' travel choices and typical drivers' behaviours, in addition to traffic data provided by the sensors and historical data. The prediction algorithms need to forecast the occurrence of shockwave phenomena and accurately localize the shockwave front in the traffic network. Advanced forecasting could include an offline determination of the typical trends of the traffic evolution in specific moment of the day or of the week, based on the extraction of specific patterns from historical data. This implies the ability of the forecasting system to deal with "big data" and the numerical aspects related to this topic. The offline computed forecasts will be complemented with efficient online adjustments based on the measured data which requires significant efficiency in data acquisition, transmission and processing.
- **Optimal routing for dynamical traffic networks**: This research will focus on new online optimal planning algorithms accounting for traffic flow congestion and modern vehicle-to-network communication policies. In the literature, several road pricing methods have been proposed to reduce congestion in traffic networks with the aim of inducing drivers to follow specific routes so as to reach a system-optimal distribution of traffic on available roads. This can minimize the total time spent by vehicles in the network. In general, there is a non-correspondence between a system-optimal traffic distribution and the behaviour of traffic participants following their individual interests and being insufficiently informed on the current and predicted traffic situation. This lack of information can be overcome in a still futuristic yet realistic scenario by massive vehicle-to-vehicle and vehicle-to-infrastructure, i.e. vehicle-to-network, communication. In this new paradigm, one expects the road authority to act as a leader player by pricing the roads which are the links of the considered traffic network. The drivers, as follower players, are asked to adapt their route choices according to the pricing strategy. The vehicle-to-network communication can contribute to increasing users' confidence in the pricing system and their consequent adherence to the system-optimal routing.
- **Resilient traffic control:** This aspect requires control and communication strategies designed to account for vulnerabilities introduced by subnetwork interconnections and feature resilience against malicious attacks on actuators (e.g., traffic lights) and sensors. Malicious attacks are facilitated by the modern trend in interconnecting supervisory control and data acquisition systems which uses the standard TCP/IP suite of protocols to design, maintain and troubleshoot the communication infrastructure, as well as interconnect different access points through the public internet. Yet, this use of well-known protocols and the interconnection with the public internet, expose the overall control system to malicious attacks. Therefore, even in traffic networks, it is fundamental to guarantee secure operation and provide fault detection and isolation capabilities. Finally, it is advisable to develop strategies to reconfigure the sensor and actuator networks in order to maintain acceptable performance of traffic control systems, even in the presence of faults or attacks. In other terms, traffic systems must be able to defend themselves from attacks by promptly changing their control paradigm and communication topology so as to avoid cascading failures.
- **Exploiting new data sources:** Recent advances in technology have resulted in numerous new monitoring systems which

expand the amount and type of traffic measurements. One may now receive data from radar, mobile phones, Bluetooth-equipped vehicles, video cameras, magnetometers, etc. In order to reconstruct more accurate traffic variable evaluations, heterogeneous traffic data must be fused in a parsimonious way. Moreover, since technology advances in have opened the door to large fleets of probe vehicles, it is also necessary to consider that, together with traditional flow and time mean speed measurements relative to a local section monitored continuously in time, probe vehicles can provide additional data, such as space mean speed and travel time relative to road segments monitored in specific time intervals. Several approaches, mainly based on the Extended Kalman Filter, have been proposed in the literature to address data fusion issues in traffic systems when new data sources are considered. Yet, there is still room for improvement in terms of algorithm efficiency and estimation accuracy. Novel data fusion techniques need to be developed to also improve traffic monitoring systems aimed at automatic incident detection. Finally, more efficient data fusion capabilities will benefit traffic forecasting and traffic control.

- **Secure and privacy-preserving data sharing:** In the emerging scenario in which data sources can be exploited, one must address privacy issues and develop traffic control strategies with communication constrained by secured real-time information sharing and privacy-preserving data aggregation. This is particularly crucial for probe vehicle-based automotive traffic monitoring and control systems. In such systems, it is of utmost importance to guarantee anonymity in a dataset of location traces while maintaining high data accuracy. Well-known anonymization algorithms may fail to provide privacy guarantees for drivers in low-density areas (i.e., where the user density is low, the spatial-temporal characteristics of the data can allow tracking and re-identification of anonymous vehicles) or may not meet the prescribed accuracy requirements. To achieve secure, accurate and privacy-preserving data sharing, and, specifically, to guarantee privacy in a location dataset, it is necessary to design algorithms that hide location samples or modify location traces, while maintaining the accuracy of the original time-series/location data. Metrics describing how long an individual vehicle can be tracked in the data set should also be introduced. Based on these metrics, reliable privacy algorithms capable of guaranteeing a specified maximum time-to-oblivion will be formulated to overcome the dilemma between privacy and accuracy in data sharing.

5.1.2. Air traffic management

Air traffic volume has steadily increased over the past four decades, accelerated by worldwide deregulation of the industry in the 1980s. According to the IATA (www.iata.org), nearly 3 billion passengers and over 50 million metric tons of cargo were transported by air in 2013. During that year, aviation supported 57 million jobs and generated over US $2.2 trillion in economic activity, worldwide.

By some estimates, world aviation is expected to grow by 25% to 30% in the next decade. The accompanying increase in the number of aircraft utilizing air transportation resources will require substantial modifications to the present air traffic control configurations and procedures. Even if air transportation safety metrics manage to remain at their present levels, this large increase in traffic volume will adversely impact system throughput. In anticipation of this fact, Federal Aviation Administration in the United States and the EUROCONTROL Organization have initiated the NextGen and the SESAR programs, respectively. The objective of these efforts is to facilitate a safe path to scaling the air traffic control system without compromising performance. In view of the sweeping changes required to enable this transition, the system has been renamed as the Air Traffic Management System in recent years.

The air traffic management system is a human-centred automated system in which controllers monitor the air traffic and communicate with the pilots to ensure conformance with pre-filed flight plans and approve any changes to them, while safeguarding aircraft separation . An important objective of the system is to maintain throughput under weather and traffic flow perturbations.

Radar-transponder based surveillance and simplex VHF/UHF voice communications were at the core of the system. Until recently, traffic volumes were low enough to enable the airspace to be segmented into the air traffic control centres and sectors, with individual sector controllers ensuring aircraft separation while aligning traffic flow objectives with centre-level traffic coordinators. This approach breaks the problem into a series of scalar flow and separation control problems, amenable to manual control requiring virtually no automation.

Increased air traffic volume requires complex simultaneous interactions between multiple traffic streams to ensure conflict-free merging and spacing. Complete manual control may not be practical without substantially increasing the number of sectors and with them, the attendant communication and coordination difficulties. Availability of widespread Global Navigation Satellite Systems (GNSS) and wireless data communication technologies provide the basis for automating several of the lower-level controller tasks, enabling the eventual elevation of the human controllers to role of traffic managers.

- Just as automatic flight control technologies have substantially reduced pilot workload in the cockpit, while significantly improving aviation safety over the past five decades, emerging automation tools are expected to reduce controller workload and enhance throughput. High-level decisions may continue to be under manual control, with more routine activities such as separation assurance being automated both on the ground and on board aircraft.
- In the flight control arena, cockpit automation began in 1912 with a two-axis autopilot developed and demonstrated by Lawrence Sperry. Flight control technology has now reached a highly-advanced state with the full authority fly-by-wire digital flight control systems on modern-day airliners, in which the pilot's role is largely that of a flight manager responsible for selecting the modes and commands to be executed by the flight control system. On some of the more advanced airliners, it is possible to auto-taxi to the runway, take off, cruise and land automatically, with little or no pilot intervention. However, it is interesting to note that even with such high levels of autonomy, the pilot is still expected to intervene and takeover control whenever the automation has difficulties in resolving ambiguous situations.
- Automation in air traffic management appears following a similar development pathway. Research over the past three decades have been focused on developing decision support systems for the controller, wherein the automation synthesizes advisories based on the sensor data, which the human controller decides to either discard or implement. Algorithms from the Systems & Control discipline form the basis for synthesizing the advisories. Techniques such as model-predictive control, linear and nonlinear programming algorithms, dynamic programming and advanced state estimation techniques are being employed for the synthesis of these advisories.
- After gaining adequate experience with this approach to graduated automation, down the road it is likely that human controllers will be relieved of certain lower level functions such as the separation assurance. The air traffic automation system will then form the "outer loop" around the flight control systems on

board individual aircraft to automatically meet most of the air traffic management objectives.

- The emergence of low-cost unmanned aircraft systems (UAS) is accelerating the trend towards automation, due to the sheer number of aircraft that will soon the airspace. This fact has prompted several industry experts to speculate that it is higher likely that extensive air traffic management automation may occur sooner than anticipated. Systems & Control technologies will be central to this transition.

It is difficult to imagine the modern world without aviation: it enables delivery of personnel and supplies for disaster relief around the world; it provides better coordination of human and material resources to manage global pandemics; it offers access to widely dispersed markets for perishable goods from all over the world; and it enhances opportunities for more frequent cultural interchange, improving understanding between peoples and nations. As population increases and the standard of living improves worldwide, aviation can become the main mode of human transportation and commerce. It is essential that this industry continues to grow and evolve, to enable much tighter integration of world economies and ensure continued prosperity of all humankind.

5.2. Automotive control

Basic concept

Depending on the traffic situation, at times, driving may be a pleasurable experience. Otherwise, in congestion or on long monotonous trips, it may become a source of fatigue and irritation. According to the 2013 European accident and safety report, approximately 1.2 million individuals worldwide perish in traffic accidents every year, with over 90% of accidents caused by human error. These figures could be significantly reduced if the human driver is helped by a control system which is free from inherent human biological limitations like drowsiness, fatigue and inability to focus on a task for prolonged periods of time.

One solution is the development of autonomously driven vehicles. The Eureka PROMETHEUS project (PROgraMme for a European Traffic of Highest Efficiency and Unprecedented Safety, 1987–1995) was one of the early demonstrators of the fully automated road vehicle. Since then, significant advancements in on-board computing capacity have enabled real-time image processing such as object/edge detection, recognition, classification and tracking. This presents a possibility to apply automated lateral (vehicle direction) and longitudinal (vehicle speed or acceleration) control for road vehicles (Shift2Rail Multi-Annual Action Plan 2015).

Control challenges

Maximum safety and minimum energy expense with comfort characterizes the definition of ideal human mobility. Advanced Driving Assistance Systems (ADAS) have been traditionally associated with driver/passenger comfort and safety enhancement. On the other hand, energy efficiency which is significant with respect to CO2 emissions and carbon footprint of mobility, has been largely addressed through Powertrain or Energy Management Systems (PMS or EMS) and not through ADAS.

As of today, the automobile "as a machine" is on the verge of reinventing itself, and its conversion to a fully automated mobile-robot may happen sooner than projected. In an Autonomously Driving Vehicle (ADV), the architecture must support simultaneous lateral and longitudinal autonomous control. Hence, the active safety systems (integrated into motion control) must co-operate with the powertrain control systems. In fact, the concept of ADV, where the driver himself is replaced, renders the original definition of driving assistance obsolete. Therefore, a completely new and expanded architecture is emerging where the advanced driving assistance systems fully co-operate with the powertrain management systems to generate a safe motion control vector for the vehicle as shown in Fig. 8.

In parallel, there has been high-level penetration of ICT in the automotive field resulting in development of automotive-specific data exchange platforms like Vehicle to Vehicle (V2V) and Vehicle to Infrastructure (V2I). Furthermore, it enables electronic mapping and navigational guidance through the eHorizon and Traffic control algorithms. This has resulted into advancement of a new approach towards transportation management coined Intelligent Transportation Systems (ITS). ITS facilitates the supply of relevant information to the vehicle motion control system so that intelligent decisions regarding safety (trajectory planning, safe headway, lane keeping, etc.) and efficiency (eco-driving, energy saving, etc.) (PReVENT Project official website. 2017) can be computed and implemented in real-time. From the scenario described above, it is clear that the ADAS of the future presents higher capabilities and more advanced characteristics than those implemented today. Hence it is important to revisit the defining properties of ADAS in order to establish the new capabilities, requirements and limitations for the ADAS of future. Automation is the key to realizing this improvement.

- **Safe and efficient mobility:** The new generation of ADAS is expected to improve both safety, and overall efficiency. While driver and pedestrian safety form an important part of human mobility (EU FP6, FP7 programs), it is also imperative to balance it with environmental safety. Thus, the newly coined term "eco-driving" comes into existence. Hence, ADAS having eco-driving capabilities are classified as Eco-Driving Assistance Systems (EDAS). Researchers have focused their efforts on coaching the driver to drive as efficiently as possible. However, for many individuals, driving style closely resembles their personality, and constant advice on driving style may often become irritating.

Theoretically, autonomous driving presents several advantages over human driving. For example, control hardware will never experience fatigue or be distracted from given tasks. However, safety is not the most significant improvement that autonomous driving can offer. Embedded processors are better suited for processing information relevant to future driving conditions and compute intelligent decisions quickly without compromising the given driving task. This is impossible for a human driver given the complex data and multitasking requirements demanded from him while he is already occupied with the basic driving task. Moreover, there are certain eco-driving manoeuvres like intelligent traffic signal to vehicle communication, efficient velocity trajectory decision and tracking, efficient acceleration control, intelligent intersection management, etc. that are better managed by the autonomous controller than by human intervention.

The modern car is being loaded with so much technology that more information than a human brain can process is being presented to it every instant through the human machine interface. This might, in a strange twist, lead to reduced safety. To balance safety and efficiency, a more logical solution may be to eliminate the driver from the equation and load him with an easy supervisory task while the autonomous car controls itself laterally and longitudinally.

Although most research related to the EDAS to-date has been concentrated in area of conventional Internal Combustion (IC) engine vehicles, the development of hybrid and electric vehicles has added another dimension to the capabilities of EDAS functions. The alliance of the dynamic programming method and ECMS (Equivalent Consumption Minimization Strategy) manage the compromise between IC and Electric engines of hybrid vehicles.

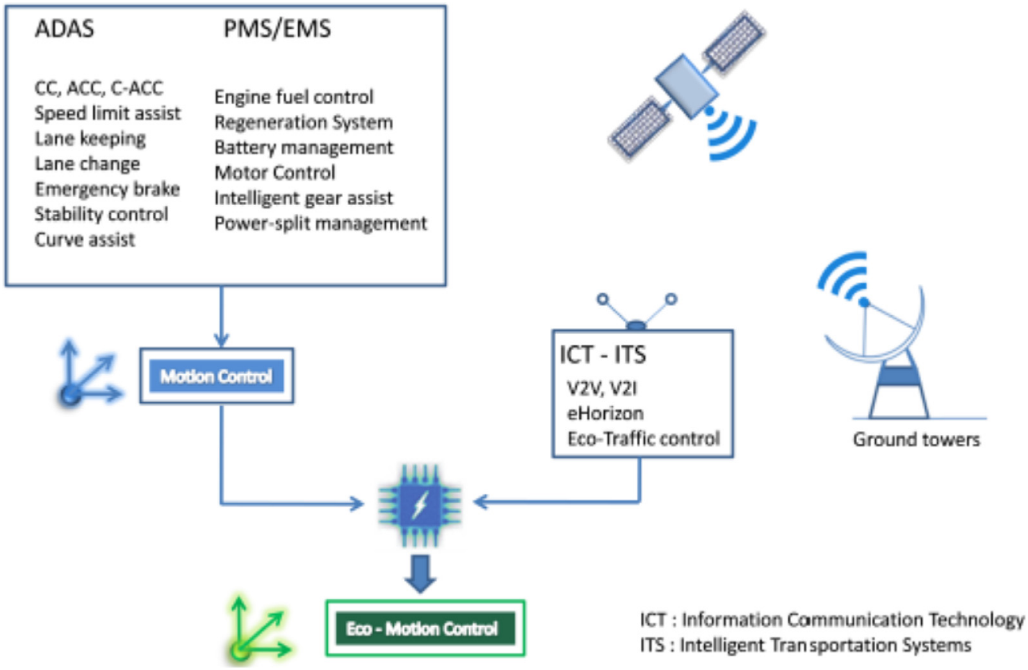

Fig. 8. ADAS, EMS/PMS and ITS co-operation.

Hybridization allows for part of the power requirement to be supplied from an electric source which is clean and distinctly efficient as compared to IC engine. Therefore, in any development of new generation EDAS functions, it becomes necessary to integrate the hybrid powertrain controller or energy management systems together with motion control systems. Moreover, the tradition of designing the EDAS as a separate system needs to be modified. The association of control and optimization methods in real-time will permit it.

- **Comfort and efficiency:** The EDAS must also consider the limitations of longitudinal and lateral acceleration, as well as the jerk limits thereby ensuring a smooth trip for the driver and passengers. Commercially available intelligent systems like highly-automated ACC (Advanced Cruise Control) are regulated by international standards which consider driver and passenger comfort characteristics such as acceleration/deceleration and jerk limitations. Another example is the Mercedes-Benz Intelligent Drive which goes one step further to ensure driver comfort with extensive vehicle system automation ranging from lighting systems to active chassis control.
- **Automation and component life:** Even with additional information about beyond-horizon traffic and road conditions, the effectiveness of EDAS is mainly characterized by the powertrain. A conventional IC engine powertrain offers far less flexibility to effectively use the ITS available data than the hybrid electric or even the pure electric powertrain. The main reason for this is that hybridization not only offers access to clean and efficient energy but also provides regeneration of the excess spent energy. However, the single most important component in the hybrid system is the battery which is not only prohibitively costly, but also very sensitive to operational deviations. Testing at Chalmers University (Sweden) suggests that the lithium-ion battery aging process is closely related to the loss of cyclable lithium. The loss of cyclable lithium is attributed to very high charge or discharge current which means that very high acceleration or high power demand from battery may damage it. Therefore, the EDAS must control the powertrain to operate within the allowable limits of such sensitive and expensive components.

New generation lithium batteries allow for a high charge acceptance without irreparable damage to their chemistry. It is normally defined in terms of time constraints as a 10 s or 15 s regeneration pulse. This phenomenon limits the time and in turn the amount of energy that can be harvested during the regenerative braking. Therefore, it is matter of complex optimization between saving regenerative energy and conserving the battery life. The modern EDAS must consider several factors in balancing the component life and preserving efficient operation while maintaining driver safety and comfort.

Automation and system architecture: The road towards a full vehicle automation is still quite long. The Society of Automotive Engineers (SAE) has defined the following six levels of driving automation:

1. No automation (driver alone)
2. Driving assistance (anti-lock braking system (ABS); electronic stability program (ESP))
3. Partial automation (longitudinal or lateral modes, Lane Keeping Assistance System (LKAS), Speed Control)
4. Conditional automation (Adaptive Cruise Control + LKAS, Park assist)
5. High automation (supervised)
6. Full automation (Google car)

These control strategies require driving to be shared between the driver and the automat to realize braking, engine or steering control. This implies the conception of new components and the decoupling of command and control with X-by-Wire systems (brake-by-wire with electro-hydraulic braking, steer-by-wire with backup steering column, redundancy of sensors and actuators). This new mechatronic architecture will help the development of new ADAS in addition to reducing vehicle mass and improving vehicle design. But electronic systems are prone to be inoperative, and in case of failure, criticality is high. In case of failure (braking or steering), could the system give back the hand to the driver?

To use this kind of architecture, failures must be diagnosed and automatically corrected in real-time. This uses the fault tolerant diagnosis techniques (fault detection and localization, strategies of accommodation). Several methods exist:

- Vehicle model-based fault detection (analytic redundancy)
- Independent braking/steering actuators (for trajectory control):
  - Differential braking in case of steering failure
  - Active steering in case of braking failure

In addition, the observer designs can help the diagnosis part to realize a complete supervisor.

In theory at least, automated systems promise to increase safety levels and passenger comfort. Therefore, preliminary research efforts were concentrated to achieve a high degree of confidence in lateral control of the vehicle. However, vehicle automation offers much more than just safety and comfort; it offers an efficient vehicle.

5.3. Control on railways

This section explains the operational and technological background to railways. It identifies currently anticipated developments, longer-term possibilities arising from Systems & Control, and concludes by suggesting related research challenges.

Attributes of Railways

Railways throughout much of the world have undergone substantial change in the last two decades, and now they are not only accepted as an essential part of national infrastructure but also contribute significantly to a sustainable future. Railways can be broadly classified as metros for cities, light rail for urban areas, regional railways for more rural areas and high-speed trains for intercity travel. All use narrow corridors of land to provide extremely safe, high-capacity transport, both for passengers and some types of freight. Many existing and almost all new railways use electric trains, and rails' "green credentials" are well founded, albeit this is under challenge from competing transport modes, especially in light of automobile improvements. However, unlike automobiles, electric trains are already grid-connected and therefore able to take advantage of the progressive decarbonisation of national energy supplies, which gives them a significant head start over electric road vehicles from a sustainability viewpoint.

Despite this unique selling point, in many countries the price to passengers and the cost to government, in terms of subsidy, are a disincentive to achieving a much larger market share, added to which many critical parts of existing networks are already operating close to maximum capacity. Also, passengers' expectations in terms of flexibility, quality and reliability of service inevitably rise as time progresses. Consequently, railways must face some key challenges in order to deliver their potential in the future. Europe in particular, has ambitious technical strategies aimed towards meeting these various challenges (Shift2Rail Multi-Annual Action Plan 2015), and the UK uses the concept of the "4 Cs" (Cost, Capacity, Carbon, Customer) to motivate R&D agendas (Railway Technical Strategy 2012).

Principles of rail traffic management

Rail traffic management involves both a safety layer and a management layer. The safety layer ensures safe separation of trains, traditionally by means of lineside signals protecting fixed blocks of track. Ensuring that only one train is in any block provides a safe stopping distance to a following train in the case of a problem with the preceding train, and the signals also protect the approach to track switches, junctions and stations. The principal challenge is the very high integrity level required to preserve the safety standard: railways have traditionally worked upon a "fail-safe" approach so that if anything goes wrong all the relevant trains are stopped, but this is generally not consistent with achieving high reliability.

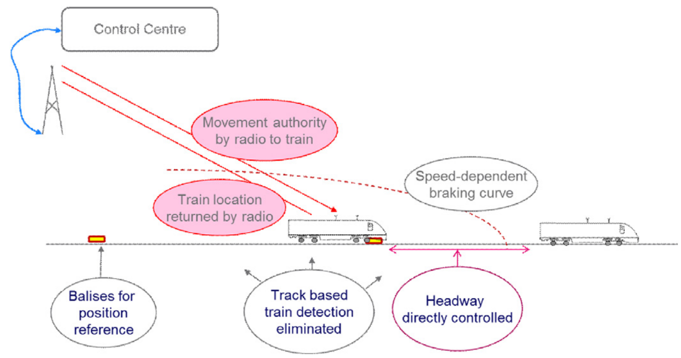

Fig. 9. Communication based train control.

The management layer involves an agreed operational timetable to determine normal operation, combined with automatic route-setting in localised regions around nodes (junctions, stations, etc.), by which common train movements are fully automated according to a predetermined schedule or script. Although control used to be localised in signal boxes, etc., nowadays overall control has progressively become centralised into control centres covering many kilometres of the surrounding network. The location of the trains is generally via track-based techniques such as track circuits and axle counters which send information directly to the control centre. The trains themselves are therefore essentially "dumb": train drivers have to obey running instructions from the lineside signals, although in some cases these are repeated within the driving cab.

The theoretical capacity for a plain line (e.g. trains per hour, or the "headway" times between trains) can be calculated, but as soon as there are any fixed nodes (principally stations and junctions) this theoretical level is not achievable, and in practice capacity is very difficult to quantify which means that designing of the operational timetable to accommodate the greatest number of trains is a complex process. Also, there is a trade-off between capacity and reliability: attempting to operate very close to maximum capacity means that even minor problems with the track or trains may create a large disruption as the effects propagate around the surrounding portions of the network (i.e. the railways' equivalent of congestion on roads). For this reason, railways typically aim to run at 70% of maximum capacity.

Current and anticipated developments in command and control

Direct capacity enhancement, as a result of more sophisticated train control, is a current target, essentially changing from the "fixed block" method based upon lineside signals mentioned above, towards a "moving block" approach in which the spacing between trains is directly controlled to a safe level. In fact, many modern, highly-intensive railways such as metros now use Communication Based Train Control which communicates parameters such as the exact position, speed, travel direction and braking distance via radio to the wayside equipment distributed along the line – see Fig. 9. These trains continuously receive information regarding the distance to the preceding train and are then able to adjust their safety distance accordingly. Other types of railways are progressing in this direction, e.g. via the European Train Control System. Train location must then become train- instead of track-based, using combinations of techniques such as GNSS with odometers and transponders; the derived location is then communicated to the control centre. These systems are significantly more complex than traditional systems, which means that the traditional "fail-safe" approach is not appropriate.

Developments are under way to improve disruption management. The ambition is to create highly reliable and resilient Communication, Command and Control systems offering network-wide traffic management capabilities for intelligent, predictive and adaptive operational control of train movements. These systems will track the precise location and status of every train on the network, and when data for all the trains' speed, acceleration, braking and load are available at all control centres, this will lead to improved operational decision-making. This is essentially real-time-optimised control of train movements to meet a variety of goals, meaning that perturbations can be resolved rapidly so that there is a minimum impact on the scheduled operation.

The previously-mentioned trade-off between capacity and reliability means improved signalling technology can either be used directly to increase capacity or to maintain capacity with higher operational reliability; conversely specific measures aimed at reliability improvement (for example, an improved ability to recover quickly and effectively from disruption), could also be used to enhance capacity while delivering the same level of overall operational reliability.

An obvious global trend is towards widespread automatic driving, which is already being exploited in railways that operate with very short headways, principally because human control is insufficiently precise in such circumstances. It is probably inevitable that automatic control will not only spread to all types of railway, but also expand in capability. Various "Grades of Automation" (GoA) are envisaged: GoA 1 describes current manual train operation where a train driver controls starting and stopping, operation of doors and handling of emergencies or sudden diversions; GoA 2 is semi-automatic train operation where stopping is automated, but a driver in the cab starts the train, operates the doors, drives the train if needed and handles emergencies; GoA 3 is driverless train operation where starting and stopping are automated but a train attendant operates the doors and drives the train in case of emergencies; GoA 4 is unattended train operation where starting and stopping, door operation and handling of emergencies are fully automated without any on-train staff.

Control technologies for rail vehicles

There are a number of possibilities for control engineering systems on railway vehicles, but here the focus is upon those that can have a major impact at the overall system level: principally power control and energy storage, which are related to improving energy efficiency; also mechatronic solutions for the suspensions and running gear, most immediately related to providing improved "track-friendliness" of the vehicles but also, as will be explained, potentially giving significant operational benefits.

Electric trains are already widely available and technologically mature, and for this reason are in an excellent position to take full advantage of progressive decarbonisation of the national energy generation facilities, and therefore well-placed to maintain their position as the most energy-efficient form of ground transportation.

Over the last two decades, developments in power converter technology have enabled a fundamental transition from DC to AC motors which are lighter, more compact, more efficient and more reliable. It has also led to the transition from locomotive-hauled trains to trains with distributed traction, i.e., where most or all vehicles have their own traction equipment.

The capability to provide braking using the electrical traction motors, and thereby recover the kinetic energy, is intrinsic, but full exploitation of regenerative braking relies upon the availability and receptivity of the trackside supply. Given that braking is a safety-critical function, all trains currently still require conventional friction brakes. However anticipated developments in energy storage technology will lead to devices that are usable on trains in terms of energy densities (both mass and volume). This will not only enable further improvements in energy efficiency, but also bring the possibility of wholly electrically-braked trains, i.e., no friction brakes. Energy storage combined with future developments in power converters will therefore provide the opportunity to optimise energy management by altering the flow of energy between the trackside power supply, the traction motors and the energy storage devices, including accommodating the need to both maintain the timetable and minimise energy via smart driving control – a challenging multi-objective control optimisation problem.

The idea of "design for control" has already been extensively employed in the aerospace and automotive industries and has proven highly beneficial. Aircraft and cars are now significantly different mechanically to what they were 40–50 years ago, whereas the conventional structure of rail vehicles (a car body suspended on two bogies (sometimes called trucks), each with two solid-axle wheelsets) is substantially unchanged. Mechatronic concepts for rail vehicles are now on the research agenda in a number of countries, and this will enable a fundamental re-think of some of the traditional design trade-offs, e.g. between running stability and performance around curves, and will lead to simpler, lighter, more energy-efficient and track-friendly vehicles.

Long term prospects and challenges

Beyond the current vision explained above, control offers other opportunities. High-integrity communication between vehicles enables the opportunity of trains operating in "flights," in which only the leading train is under centralised control and other trains follow their predecessor. It's also feasible to consider convoying, in which individual vehicles (or small trains) run close together, "virtually coupled" by means of bi-directional communication and electronic control of their traction and braking systems (much as envisaged for future automotive technologies). Fully autonomous trains/vehicles can also be conceived whereby centralised control becomes less and less important. For example, several trains negotiate directly amongst themselves on the approach to junctions based upon a variety of measures such as the number of passengers and their speed of approach. Too much autonomy is inconsistent with a train timetable, which is essential for passengers, so whether trains are running ahead or behind schedule would also need to be accommodated.

One of the very interesting long-term operational possibilities is mechatronic vehicles that control direction through junctions from on-board the vehicle, i.e. vehicle- instead of track-based switching. When this is combined with electronically- rather than mechanically-coupled trains, it is possible for individual vehicles to come together from different origins to form a train on the main intercity routes and then to diverge to different destinations. While very much a long-term idea, this could be the natural consequence of fully incorporating and exploiting mechatronics and control ideas. Importantly, this type of flexibility would be welcomed by future railway passengers and would undoubtedly help the competitive position of the railways compared with other transport modes.

Research challenges for Systems & Control in railways can be summarised as follows: further application of conventional control in new systems or sub-systems; real-time optimisation of train operation; determining the most appropriate level of autonomy between totally centralised and fully autonomous, i.e. another aspect of optimisation but in this case relating to the fundamental design of the Command, Control and Communication architecture. All these ideas must incorporate fault-tolerance, redundancy and model-based approaches to preserve the high safety standards that are a key attribute of today's railways.

5.4. Spacecraft control

Space systems represent a primary domain for applications of control theory and technology. At the present, optimal control has become an essential tool for spacecraft trajectory optimization, classical and modern feedback control techniques have been exploited for spacecraft attitude control and for vibration suppression in launch vehicles, while estimation algorithms, e.g., those based on Extended Kalman Filters, have been implemented to determine spacecraft states. Control will remain a critical enabler of current and future space missions. Spacecraft capabilities can be improved based on developments and advances in stability theory, optimal control, model predictive control, nonlinear control, control of time-varying systems, adaptive control, estimation theory, autonomous systems, time-delay systems, coordinated control, etc. Some directions for spacecraft control research (not an exhaustive list) are highlighted below.

- **Spacecraft orbital manoeuvring**: The traditional approach to spacecraft orbital manoeuvring exploits "open-loop" trajectories with occasional trajectory correction manoeuvres (TCM) providing a form of feedback. Spacecraft trajectory optimization techniques must address nonlinear, non-convex, and mixed-integer programming problems that are becoming more difficult with growing mission complexity. Such a complex mission may, for instance, involve a spacecraft carrying and releasing multiple probes which, exploiting thrust and gravity assist, must visit a maximum number of asteroids and return to the main spacecraft. The development of systematic feedback control approaches for orbital manoeuvring represents a promising research direction. Examples where such feedback solutions may be useful include manoeuvring near/landing on an asteroid/comet with uncertain gravitational field and outgassing pressure disturbances, relative motion manoeuvring in Low Earth Orbit (LEO) in the presence of disturbances such as air drag or thrust errors, and maintaining (open-loop unstable) Halo orbits near Lagrange points. Receding horizon/model predictive control can be exploited to provide feedback solutions that satisfy stringent actuation constraints (thrust or delta-v limits) and do not lead to substantially increased fuel consumption. Furthermore, stabilization and tracking techniques for time-varying systems can be applied to spacecraft control thereby exploiting time-varying linearization along non-circular orbits.
- **Spacecraft attitude control**: The time and effort for spacecraft characterization can be reduced by leveraging advances in adaptive control and on-line system identification. In particular, uncertainties in spacecraft inertia matrix, reaction wheel alignments, thruster generated forces and moments, can be handled through applications of these techniques. Further research can enable more effective and efficient use of attitude control actuators such as reaction wheels that suffer from zero speed crossing issues and Control Moment Gyros (CMG) which have kinematic singularities. To extend spacecraft life, attitude control techniques that can accommodate actuator or sensor failures are of increasing interest. Such techniques may need to exploit "higher order" physical effects, such as solar radiation pressure-induced torques, and approaches that take advantage of nonlinearities in spacecraft kinematics and dynamics through nonlinear control. Finally, the combined treatment of translational and rotational dynamics of spacecraft can be exploited for spacecraft relative motion control.
- **Control of multi-body and flexible spacecraft including formation control and space robotics:** Challenging nonlinear control problems emerge for multi-body spacecraft. Such spacecraft may consist of multiple rigid links connected by actuated joints or tethered links. Accounting for the effects of flexibility, as in antennas or booms, is also necessary in certain applications. When flying in formation, multiple spacecraft are connected through feedback laws and communication links rather than physical joints. Control objectives for formation flying spacecraft may include efficiently maintaining the desired shape while avoiding debris collisions. Advances in coordinated control and synchronization can be exploited for translational and attitude control of spacecraft formations (Fig. 10).

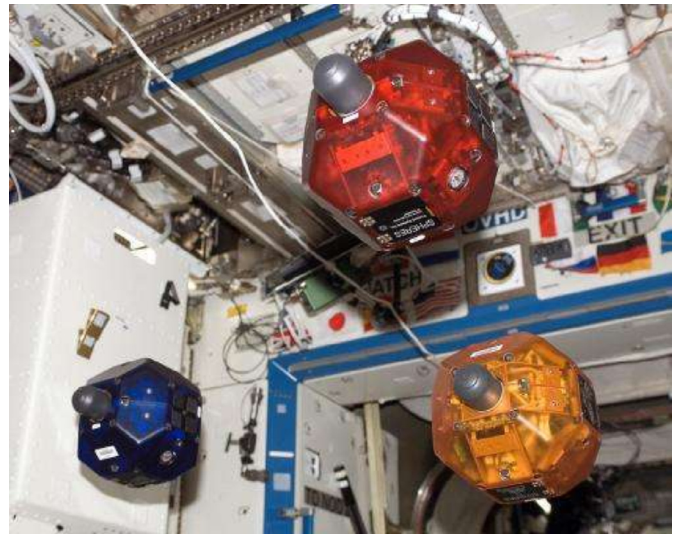

Fig. 10. Three satellites fly in formation as part of the Synchronized Position Hold, Engage, Reorient, Experimental Satellites (SPHERES) investigation. This image was taken aboard the International Space Station during Expedition 16 in the U.S. destiny laboratory module. Credit: NASA. Read more at: http://phys.org/news/2014-03-breakthrough-robotics-space-exploration.html#jCp.

- **Enabling and expanding spacecraft autonomy:** There is growing interest in increasing spacecraft autonomy and expanding spacecraft capabilities in the areas of automated reasoning, decision-making, intelligent handling of failure modes, on-board repair and re-planning of spacecraft missions. Increased spacecraft autonomy will improve robustness, extend life and maximize scientific value of spacecraft missions, especially when communication and software/command uploads from the ground are infeasible or impractical. The ability to solve constrained optimization and optimal control problems on-board in real-time is considered an enabler of spacecraft autonomy.
- **Spacecraft state estimation.** Challenges in estimating states of spacecraft (and of other objects in space) are numerous and include infrequent/intermittent measurements, weak observability of certain states, limited sensor accuracy or field of view, tracking many objects with a small number of sensors, etc. Many effective techniques for spacecraft state estimation have been developed, and this area of research has been as important and vibrant as spacecraft control itself. Additional opportunities exist to account for the unique challenges and characteristics of estimation problems in the spacecraft domain.
- **Spacecraft power management.** Spacecraft power management addresses optimization of electrical energy generation, consumption/load management, protection of components and life extension of components. Recent research activities in control of hybrid and electric vehicles, micro-grids and smart grids have also stimulated the growing interest in spacecraft power management problems. In particular, a fraction of power consumption that supports on-board computing and communication can be substantial and should be managed by a vari-

ety of approaches including "processor de-clocking" in the current spacecraft. The development of algorithms that maintain a low/reduced power consumption footprint while generating feasible control actions represents an interesting research topic.
- **Control allocation and plug-and-play spacecraft control.** The actuation configuration of large spacecraft, such as the International Space Station, may vary in time as some of the actuators may fail or propulsion system of visiting spacecraft may be used for orbit raising or attitude control manoeuvres. Spacecraft control systems can thus benefit from control allocation solutions that can optimally manage time-varying actuation configurations and provide plug-n-play functionality to such actuators.
- **Control problems for small spacecraft.** Small satellites (small sats and cube sats) are increasingly considered for a variety of missions. Control solutions for such spacecraft must accommodate their relatively small actuation authority, stringent constraints, cost optimized hardware and relatively large influence of disturbances.
- **Validation and verification.** Further research into effective validation and verification techniques for spacecraft control systems and software will facilitate the integration of advanced control technology into spacecraft applications and alleviate current conservatism in replacing legacy spacecraft control solutions.

5.5. The future of marine automation

There have recently been two reports on trends in the marine industry in general (Lloyd's Register, QinetiQ, & University of Strathclyde 2013), and shipping in particular (Hodne, 2014). We will here try to summarize some of the main findings in these two publications, and add some more specific comments related to the future of control in marine. However, for more details we strongly recommend the reader consult these two excellent reports.

The shipping industry moves about 80% of world trade volume, making it a very important part of the global economy. It is an extremely efficient means of transport which has the lowest carbon footprint per ton of cargo. That does not mean it is without challenges, however. The keyword for the future is sustainability. Even with a low environmental impact, there is still room for substantial reduction of CO2 emissions. There is also the constant pressure to reduce the freight costs.

Finally, safety is a major concern. The number of fatalities at sea is currently very low. However, according to the Statistics Portal, after peaking at 445 in 2010, there were still 245 pirate attacks reported in 2014 causing a staggering $12 billion estimated cost to the shipping companies. Hence, improving safety for all kinds of ships – cargo as well as passenger – is also very high on the agenda.

Below is a brief overview of some of the trends (Fig. 11) that are expected to have a major impact on the future of marine automation:

- **Electrification.** There is already a trend towards increased use of electric propulsion in ships In which diesel engines are no longer mechanically connected to the propeller but merely driving the generators used to produce the needed electricity for propulsion as well as all other electrical needs of the ship. Hence, it is expected that electric propulsion will be common place for many ship types by 2020. Furthermore, we have already seen the first applications of direct current (DC) grids on vessels which allows generators to operate using variable speeds, enabling optimal fuel consumption. Moreover, ship power management will be enhanced by the introduction of energy storage solutions. Due to cost, life cycle and size, electricity storage in batteries currently has few marine applications. But with prices expected to fall, we could see larger vessels incorporating them as part of a hybrid power solution. In addition to hybrid ships, we have already seen the first full electric ferries, and we expect full electric solutions will become more and more common in short sea shipping.
- **Alternative energy sources.** Increased demand for reduced CO2 emissions will be a strong driver towards the use of new fuel and energy sources. Current dependency on heavy oils is clearly not sustainable. There will be increased use of alternative fuels: first, mainly liquid natural gas (LNG), but later, different types of biofuels as well as fuel cells. At the same time, developments in energy harvesting and recovery are expected to drastically improve energy efficiency. Energy can be harvested from thermal, solar, wind and mechanical energy sources and stored for later use. As an example, [35] describes auxiliary wind propulsion using a rotary sail solution which, based on the first sea trials stared in late 2014, shows a potential 20% fuel savings.

  Also, utilising the waste energy of a power production system can drastically improve a ship's energy efficiency. The most common method of recovering energy today is waste heat recovery systems, but low temperature recovery systems are expected to become available.w

- **New ship design.** Energy efficiency will also drive ship design evolution. At present, Model-Based Systems Engineering (MBSE) is used in the design of ship machinery and is a key enabler for the introduction of new technologies. In the future, MBSE will expand in scope to include other areas of ship design, such as structural and hydrodynamic elements. At the same time, developments in computational fluid dynamics (CFD) provide great potential for ship design. In the future, CFD will be able to simulate complete scenarios including varying wind and wave conditions.
- **Digitalization.** Parallel to the expected advances in ship design, we will also enter the era of virtual commissioning, where whole ship system models will be built using subsystem simulation models. Every component or device will have a "digital twin," i.e., a software model ready to be incorporated in either a pure computer simulation or a hardware-in-the-loop simulation thus enabling the entire control solution to be tested in simulation prior to commissioning. Digital twin could also be a full ship (or some component or system) model that is run parallel to the actual ship, perhaps in the cloud.
- **Cloud and IoT.** Ongoing digitalization will also mean that all vessels, as well as their individual devices, will be directly connected to the internet in the future. Certain monitoring functions are already implemented as cloud solutions, but with increased bandwidth, the sheer volume of information available regarding individual vessels, as well as entire fleets, will enable ship owners to make better-informed decisions. A first step could be to have a local cloud on-board the ship with all devices connected to the ship intranet.
- **Optimal system-level operation**. Ship digitalization enables development of automatic optimization solutions that rely on continuous weather and sea forecast information. Currently, several decision support and advisory systems provide the onboard and onshore crew assistance in improving ship or subsystem operation efficiency. It is expected that there will be more and more closed-loop optimization solutions where, for example, weather routing and speed optimization are directly connected in an autopilot system enabling the ship to run at optimal speed and route given the latest weather and sea forecasts. Similarly, operation planning and forecast information can be

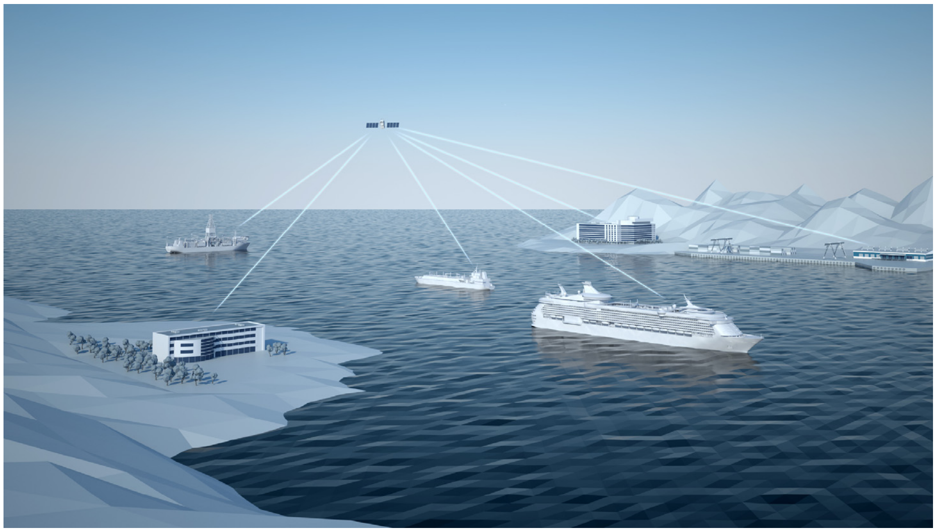

Fig. 11. Connectivity, modern communication protocols and cloud solutions enable continuous monitoring and remote control of ships.

used to optimize the ship's energy consumption and production. It is expected that the ship's systems will be more connected and operated more optimally from the perspective of total efficiency.

- **Autonomy and remote operation.** Following the trend of the railway, automotive and airline industries, marine is also moving towards autonomous and remote operations. The main drivers for this development are costs, efficiency and safety. By 2025, we expect to see remotely controlled, unmanned coastal vessels. Some ship operators already utilise remote operation centres which provide real-time views on the whole fleet of vessels. It is expected that more decision-making and operation planning and even manual control will move to remote operation centres [36]. In addition to ship operators, system and component makers have remote operations centres to manage their own installed base and provide assistance and real-time support for the ship crew.
- **Optimization of logistic chains**. Connecting land, port and marine operations in the same logistic optimization problem will offer significant opportunities for increasing the efficiency of the logistic chain. In this way, one is able to optimize, e.g. the container transportation from the first sender to the last receiver.
- **Sea traffic control.** Closer to ports, the industry may also benefit from a traffic control system, modelled on systems now in use by the aviation industry. Such a Sea Traffic Control (STC) system could provide clearance through congested areas on assigned routings, or provide recommendations to alter heading and speed when appropriate. The STC system would be responsible for helping vessels maintain a safe distance from land and other marine traffic.
- **Increasing arctic traffic.** Global warming will increase the use of North-East Passage for cargo traffic between Europe and Asia. Operation in arctic conditions requires ships which are designed for arctic operation. In addition, the need for new type of ice management, forecasting and routing services will increase. In addition to dry cargo, also oil and gas transportation and drilling in arctic conditions is expected to increase.
- **Robustness, fault tolerance and multi-objectivity.** The trend towards remotely operated and autonomous ships will require robust solutions on various levels of operation. Operation planning systems (e.g. weather routing) ust account for the uncertainty in weather and sea forecasts, as well as the current condition of on-board equipment to decide the optimal operation plan. In addition, as most ship operation optimization problems are multi-objective in nature, it is expected that multi-objective methods where the operator can easily compare various Pareto optimal solutions become more common in the ship operation.

The presentation above primarily addresses civil marine vessels, but of course many of the trends apply to naval vessels. However, the naval sector is probably the least predictable of all marine sectors.

In summary despite being a very old and seemingly mature business, the shipping industry will go through fundamental changes in the coming decades. In this process, control will play a crucial role as an enabling technology.

5.6. Renewable energy and smart grid

5.6.1. Mixed AC/DC for renewable energy sources integration – from microgrids to supergrids

Electric power networks are among the largest and most complex man-made systems. They connect hundreds of millions of producers and consumers, cover continents and exhibit very complicated behaviours. The transition towards a low-carbon economy leads to an even more complex system of systems; as a result, there are still many poorly understood phenomena caused by the interaction of such a large number of devices and the large spatial dimensions. Currently, major changes to the grid structure are being implemented, in particular to support the large-scale introduction of renewable energy sources (renewables), such as wind farms and solar plants, to reduce CO2 emissions. Two crucial features of these electric power sources must be addressed: most renewable sources are small generating units, dispersed over a wide geographical area; and their primary energies (wind, sun, tides) are by nature not controllable and fluctuate over time. Integration of these highly intermittent, randomly variable and spatially distributed resources calls for new approaches to power system operation and control.

One possibility for the massive integration of renewables is the use of Direct Current Networks (known as Multi Terminal High Voltage Direct Current MT-HVDC), in support of the current AC (Alternative Current) grids. Unfortunately, MT-HVDC grids are yet not a reality, even though significant research efforts have been devoted worldwide to this field. One of the reasons is DC (Direct Current) lacks a global variable represented by frequency, as in the case of AC, to help in estimating locally the required power to balance the system. Other more important factors include the difficulties in designing DC breakers and DC/DC transformers, as well as their associated huge cost. Furthermore, DC grids' dynamics may be much faster than AC ones. As a consequence, their stabilization may be much more delicate. In addition, since DC networks connect to AC grids, it is important to study their interactions and develop control mechanisms that allow DC grids to bring support (called Ancillary Services) to the main AC grids. This is a very complex topic because of the mixed time scales and very different dynamics and concepts between AC and DC.

Recently, point-to-point HVDC transmission lines are used increasingly in high-voltage electricity grids, an evolution that will continue in the future. HVDC is already the transmission technology of choice for submarine interconnectors, but it is also increasingly used for on-land interconnection lines (in particular, underground cables). As offshore wind farms are constructed farther from shore, it becomes more attractive to connect them to the main grid through HVDC cables. The first offshore wind farms connected through HVDC are in operation, but they are not very reliable yet (Borwin 1 is facing problems of harmonic resonances leading to wind power curtailment!), and many more are being constructed or planned. But still many problems are to be solved in order to construct multi-terminal DC grids.

Future power networks, already called SuperGrids, will be mixed AC/DC grids. These SuperGrids will connect offshore renewables to the main AC grid, as well as reinforce these main grids. They are meant to be backbones for the power system, interconnecting different electric regions (synchronized or not). Energy storage could also be connected to the grid and bring flexibility and reliability to these grids. They will mitigate the effects of intermittence and variability from renewables. Furthermore, DC allows much longer lines, and such grids will be better fit to integrate the wide geographical dispersion of renewables. This will also help counteract variability, since the probability of having sun or wind in each given moment is much higher if one has a larger geographic dispersion.

Such a vision will lead to a power system of increased complexity, where many converters are connected to the AC power system. This will ultimately bring an important reduction of inertia, since current AC grids are mainly composed of very large rotating masses (power plants) with very large inertia, which mitigates the effects of the disturbances that affect the system, which gives time for control systems to counteract them. These large rotating masses are currently being replaced by renewables connected through power electronics, which have no inertia at all. On the other hand, these new power electronic and other ICT (Information and Communication Technology) devices open new possibilities for control and will ultimately be the tools for stabilizing these new smart grids, created by the application of ICT and modern control methods to the electric power grid. This increase of power electronics can bring up an extreme case where all power sources are composed by converters. This is indeed the case in some configurations of wind farms connected through HVDC to the mainland and some MicroGrids. This case, where no rotating source is providing inertia, becomes a completely new (and yet open) problem in which frequency does not represent power balance anymore. There is no "global" variable to indicate to each source how much power it must feed to the system in order to keep in balance. Even though synchroconverters (converters that mime synchronous machines) can attenuate this problem, this is certainly not the optimal solution, and the solution is still widely open

Even beyond, we can easily envisage the use of MT-DC for other situations where it is well fit. For example, in medium or low voltage situations, DC is already the electric power inside photovoltaics, batteries and electric vehicles. It is very natural to propose the use of medium voltage DC networks interconnecting photovoltaic plants and storage, towards one or several points in the AC grid, electric vehicles and several other loads. In the same way, a large number of individual photovoltaic arrays (household individual panels in a neighborhood for example) could be interconnected by a DC grid, mutualizing storage devices like batteries, super-capacitors and fuel cells, and loads like households and electric vehicles. This would constitute a DC MicroGrid to be connected to the local AC one. Such mixed AC/DC MicroGrids could be very important to address the integration of renewables at the distribution system's level. In this way, the electric grid would become a constellation of MicroGrids interconnected by a main SuperGrid. The main SuperGrid would deal with each MicroGrid's production/consumption customers, where energy could be exchanged at a real-time power market. Here again, ICT and control in particular, is key to stabilize such complex smart grids.

5.6.2. Wind and solar energy systems

Wind flow and solar radiation and are main sources of energy offered by nature to human beings. In the 21st century, their employment in industrial power systems is more and more realistic, with a non-negligible share of total produced power.

Control in wind energy systems

Wind energy has evolved rapidly over the past three decades and continues to contribute to the power grid around the world. Control is critical for reliable and economic operation of modern wind power systems. In fact, it plays an important role in almost all sectors of wind industries at all stages of development ranging from individual wind turbines (WTs) to wind farms, from installation on land to offshore, from operation at ground level to high altitude. Various control techniques are being developed to achieve targets such as high-capacity factors from available resources, efficient and safe operation under uncertain environment, increased lifetime and efficiency of large offshore wind farms, reduced operation and maintenance (O&M) cost, enhanced voltage and frequency stability for grid connection, etc. At a higher level of wind power production, modelling, control and optimisation also contribute to reducing environmental impact, mitigating land use conflicts and achieving optimised energy policy. Modern wind power production has been driven towards the scale of multi-megawatt (MW) turbines operating in large wind farms. The trend of increasing size raises new challenges to control system development, for example, the role of controllers has been extended to include the alleviation of loads. Control focus has been moved from individual machines to arrays of turbines, where whole-farm modelling is required that describes the loads and the time-dependent interaction between turbines imposed by the wakes and the overlying wind flow. Wind farm-level controllers, acting through power adjustment among turbines, need to be developed to achieve the long-term objectives of providing ancillary services to power networks, maximising energy capture and minimising O&M costs.

Developing new generation wind turbines to produce more energy at lower cost with low maintenance is always a challenging topic in wind industry. This ambitious goal can only be achieved by an integrated approach to the design and control of wind turbine systems. One such recent exploration is to create new and better WTs by harnessing the benefits of variable stiffness materials for aeroelastic tailoring, for which servo-aeroelastic — rather than just aeroelastic — principles need to be exploited by including aspects of control and electro-machinery.

Airborne wind energy (AWE) systems is another novel design concept that seeks to attain operating altitudes (200 m – 1000 m) where there is a greater on-average wind resource and, by consequence, a greater potential power output compared with similarly sized horizontal axis wind turbines. Technically, factors such as wind speed distribution, turbulence and extreme winds, temper-

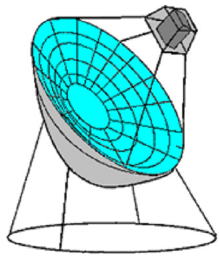

Fig. 12. Dish stirling solar plant.

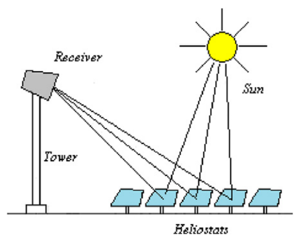

Fig. 13. Solar power tower.

ature and weather conditions (thunderstorms, hail, tornados, hurricanes, icing, etc.) must be considered in the control system design. Therefore, comprehensive approaches to safe and high-efficiency operation of AWE systems are needed to meet the diverse and ambitious requirements.

Control in solar energy systems

Sunlight can be converted into electricity mainly by photovoltaic (PV) or thermodynamic solar power (TSP) plants. In PV, the light is directly transformed into electricity using the photoelectric effect, producing a direct current. Such plants have no moving parts and are noiseless. One of the main control problems is extracting the maximum power available in relation to the solar radiation. This is known as MPPT (Maximum Power Point Tracking). PV technology is growing at 30% a year, thanks to governmental aids and cost reductions. TSP plants focus a large area of sunlight into a concentrated heat receiver. These plants make use of mirrors as concentrating tools, with different shapes and different arrangements. The receiver is a heat exchanger where the solar radiation is transferred to a fluid.

Among the concentrating technologies, the following ones are often used:

a) **Parabolic trough**, in which a linear parabolic mirror focuses the solar rays on a pipe located in the focal line. A heat transfer fluid is heated as it flows along the pipe and is then used in a thermodynamic cycle in order to produce electricity,
b) **Dish Stirling solar plant** (Fig. 12), where a parabolic mirror focuses the incoming sun rays towards a receiver acting as a thermal source for a thermo-dynamical machine based on the Stirling cycle (two isovolumic and two isothermal transforms),
c) **Solar power tower** (Fig. 13), characterized by an array of heliostats, each of which points to the top of the tower where the solar radiation is converted in heat flux. To this purpose, the heliostats have to be suitably controlled.

In order to address efficiency, reliability and lifetime in large-scale or new generation energy production systems, wind and solar energy systems' control requires joint efforts of researchers from different backgrounds including energy meteorology, aerodynamics and aeroelasticity, fatigue and structural mechanics, optics, and of course Systems & Control.

5.6.3. The energy internet

The electric power grid is experiencing dramatic changes, and the future grid is yet to be defined. The past decade has seen unprecedented growth of renewable energy with significant new capacity being continuously added in major grids around the world. We are experiencing the changeover from the traditional model of centralised generation to a mix of centralised and distributed generation in the power grid. The renewables bring with them their own challenges of intermittency. These problems are only worsened by the recent increases in the adoption of DC loads, such as electric vehicles and their associated vehicle-to-grid potential, small electronics, LED lighting, flat screen televisions and computing. Each of these devices incorporates advanced power electronics that inject small disturbances into the electric utility grid. Plug-in electric vehicles will offer challenges but also opportunities, thanks to the possibility to exploit their energy storage capabilities. The pace of this change is not predicted to slow, however, without the ability to effectively manage these intermittent resources, our aging electric utility grid is hurtling towards a series of challenges of epic proportions.

While today's grid model is based on uni-directional power and information flow, the future grid will have bi-directional flow of both power and information with new generation and new loads. While the premise for current grid control is open loop control of power flow between sources and loads, the distributed generation and new loads cannot be reliably and economically integrated into the new grid without the research and development of sophisticated distributed controls for the future grid. This is a challenge as well an as an opportunity, since distributed generation has the capability to ensure the desired level of resiliency, reliability and security in the future grid. Establishing a cyber-physical infrastructure (Fig. 14) with a critical amount of storage in the system and information flow among the distributed controllers that provides ubiquitous sensing and actuation will be vital to achieving the responsiveness needed for future grid operations. Sensing and actuation will be pointless without appropriate control laws that enable operators on the energy market to optimally manage power flows, maximizing profits (in particular through the use of renewables) under setpoint regulation at the nodes, and power and energy balancing constraints.

The FREEDM System concept is one approach of addressing the challenge of effectively managing the new loads and generation using the key elements of Energy Cell, the Energy Router, and Isolation Device. The Energy Cell consists of load, generation and storage which resides at the very end of the distribution system. The Energy Router is a controllable transformer envisioned to be made of power electronic devices, while the Isolation Devices are in place for rapid reconfiguration of the system in case of faults or failures in certain segments. Both the controllable transformer and the isolation device are digital devices that can store data and communicate with their peers for making control decisions. The basic construct of this cyber-physical system is shown in Fig. 14. The controllable transformer is also capable of managing both AC and DC Energy Cells in the LV side of the distribution system. The FREEDM System allows the Energy Cell to be replicated across the utility distribution system. These energy cells are essentially dispatchable and will work together to provide the desired level of reliability, security, and seamless integration of distributed renewable resources, loads and energy storage. These elements together form what we call the Energy Internet, within which Energy Cells can participate in the energy market based on economic, social, environmental and security considerations. The distributed control algorithms needed are yet to be developed, but are essential for automated and robust power, energy and fault management. The algorithms to be developed should be generally applicable for any MicroGrid system.

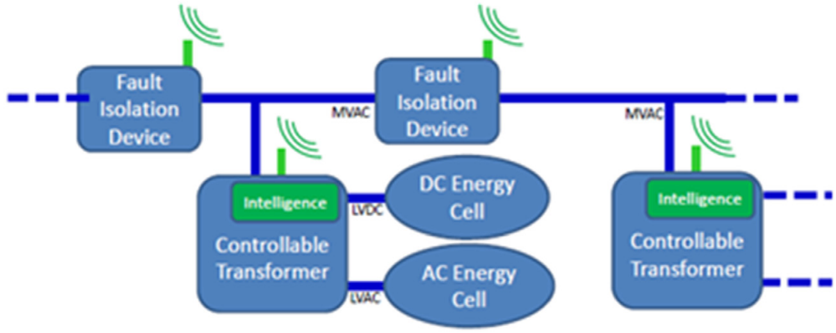

Fig. 14. Power cyber-physical infrastructure.

5.7. Energy and resource efficiency in production systems

Worldwide, 51% of the total global energy consumption occurs in the industrial sector, and 90% of that happens in manufacturing. With appropriate optimization techniques, a reduction of at least 50% in energy consumption may be possible. Techniques such as selective actuation of devices (e.g., turning components off when not in use), reducing idle production time (which also increases efficiency), as well as process planning and scheduling based on energy and resource efficiency are all useful to help achieve these significant energy savings. On the other hand, when machines are power-cycled frequently, they often require more maintenance, which adds cost and reduces efficiency.

Generally speaking, manufacturing, i.e., production of goods that have a weight, shape and other physical properties to fulfil their functions, contributes about 22% of the European Gross Domestic Product and 12% of US GDP. In Europe for instance, 70% of all jobs are directly or indirectly dependent on manufacturing activities. Manufacturing can be divided into two types of activities: *discrete manufacturing*, which produces investment goods and consumer products and their parts and components; and *process industries*, such as the chemical, steel, glass, ceramic and metal industries. Despite increasing pressure and tight regulations, the manufacturing industry has remained competitive in most areas of Europe, thanks to continuous innovation and a high level of automation. Industry as a whole, contributes 20–25% to the final energy use in Europe, with the larger part consumed by the process industries. Energy and raw material use reduction is of prime concern in the process industries and represents a major contribution to a greener and sustainable society in Europe. In addition to improvements in the construction and configuration of the production equipment (e.g., better insulation, heat recovery), control plays a major role in energy saving. As an example, crude oil is separated into fractions for final use (gasoline, diesel fuel, kerosene, naphtha, etc.) mainly by distillation. In distillation columns, a large amount of thermal energy is used to boil up the liquid: the higher the purity required of the products, the higher the consumption of energy. If these columns are operated automatically under continuous feedback control, the variability of product purity is reduced and operation can be closer to the specifications, leading to significant energy savings. Advanced configurations as, e.g., divided-wall columns and the integration of chemical reaction and separation, offer further potential for energy savings but pose challenging operation and control problems.

The energy usage of discrete-part manufacturing such as machining is also significant. Thus, there is a significant opportunity for energy savings. Automatic control methods can optimize the production schedule to maximize energy savings.

The extension of the system view and model-based decision-making methodology will lead to a new generation of functionalities in the MES (Manufacturing Execution System) layer. Now, most of the applications at this level are related to production management and process supervision, but new areas such as plant-wide energy management, production flow-adaptation to changing conditions, for instance by on-line re-scheduling and explicit consideration of the plant-wide dynamics, will significantly improve the performance and efficiency of today's production plants and factories.

Raw material processing industries are fundamental and unreplaceable in their role in social and economic developments worldwide. At present, China has become the largest country in this industry sector with complete and large-scale production infrastructure. To maximize the use of resources, it is imperative that low-grade raw materials with large variations should be employed in production. In this context, the operation of some industrial processes exhibit complexity in terms of variable dynamic characteristics, strong nonlinearities, heavy coupling, unclear mechanisms, mathematically un-modelable and online un-measurable key parameters. This constitutes challenges to existing control theory and technology. It should be noted that such a situation will not be unique for China. Indeed, with the increased unavailability of quality raw materials, this will become a global problem in future years.

Indeed, the above raw material processing industries are part of typical process industries, and they are clearly different from machinery industries such as discrete manufacturing. To enable the Systems & Control discipline to play a key role in the optimized smart manufacturing for the processing industries, it is imperative to promote the following research areas:

- **Theory and technology of intelligent optimal control systems for major manufacturing utilities.** In this context, research needs to be carried out on intelligent sensing for production variations, active and self-decision making for control systems and automation systems. It is expected that when an abnormal condition presents, such a system should be able to predict, self-recover and repair the abnormal condition in real time. Indeed, research into safe operation and optimal operational control theory and technology needs to be conducted so as to realize maximized product quality and efficiency and minimized consumptions of energy and raw material costs during the production phase.
- **Theory and technology of big-data driven intelligent optimal decision making systems for enterprises operational and production processes**. It is necessary to develop new decision making systems that can automatically and effectively acquire data and information on market variations and resource properties in response to variations of market needs and raw materials. Such a decision-making system can intelligently measure the state of material, energy and information flows with active

learning and responsive capabilities so as to adaptively optimize the decision making in terms of production and relevant performance indexes. It is also expected that such a decision-making system can optimally arrange the resources and the recycling of consumed energy so that the realization of production planning, scheduling and plant-wide control can be optimized together with relevant technologies.

- **Theory and technology of data, knowledge and model driven intelligent optimal decision systems for production structure.** This area needs to focus on the study of intelligent modelling and digitalization of process products and its production phases, so as to explore the dynamic performance analysis and visualization of the complicated interactions between energy and information flows together with effective experimental systems for production structure and systems.

5.8. Controlling water distribution networks

The World Water Group 2030 report indicates that the world is headed towards significant economic water scarcity unless we manage our fresh water resources better. Water for food, or irrigation water represents, 70% of the world's fresh water usage.

Given the typical size of irrigation water distribution networks, some non-trivial control engineering problems emerge. Periodic sampling of actuators and sensors is not a sensible approach. Most of the time there is simply not enough dynamic variation to warrant this, and when things change, sampling has to be relatively fast to capture the essence of the transients, e.g. to route flood waters and avoid flood damage. Measuring and actuating by exception, adaptively as the circumstances require, is more economic and will suffice. Equally, because of area and distances involved, a fully centralised controller that schedules an entire water distribution system is not advisable. Despite the fact that it may yield the "best" solution, the required communication infrastructure and bandwidth becomes prohibitively expensive. Minimal communication, again scheduled by exception, improves reliability and brings costs down. From a control point of view, overall system stability is typically relatively easily ensured. On the other hand, chain-stability or the amplification of disturbance-induced responses must be addressed. Chain-stability requires one to limit the gain from a disturbance to (a distant) control action. This poses the more challenging control objective, especially for large-scale networks.

More practically, dimensioning actuators and limiting control actuator interventions (especially to remain within the limits of locally harvested solar energy as reticulated power is typically not available across most of the irrigation system), and yet to achieve expected performance, represents interesting engineering design problems.

Also, cybersecurity is important because water networks are critical infrastructure. This is more difficult when the control is performed over wireless communication networks. In this instance interference, which may be malicious as well as naturally occurring, cannot be avoided. Hardware and software enabled redundancy play a critical role to achieve expected performance despite interference. Much remains to be explored in this area to achieve secure and reliable network performance even when the system experiences a cyber-attack.

- **Potential for impact.** Testimonials of commercial and fully automated irrigation districts report overall conveyance efficiencies near 90% (up from 70%) and on-farm water productivity gains of close to 100% (double the harvest with the same amount of water, largely due to precision timing in response to crop needs). In principle, this goes a long way towards making water use more sustainable. A schematic of the system is displayed in Fig. 15.
- **Beyond irrigation systems.** It should be observed that water-use efficiency in industrial or urban water settings is not much different from water use in irrigation whose use efficiency is also is less than 50%. Although the urban and industrial water only affects 30% of the overall water use, the associated economic impact is large. Simply contemplate the consequences of a city running out of water! Significant economic gains can be pursued through industrial and urban water use efficiency. Important in this context is that a proven reduction in overall water use may provide the licence to operate, whereas without proof of appropriate efficiency gains, the water could not be made available. Not all water use efficiencies require much additional technology. Rather they may be achieved through behavioural changes. The city of Melbourne, Australia, demonstrated as much during the last draught period 1995–2009. In this period, Melbourne better than halved its daily water usage simply by enforcing a number of water use restrictions.

Nevertheless, automating the management of pressurised water systems can yield significant benefits. Unfortunately, retrofitting existing urban infrastructure with the necessary information infrastructure to enable an automated operational regime is much harder than for open channel systems. In the first place, infrastructure access in a sprawling city is at best hard and always expensive. Secondly, pressurised water systems have much faster dynamics and much less storage, making measurement and control hardware more expensive, and in general such systems pose much harder to attain control objectives. Loss estimation, maintenance scheduling and improved water services in terms of quality monitoring, and service security are important and carry significant economic and ecological impacts that deserve the attention of the control community.

At the scale of an entire catchment, the natural scale on which the world should manage its water resources, the combined technical and economic management challenge is to identify the sustainably available water resource limit as a function of time and to implement an exploitation regime that ensures the community lives within this limit. This receding horizon planning and management problem remains unsolved. As for impact, managing water at the catchment scale is an essential component in addressing food security, as well as the millennium goal of ensuring people have access to sufficient water for sanitation purposes. Presently, as noted in the most recent World Water Report, more than a billion people experience severe water shortage.

5.9. Dynamics in neuroscience

Current State and Need for the future

Theoretical and computational neurosciences are witnessing spectacular growth. Theoretical neurosciences include formalisms inspired from mathematics and physics. They are generally distinguished from applied computational neurosciences, which are more data-driven than theory-driven. However, their object still concerns theoreticians, through the identification of generic principles of neural computation (integration and plasticity), allowing in some cases to reach a level of abstraction independent of the specific biological substrate. These neurally inspired principles of computation can then be transposed with profit to scientific fields other than neuroscience.

The larger impact of engineering methods and applied mathematics in the field of neuroscience may be explained by the combined effects of the rapid increase in high-performance computing and large-scale simulations, and the development of experimental techniques (often combining imaging and electrophysiology

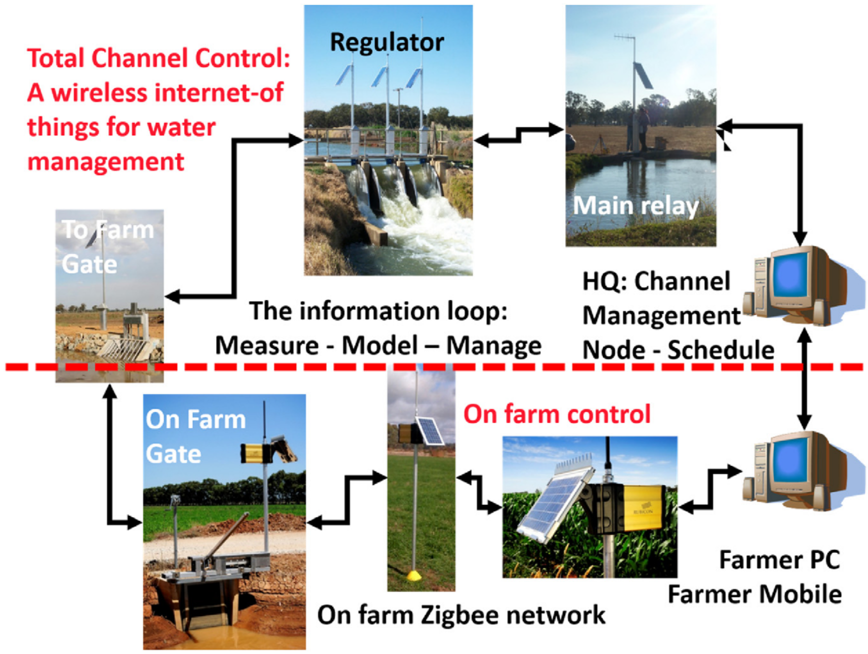

Fig. 15. Schematic of an automated irrigation system using TCC® developed by Rubicon Water and the University of Melbourne.

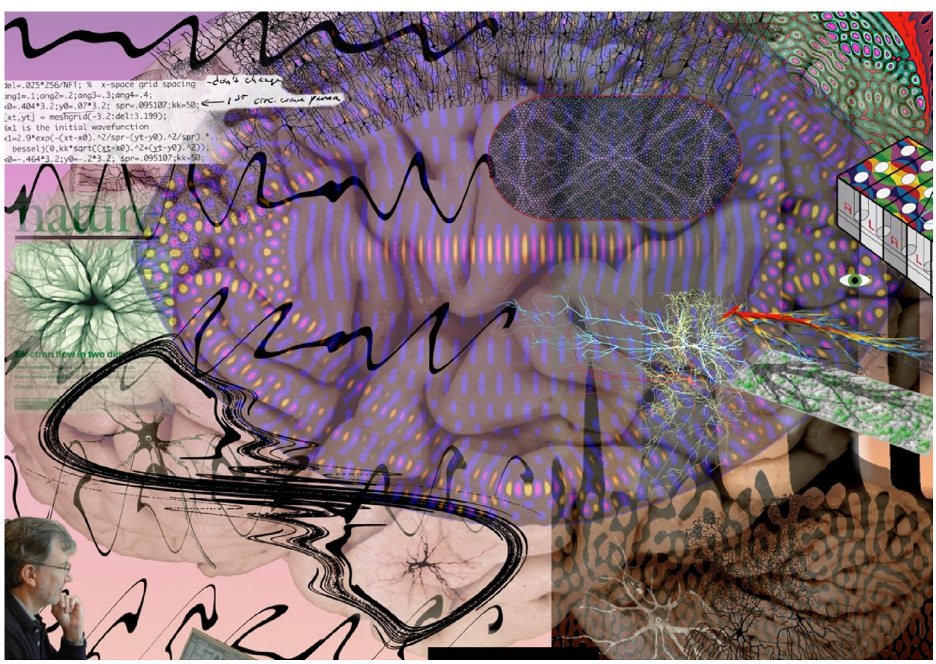

Fig. 16. Theoretical approaches of Brain dynamics. Collage, Eb Fetz (with permission).

such as two-photon) that allow and simplify the acquisition of large amounts of data. Over the past 15 years, significant advances have been obtained in our understanding of the brain (Fig. 16) from intensive collaboration between neuroscientists and computational/mathematical researchers, leading to the emergence of new scientific interfaces between biology and the world of information and technology.

Engineering methods have become the necessary tools for neuroscientists to tackle major scientific issues which up to now constituted often insurmountable bottlenecks, such as the ubiquity of dynamical processes ruling neural behaviour and their underlying recurrent connections, the design of sophisticated methods for non-linear model identification from experimental data and for parametric space reduction, the development of formal analysis of neural dynamics applied to closed-loop interactions, and the construction of real-time hybrid interfaces between computers and the living brain (dynamic clamp, brain-machine interfaces).

The importance for the future in developing interfaces between neuroscience, applied mathematics and computer science, is attested by the recent (2013) launching of long-duration (> 10 years) and large-scale interdisciplinary projects of unprecedented ambition at the international level: BRAIN (US) launched by NIH, private parties (Google) and donations (Kavli); The Allen Institute, financed by one of the Microsoft founders; the European Flagship "The Human Brain Project," financed by the Future Emerging Technologies Program of the IT division; MINDS (Japan) and a Chinese project yet to come. The common ambition of these different projects targets a better understanding of the brain and its realistic simulation.

However, the impact of Systems & Control theory on neuroscience research and, more acutely, interdisciplinary teaching, remains in its infancy. Below is a list of possible directions in which the discipline of Systems & Control can contribute to make advances for facing these challenges.

Bio-engineering and control theory to understand and simulate

One key limitation in our current understanding of brain dynamics stems from the fact that the brain can be seen as a nested hierarchy of various levels of integration, interacting together from its most molecular components to the most holistic features of cognition and behaviour. At the microscopic scale (subcellular and cellular) the mechanisms underlying neuronal communication (such as synaptic transmission and generation of an action potential) already exhibit strong complexity due to the diversity (electrical, physical and chemical) and nonlinearity of the processes involved. At a more integrated scale, the activity of the neurons of a given population gives rise to mesoscopic features that may be quantified using signals such as firing rates of single or multi units, local field potentials or their metabolic/fluorescence impact seen with intrinsic/extrinsic imaging measures. At the macroscopic level, the contributions of the activity of different neural populations that can be collated through EEG, MEG, and more indirectly fMRI measurements, can be correlated to a specific behaviour of the individual.

Classically, the study of neural phenomena is limited to a given scale. It usually relies on experimentally tuned models and gives rise to fascinating challenges from a dynamical systems perspective. Ground-breaking experimental advances (in optogenetics, electrophysiology and optical recordings) provide unprecedentedly fine data to nourish dynamical models. Observers, advanced identification methods, and nonlinear model reduction still need to be developed or adapted to the specificities of neuroscience data. The ubiquitous feedback interconnections, uncertainties, nonlinearities, delays and spatiotemporal evolutions all plead for the use of advanced analysis tools from systems theory, especially in an infinite dimensional or multi-scale context. The development, in the control community, of methods that inherently cope with model uncertainties, such as monotone systems, could provide precious analysis tools for experimentalists. Moreover, the transient nature of typical brain signals (for instance in sensory, perceptual motor tasks, speech, recognition or memory) or their oscillatory features (regular spiking, bursting, central pattern generators or brain waves within specific bandwidths: alpha, beta, gamma, delta…), require novel modelling and analysis approaches that go beyond stability of a sole equilibrium point and raise new synchronization-asynchronicity related questions. Optimal control theory may also constitute a relevant tool to understand the principles underlying neural coding.

Furthermore, major developments in mathematical neuroscience rely on stochastic approaches. They assume that internal random noise sources add to deterministic signals that can be causally locked with a sensory input or an internally generated feedback. At the molecular scale, the dynamics laws ruling the opening of ion channels, responsible for the current exchanges between intracellular and extracellular media, are mostly probabilistic. At the cellular scale, stochastic modelling allows for accounting of the inherent unreliability of spike generation. At the population scale, it permits better fitting to *in vivo* recordings and is instrumental in the understanding of some neuronal functions. Advances in the analysis and control of stochastic dynamical systems is therefore of relevant use in a neuroscience context and sometimes leads for unexpected use of noise in a Shanonnian perspective such as stochastic resonance.

In contrast, other approaches exploit deterministic dynamical systems. This second class of theory, although more phenomenological, is more provocative and proposes that the irregularity of activity patterns in the brain comes from the recurrence of the network and asynchronous firing. Consequently, perseverant on-going activity (seen in the absence of external drive) has the same signalling value as the sensory drive itself. New data suggest that, in spite of their limited biophysical performances, the subthreshold membrane potential trajectories in sensory cortical neurons become highly reliable when the full network operates in closed loop in natural conditions. This opens the possibility that there is nothing such as noise in cortical networks and that on-going states recapitulate with different time scales activity patterns evoked by past sensory-motor experience. This suggests that the dynamics of sensory cortical networks switches to an irregular asynchronous but highly deterministic mode (near the edge of chaos) when the dimensionality of the input drive reaches the internal memory capacity of the network. This opens the need for new concepts to define the relation between external constraint levels and versatility in network dynamics.

One of the main difficulties that neuroscientists have to solve is that observation of the living brain is done through multiple observable variables of different biophysical natures. Experimental measurements still obey a kind of generalized Heisenberg principle that jeopardizes comparison between physical measurements, each characterized by very diverse precisions in space and time (for instance, comparison of intracellular, voltage-sensitive dye imaging and brain fMRI).

But the main obstruction to a deeper understanding of brain functioning probably lies in the relationships between one scale and another. Existing analysis tools do not yet encompass all mechanisms from the molecular regulation to the resulting behaviour at the macroscopic scale. For instance, while neuronal population models satisfactorily render the averaged behaviour of a given neuronal structure under appropriate assumptions, they are not yet able to satisfactorily model the influence of a behavioural change of a given neuron or a given ion channel. This challenging question of a cross-scale unification requires radically new paradigms to which control theory may contribute significantly. The emergence feature of complex systems is probably the easiest to illustrate in neuroscience (e.g. maps of orientation preference in higher mammalian visual cortex), but recent approaches have targeted the converse immergence feature (typical of complex systems) where mesoscopic constraints (such as specific brain states) change the transfer function of more microscopic elements. This explains why pure LEGO-like bottom-up approaches are doomed to fail in simulating an entire brain. A new field of applied mathematics in which dynamical aspects are expected to be central, has yet to be invented.

Control theory to palliate and possibly cure

While loops are naturally omnipresent in neuronal processing chains, artificial loops may be added for treatment purposes. The actuation can take several forms: chemical in the case of pharmacological treatments; electrical in the case of activity neuromodulation triggered either electrically (including deep brain stimulation), magnetically in the case of transcranial magnetic stimulation, or optically in the case of optogenetics. The artificial loop is closed when the stimulation signal depends on real-time measurements of the subject's activity, for instance muscular activity or neuronal activity in specific brain structures through implanted electrodes or surface EEG sensors. This type of control finds its roots in the impressive development of brain-machine interfaces and the rapid growth of a new field, neuroprosthetics.

The possible gains expected by such a closed-loop approach in terms of efficiency, robustness and adaptability would be of no surprise to a control engineer. Neurological treatments exploiting closed-loop devices have started to appear in the neuroscience community, but most of the experimentally-tested approaches rely on rather simplistic views of feedback and often boil down to on-demand stimulation, meaning open-loop treatment once a pathological condition is detected.

In the case of electrical neuromodulation, one obstacle that probably slowed the use of more advanced closed-loop policies stands in the unavoidable stimulation artefacts: when sensing and stimulation take place in neighbouring brain structures, the measurement signal may be drowned in the applied stimulation signal (usually of much larger amplitude). The recent development of new stimulation techniques, especially optogenetics, allows removal of these artefacts by relying on a stimulation signal whose nature (light in the case of optogenetics) is different from the measurement signal (electrical). This latter technique also allows activity control to couple with the genetic dissection of subcircuits or specific classes of cells.

The objective of altering, or even controlling, the response of neuronal ensembles opens the door to a wide range of possible therapeutic and clinical applications, including Parkinson's disease, blindness, epilepsy, chronic depression or obesity. It raises fundamental questions in terms of controllability and observability, and may benefit from current knowledge in feedback stabilization, optimal control, or robust control.

Thus, state-of-the-art experimental technologies in neuroscience allow both refined dynamical models and easier real-time measurement and actuation, which opens up radically new perspectives for the control community.

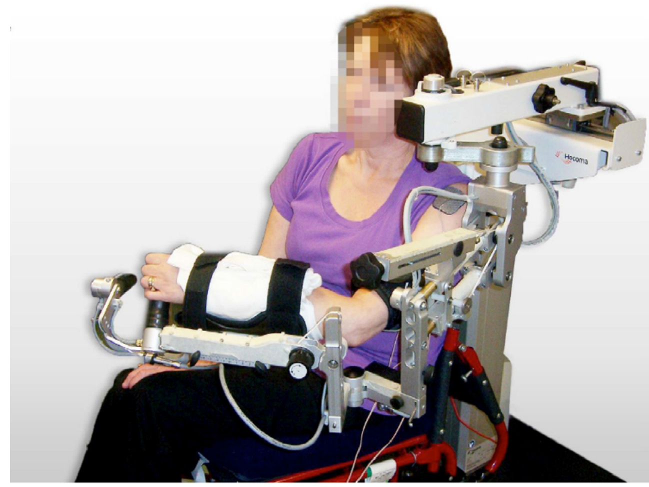

Fig. 17. Assistive device.

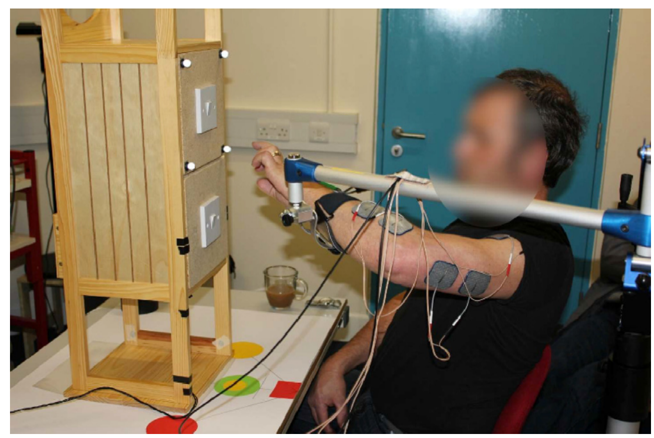

Fig. 18. Assistive device.

5.10. Assistive devices for people with disabilities

Control engineering has impacted the development of assistive devices for people with disabilities since its inception. It is noteworthy that the timeline of the historical development of powered wheelchairs aligns very closely with that of the development of automatic control. Control engineering has also played a crucial role in the development of prostheses with recent contributions applying control techniques inspired from robotics to enable powered prosthetics to adapt to the wearer's environment and to help amputees walk. Such mobility aids promote independence and enhance their quality of life. For children, the opportunities afforded by mobility aids are crucial not only to their physical development but also to their social and cognitive development.

Independent mobility is essential to maintain or regain an appropriate quality of life. The prevalence of wheelchair users continues to increase with increased survival from neurological conditions. For example, 80% of people with spinal cord injuries are expected to depend on a wheelchair for the rest of their lives. Flexible performance is required in that some users will recover from a state of almost no physical function to independent control such as those with Guillain-Barre Syndrome. Some of them will exhibit little or no change over time as is the case with Cerebral Palsy and some will have a progressive condition where they begin with independent control of the wheelchair and after some period of time have little or no independent control as in the case of a patient diagnosed with Motor Neurone Disease. In addition, within each such user group there will be individuals with a range of cognitive impairment, and cognitive ability may deteriorate over time. These challenges are compounded once the environment of use and safety factors are added to the system specifications. The control system must be appropriate for typical user environments such as the home, hospital or school, where the location of objects and people will change over time, as well as even more challenging outdoor environments.

In terms of control, the modelling challenges, particularly in relation to uncertainty, are key. There are challenges around sensors for autonomous navigation where the wheelchair must be able to locate itself in different environments and subsequently decide upon an appropriate course of action. There are also challenges regarding controller design for wheelchair systems which have varying degrees of autonomy and intelligence and must incorporate those legislative frameworks which are necessary to ensure the safety of both wheelchair users and other individuals. As has been the case in the development of drug treatments, the formulation of appropriate data sets and trials to evaluate the developed systems come much more to the forefront for the control engineer.

Assistive devices (Figs. 17 and 18), from prostheses to wearable robots, seek to further empower individuals to achieve independence where the goal may be to provide a normal gait pattern with the flexibility to transition between walking on a flat surface to climbing stairs, for example. Power requirements around motors and actuators for these systems are an important constraint for the control engineer and provide particular challenges in an uncertain environment. It has already been demonstrated that incorporating neurological signals within the control system along with the classical mechanical sensors, greatly improves performance both in terms of classical measures such as error tracking but also in terms of patient acceptance. Thus, integrating a range of sensors into the control paradigm becomes of increasing importance.

As for the wheelchair, guaranteeing safe operation in uncertain environments across a range of users provides growing challenges for the control engineer. Control is also playing an increasingly important role in the development of rehabilitation systems which assist individuals who have suffered an illness or injury to restore lost skills and hence regain maximum self-sufficiency. The demand for this technology increases with medical advances and a world population where life expectancy is increasing. In this domain, supporting rehabilitation by technology underpinned by appropriately designed control interventions, which aid individuals in performing rehabilitation tasks and exercises, can reduce the costs to healthcare providers and optimise the intensity of therapy. It also permits more formal evaluation of an ongoing intervention's progress. Recent contributions have seen control engineering underpin the development of systems to rehabilitate upper limbs following stroke where interventions are key to optimising recovery outcomes.

Underpinning the contribution of control to the development of all assistive devices for people with disabilities, is the ability to adapt performance in the presence of changing environments and uncertainty thus empowering sophisticated decision making. The variable nature of fatigue and the highly uncertain degeneration of patients suffering from neurological diseases must be accommodated.

Future challenges for control engineering in this domain are aligned with the research and innovation challenges in the domain of human cyber-physical system interaction and human in the loop control design. Patient-specific predictive modelling will be coupled with control methods which will provide a means to undertake the systematic treatment of uncertainty in increasingly complex models. Control paradigms will become increasingly centred on the individual for whom a particular assistive device is being developed which will further enhance independence and quality of life. This will ensure that a broader group of individuals will more keenly embrace the resulting technological interventions which in turn will improve success rates and longevity, reducing the need for revision surgery. The future focus will be on the application of control to restore function and optimise those physical interventions which achieve the best results with minimal invasiveness.

For the control engineer, the paradigm in this domain frequently does not requires a solution that is defined in terms of physics but a solution that is defined in terms of less explicit metrics around user acceptability. The control paradigms must incorporate elements from the arts as well as the sciences.

5.11. Health care: from open medication to closed loop control

The treatment of chronic diseases involves a repeated cycle of taking measurements from a patient and using this data to adjust medication or to implement other interventions. The objective is to achieve the best possible outcome for the patient. This cycle is a quintessential example of feedback control, and therefore great opportunities exist for combining clinical and control engineering knowledge to improve the treatment of chronic diseases.

Better treatment of chronic diseases would bring enormous benefits to society. In purely financial terms, the global cost of chronic diseases is forecast to reach US $47 trillion for the period 2011–2030 (World Economic Forum, 2011).

Close collaboration between engineers and clinicians already exists in several areas. A closed loop, implantable cardiac defibrillator is currently commercially available. There is also ongoing research on the application of closed loop control to improve treatments for type 1 diabetes, Parkinson's disease, epilepsy, high blood pressure and paraplegia.

The nature of the feedback problem varies from disease to disease. Each medical condition requires consideration of unique issues, such as the availability of measurements, the natural time constants associated with the variables of interest, and the time taken to respond to external intervention. In some cases, the feedback operates on a short time frame, on the order of minutes, while in other cases, the feedback can operate on a longer time frame on the order of hours or even weeks. An example of the former is an implantable defibrillator which senses arrhythmia and discharges, whereas an example of the latter situation occurs in diabetes mellitus.

A control engineering viewpoint of the treatment of chronic diseases, in common with all other control problems, will include consideration of issues such as dynamic modelling, prediction, design of observers, development of novel sensors, coping with uncertainty and dealing with constraints. Another key issue is that of patient safety. Engineers are well-aware of the safety-critical nature of engineering systems such as bridges, nuclear reactors, aircraft and so on. These issues are equally, if not more, relevant to chronic disease treatment since patient well-being is obviously of paramount concern. How does one guard against malfunction of drug delivery devices, measurement devices or implementation platforms? Such incidents could be life threatening if not dealt with appropriately.

As a specific example, approximately 8% of the world suffer from diabetes in one form or another, and roughly 10% of those have type 1 diabetes. The incidence of type 2 diabetes is growing at an alarming rate worldwide. Current treatments for all types of diabetes can lead to poor outcomes in the long term including cardiovascular disease, loss of sight and renal failure. Short term complications include seizures, coma and death in extreme cases. The current treatment of diabetes involves taking frequent measurements of blood glucose level (BGL) and then injecting insulin accordingly, which takes effect over the next several hours.

Engineers and medical practitioners are already working together to develop new strategies to treat diabetes patients. In particular, there is a worldwide effort to develop a so-called "artificial pancreas" to provide closed loop insulin delivery for people with type 1 diabetes. However, much work remains to be done before a truly autonomous closed loop system for the treatment of diabetes becomes a reality. Some of the necessary ingredients include:

- **Dynamic modelling:** A dynamic model is needed to accurately describe the impact of disturbances (carbohydrate, fat, protein, exercise, stress) and manipulatable inputs (insulin) on BGL.
- **Prediction:** The BGL response can extend over many hours and thus long-term prediction is needed to guard against the potential of future hypo- or hyper- glycaemic events.
- **Design of observers:** Tools are required for estimating the current state based on combining the available model with the measurements of inputs, outputs and disturbances.
- **Development of novel sensors:** The most reliable sensor for BGL is a fingerstick measurement. Continuous glucose monitoring sensors have appeared in recent years but the current devices are noisy and prone to drift. Also, accurate measurements for external disturbances are, as yet, in their infancy. For example, exercise, stress, food composition, etc. are usually estimated by patients and entered manually.
- **Coping with uncertainty:** A major consideration in diabetes treatment is that patients do not lead highly regulated lives. Thus, future food intake and exercise patterns have a random component. This inhibits making accurate predictions over a

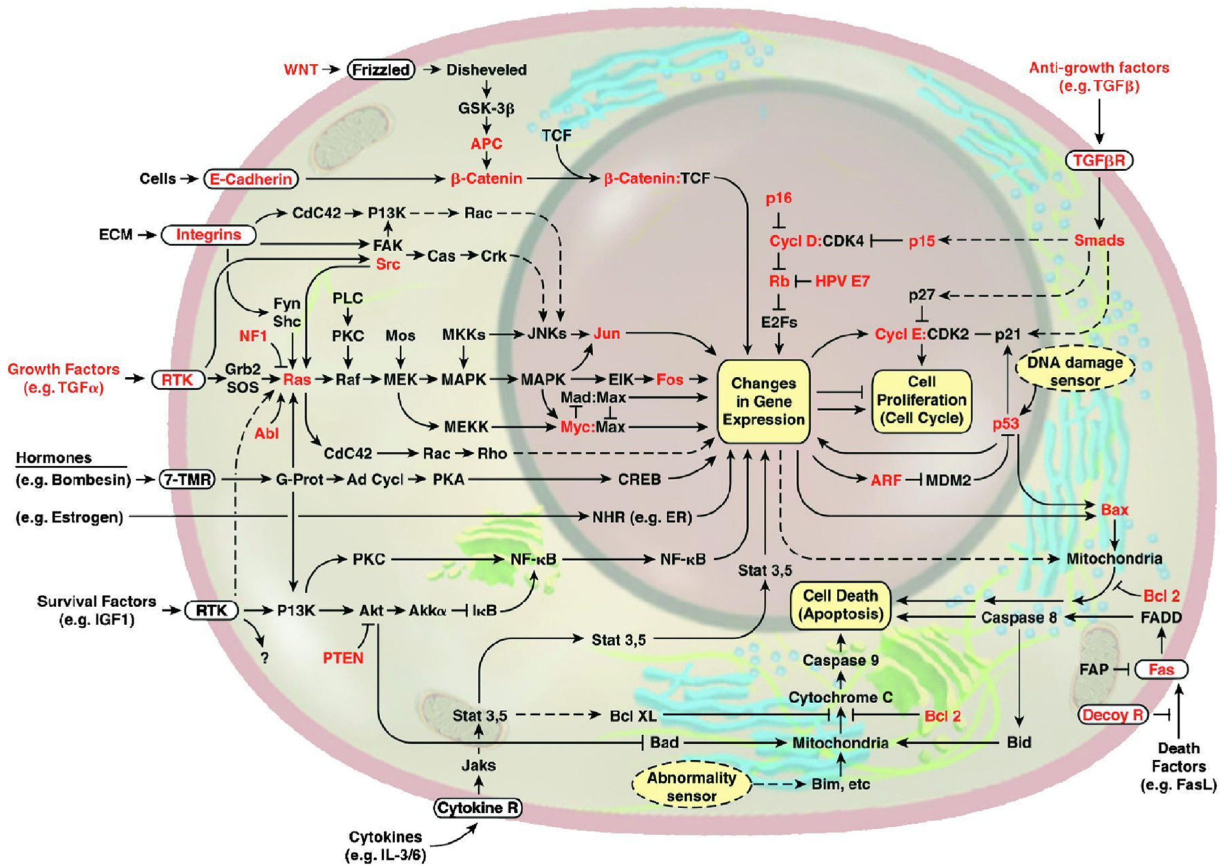

Fig. 19. A human cancer-signaling network. Genes known to be functionally altered in cancer cells are in red. From "The hallmarks of cancer," Hanahan D1, Weinberg RA., Cell. 2000 Jan 7;100(1):57–70.

long period. Yet such predictions are central to achieving ideal outcomes.

- **Dealing with constraints:** Insulin can be added but not removed from the body. Also, high BGLs lead to long-term health issues and low BGLs lead to both short- and long-term health issues. Thus, there are critical constraints on both inputs and outputs.

The above issues have a clear overlap with Systems Theory and Control Engineering.

There is a long journey ahead. However, the prospects are enormously positive. Engineers and clinicians have already developed mechanical hearts that offer the same life expectancy as donor hearts. Thus, it is reasonable to claim that better treatments for a wide variety of chronic diseases, achieved by combining engineering and clinical knowledge, are only a matter of time and dedication.

5.12. Cellular and bio-molecular research

Biological research has been collecting a tremendous amount of information about the biochemical processes that go on in living organisms and enable life, growth and reproduction. Biologists have spent a considerable effort sketching the network structure for many processes, as typified by a well-known diagram for the basic pathways involved in cancer (Fig. 19).

However, the complexity of the interactions, even in the simplest of organisms, single-cell organisms like *Escheria coli* or *Saccaromyces cerevisiae* (baker's yeast), is huge when analysed at a more detailed level. Moreover, most available information is still semi-quantitative (describing only the existence and the relative strength of interactions) and static.

For a deeper understanding of disease processes, and for a systematic design of drugs or of microorganisms for the production of pharmaceutically or technically relevant substances, an increased interaction with Systems & Control research is strongly needed. Only the creation of mathematical models that can be simulated and (at least partially) analysed will enable one to understand the dependencies of the different mechanisms fully and to distinguish between dominant and subordinate, fast and slow interactions.

Systems & Control theory offers tools to analyse nonlinear, dynamic, discontinuous and hybrid behaviour in complex systems, and in turn will benefit from the new theoretical challenges posed by the biological community. Building mathematical models also provides a common reference for the integration of results of work in the laboratory, the determination of crucial open questions, and the formulation of new biological questions. Of course, sufficient trust in "in silico" experiments has to be built as a first step. Recently, biological research has expanded from the analysis of natural organisms to the deliberate re-design of biological systems, termed synthetic biology, by removing existing components and adding new ones to the living cell. For example, one may build binary memory devices, a first step toward building biological computers, by employing genes from bacteria and viruses, see Fig. 20.

Systems & Control theory provides a useful framework to study this re-design. Control functions are crucial to the viability of such artificial biochemical systems, and the field of Systems & Control theory is in an excellent position to contribute to this area based upon a long record of synthesis of control functions of all kinds. For example, it is now possible to build genetic systems which use

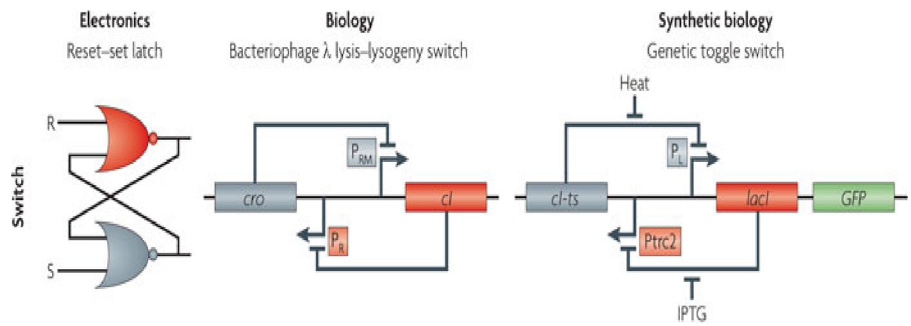

Fig. 20. Analogies among latches in electronics, binary decision systems in bacteria, and toggle switches constructed in synthetic biology. Inputs to biological systems may be optical, physical (heat, pressure), or chemical. From *"Synthetic biology: applications come of age,"* Ahmad S. Khalil & James J. Collins, Nature Reviews Genetics 11, 367–379 (1 May 2010).

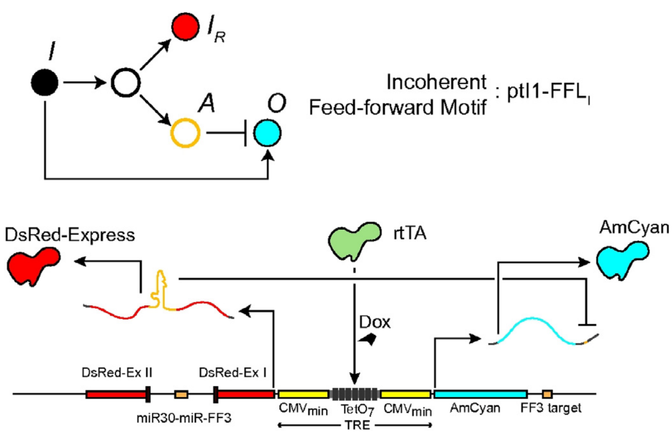

Fig. 21. A circuit that provides robustness to gene copy number (top: conceptual diagram; bottom: actual gene construct). From *"Synthetic incoherent feed-forward circuits show adaptation to the amount of their genetic template,"* Bleris, Xie, Glass, Adadey, Sontag, Benenson, Molecular Systems Biology 2011.

feed-forward architectures based on microRNAs that allow protein production to be robust to uncertainty in the number of copies of a gene present in the organism, see Fig 21.

5.13. Factory of the future and logistics systems

5.13.1. Factory of the future

The Factory of the Future is a generic concept that is part of a general awareness of the importance of manufacturing industry for nations' development. This reflection is intended to maintain and develop a strong industry, generating wealth and job creation. Hence, the Factory of the Future has to take into account several simultaneous transitions: energy, ecological, digital, organizational and societal. Factories have to transform themselves to become more sustainable in industry and more respectful of the Earth.

The first industrial revolution at the end of the 18th century was characterized by the introduction of mechanical facilities using water and steam power. The second industrial revolution appeared during the 20th century with the introduction of mass production and a division of labour with the help of electrical energy. The third industrial revolution started later in the 20th with the introduction and use of electronic systems (Programmable Logic Controller: PLC) and software (Supervisory and Control Analysis and Data Acquisition: SCADA and Manufacturing Execution Systems: MES) that achieved further factory automation. Today, the massive use of digital and information technologies like Cyber-Physical Systems (i.e. network of interacting elements with physical input and output instead of as standalone devices), Internet of Things (IoT), Machine to Machine (M2M) communication, Big Data and Cloud Computing represent what it is called the fourth industrial revolution. The names are different all over the world: Industry 4.0, Internet Factory, Smart Plant, Digital Factory, Integrated Industry, Innovative Factory, Intelligent Manufacturing, e-Factory or Advanced Manufacturing, but the concepts are the same. The convergence of the virtual world of the internet and IT (Information Technology) and the real world of industrial installations and OT (Operational Technology) will be the challenge for the Factory of the Future. Modern information and communication technologies seem a solution to increase productivity, quality and flexibility within the industry. Hence, the industry has entered a phase of big change that sees digital technologies as a key factor for the future to design Cyber-Physical Production Systems. These systems are predicted to enable new automation paradigms and improve plant operations in terms of increased facilities effectiveness. The challenges (the list is not exhaustive) are numerous:

- **Connectivity and interoperability**: the ability of cyber-physical systems, humans and factories to connect and communicate with each other via the Internet of Things and the Internet of Services is a big issue. Internet compatibility and open standards are key elements in the expansion of large-scale automation systems. Machine-to-machine communications using Internet of Things principles will define the Cyber-Physical Production Systems of tomorrow. IT security (or cyber-security), aggravated by the inherent need to open up previously closed production shops, is a major Information and Control Theory (ICT) research need. In addition, Machine-to-Machine communications offer a large range of questions for ICT research: How to guarantee reliability and stability of critical M2M with very short and stable latency times? How to maintain the integrity of production processes? How to protect automation industrial processes and know-how (like the PLC program)?
- **Virtualization:** a virtual copy of the factory which is created by linking sensor data (from monitoring physical processes) with virtual plant models and simulation models enables controllers to be checked and validated prior to implementation. Virtual commissioning, process simulation and techniques like model checking or formal methods are key components of the research agenda for making a safe Factory of the Future a reality.
- **Decentralization:** Cyber-physical systems enable collection of large amounts of data about machines and products. With the right treatment and analysis tools, the data can be used, for instance, to identify production line problems. The ability of cyber-physical systems to make decisions on their own with real-time capability to collect and analyse data in order to provide information as fast as possible is an open research question.

- **Innovative production lines and logistics:** To be competitive, time and costs associated with developing and manufacturing ever more complex products must be reduced. Part of the solution can come from virtualization, as well as from the merger of virtual planning and new physical production processes based on plug-and-play machines. Additive manufacturing is also raising many hopes. 3D printing already allows production in small series of complex parts, spare parts or custom tools. In the future, speed and precision should increase and allow additive manufacturing on a large scale and considerably modify factory design. Hence, how to get flexible adaptation of a production line to match evolving requirements is still an open question. In addition, the Factory of the Future has to be able to guarantee the quality and traceability of products, to manufacture individualized products in adapted quantities – clean, silent, saving raw materials and energy, and human-centred – or adapted to better take into account the expectations of employees. All of these goals can be linked to ICT fields, needs and directions already highlighted in this report.
- **Human-centred automation:** IT, automation and collaborative robotics can liberate a human heavy or repetitive tasks but also can support his cognitive functions. For instance, augmented reality glasses can immediately provide information on maintenance. Analysis, Design and Evaluation of Human-Machine Systems for the Factory of the Future, is a field where the control community can significantly contribute.

This major technological breakthrough with new digital tools offers a priori an extraordinary field of innovation open to researchers in Systems & Control. In the case of the Factory of the Future, the dynamic system is characterized by massive interconnection, the processing of huge amounts of data and new forms of synergy between humans and technical systems. ICT research has to focus on these particular classes of dynamic systems in order to provide tools for modelling, designing, simulating optimizing and validating them.

In the last decade, cloud computing, big data and machine learning are areas where computer science has made major advances and is recognized for that. The main challenge for the controls community could be its capacity to interact and cooperate efficiently with the computer science community in order to propose methodologies and tools integrating the two worlds: IT and OT. At the very least, control and system engineers will toned to combine know-how related to ICT with strong IT competencies that range from basic (using spreadsheets and accessing interfaces) to advanced (applying advanced programming and analytics skills). The need for multiple hard and soft skills will become more and more important. Employees will need to possess greater flexibility to adapt to new roles and work environments and become accustomed to continual interdisciplinary IT and OT learning. For this reason, control education, training and outreach must evolve and adapt to the requirements of the Factory of the Future and more generally to our society. This can be achieved by integrating teaching and research at all levels in order to promote control as a field that spans Science, Technology, Engineering and Mathematics (STEM).

5.13.2. Manufacturing systems and logistics

The major applications of control theory in **manufacturing systems and logistics** are in the Horizontal and Vertical Integrations of Decision Making Process and in a wide use of analytics at all levels of the decision and control process. Indeed, the pressure of the competitive global market has intensely affected the production systems, calling for:

- integration of the activities that cover the whole production spectrum from customers' requirements to payment;
- flexibility in the face of customer-demand changes;
- drastic reduction of production costs.

Thus, the main directions of research are on optimal supply chain management and control, integrating reconfigurable manufacturing technologies for fast adaptation to changes in the quantity and mix of products, risk analysis in global supply chains, new control and management approaches taking into account fundamental transformations to new product-based economics through internet-based service enterprises and demand-driven supply chain based on multi-criteria approaches and techniques.

Supply Chain Engineering is an emerging field based on analysis and comprehension of the essential principles of production and distribution systems. This scientific domain concerns the methodical evaluation and optimization of production systems, logistics networks, and their management policies to increase the effectiveness of multifaceted demand and supply chains.

To reach these objectives, radical changes have been introduced in production systems, thanks to new manufacturing technologies that increase efficiency and IT technologies that improve system organization and management. Furthermore, dynamical pricing and revenue management, which proposes approaches that define the price of the products based on market situations, attracts more and more researchers and practitioners. Pricing stresses the return on the investment.

Supply chains are emblematic examples of the renewal of production systems in recent decades. It is through this new paradigm that cost reduction and service enhancement can be achieved. To make this easier to implement, new types of manufacturing systems have been introduced. Examples include: reconfigurable manufacturing systems (RMS), assembly lines with workers' flexibility, bucket brigades or U-shaped assembly lines. Over the same period, new technologies arose to monitor the state of systems in real time. We can mention radio-frequency identification (RFID), internet applications or "intelligent" storage facilities, to name just a few. These technologies favour one of the most important objectives of production systems management: the ability to make a decision almost immediately.

Radical changes in the criteria that express the new objectives of production systems in the face of competition are another important aspect. The introduction of some new criteria reflects the just-in-time (JIT) requirements. For instance, conventional scheduling optimization is now restricted, in the best case, to deciding the order products are launched in production. In other words, the conventional scheduling activity migrated from the tactical to the strategic level. In actual production systems, this is replaced by a real-time scheduling, also called real-time assignment. Other criteria are used to reflect quality, flexibility and work-in-progress (WIP): adequate quality is now unavoidable to meet customers' satisfaction; flexibility is a necessary condition to remain competitive in an ever-changing market; and reduction of WIP is a factor to minimize the production cost and the probability of obsolescence.

Five challenging issues:

- **Supply chains are emblematic examples of the renewal of production systems in the last decades.** Supply Chain Engineering is an emerging field for application of control theory. This scientific domain concerns the methodical evaluation and optimization of production systems, logistics networks and their management policies to increase the effectiveness of multifaceted demand and supply chains. The major industrial problems and various effective approaches of inventory control in supply chains, use of RFID and internet applications or intelligent storage facilities are being examined. Radical changes in the criteria that express the new objectives of production systems and logistics are on-going: JIT requirements, dynamic scheduling, dynamic pricing, etc. In addition, the main concerns of outsourcing are being detailed. In particular, vendor selection and evaluation models are being developed. Certainly, warehouses are critical components of supply chains. Their usefulness is highlighted and their various functions and equipment are being analysed. The design stage is also being extensively considered via developing storage algorithms as well as examining warehouse sizing static and dynamic models.

- **Operations risk analytics will enable the growth and understanding of best practices in operations**, e.g., pricing functions. Banks are processing millions of transactions every day in order to protect against fraud and terrorist financing. Energy companies monitor operations process and customer activities to protect against abnormal spikes in demand. Risk analytics in business intelligence represents data-oriented techniques to supplement business systems for better risk-based decision making. Risk performance analysis in manufacturing intelligence uses advanced data analytics, modelling and simulation to produce a fundamental transformation to new product-based economics through internet-based service enterprises and demand-driven supply chains. Risk evaluation plays key roles in emerging areas such as bio-manufacturing, nanotechnology and energy. We see a dramatic increase in the use of predictive analytics in these and many other areas. This working group will bring together scientists who have different backgrounds and disciplines and provide a set of opportunities to discuss these open issues.
- **Ameliorating the situation of an industry requires reducing costs and maximizing customer satisfaction.** These two aims cannot be achieved without good management and good knowhow while making decisions. These decisions are generally associated with three levels of the hierarchical planning process: strategic, tactical and operational levels. Generally, manufacturing industries aim to determine the most adequate Integrated Maintenance-Production Strategies which help them optimize system exploitation and reduce costs. Releasing efficient planning urges firms to have a global vision on their production and maintenance process which may be looked upon as an inter-dependent set of subsystems performing various functions including ordering raw materials, assembling pieces, controlling quality, repairing machines, storage, etc. One of the key current issues in integrated maintenance production strategies research is to develop a set of new which integrate maintenance and production aspects while considering several environment constraints. The real goal is to face the various contemporary industrial constraints in order to optimize the system and reduce costs.
- **The new competition is a major upheaval affecting every aspect of how enterprises organize and operate.** The evolution from single enterprise with a high vertical range of activities toward enterprise networks offers new business opportunities especially for Small and Medium Enterprises (SMEs) that are usually more flexible than larger companies. However, in order to make a successful commitment to an enterprise network, expected performance and benefits must be carefully evaluated and balanced for a company to become a partner of the right network and for the right tasks. All these issues must be considered in order to find an efficient, flexible and sustainable solution. The underlying logistic networks are complex and their analysis requires a carefully defined approach. As technological complexity has increased, logistic networks have become more dynamic and complex to handle. Multi-criteria approaches have been put to use in multiple segments of manufacturing and logistics. They have taken a prominent role in integrating people, information and products across integrated supply chain boundaries including management of various manufacturing, logistics and retailing operations in manufacturing, warehousing and distribution of goods and services. Decisions involving customer profiling, new product development, retail marketing and sales patterns are immensely refined using innovative multi-criteria approaches.
- **Another important issue concerns modelling approaches for designing and management of reconfigurable machining, assembly and disassembly systems**. One of the main characteristics of these automated systems is that they use reconfigurable manufacturing technologies for fast adaptation to changes in the quantity and mix of products. Indeed, the industry has new requirements for manufacturing systems given the shorter product runs and the need for more customization. Production systems should be designed to adapt its physical configuration to answer market fluctuations in both volume and type of product. One of the principal characteristics of Reconfigurable Manufacturing Systems (RMS) is modularity: in a reconfigurable manufacturing system, all the major components are modular (system, software, control, machines and process). Selection of basic modules and the way they can be connected provide systems that can be easily integrated, diagnosed, customised, and converted. An RMS is also supposed to quickly integrate new technologies to improve its efficiency. RMS is assumed to be the perfect tool for the new era of mass customization that requires simultaneously the productivity of a dedicated system and the flexibility of agile manufacturing systems.

5.14. Control in the high-tech industry

Over the last decades, the field of control has had a major impact on both the development and performance of high-tech machines, like those used in semiconductor chip manufacturing. For example, in lithographic tools that create the fine chip patterns on the semiconductor substrate, a sub-nanometer positioning accuracy of a mask's image during a high-speed scanning motion is essential to manufacture working ICs (Fig. 22). In the past 14 years, accelerations of stages in lithographic machines have increased six-fold, while simultaneously the positioning accuracy and settling times of said stages have improved ten-fold. One could advocate that none of these improvements would be possible without the aid of control. This is evidenced by the large number of control loops in lithographic tools which can easily add up to the order of 1000. To maintain this progress within the coming decade, the following challenges are identified as playing a key role in defining the control research agenda for the high-tech industry.

- **The growing contradiction between (sub)-nanometer positioning accuracy and zero settling requirements on the one hand, versus huge actuator forces on the other hand** puts heavy demands on vibration isolation and control. In order to achieve desiredscanning velocity, large (reaction) forces must be exerted which lead to vibrations in frames. These vibrations are not allowed to enter critical machine parts such as optics and measurement systems. Components that function as vibration isolators by themselves have become multivariable closed-loop systems interacting with the high-speed moving parts. Trends toward lightweight designs in an attempt to reduce the (reaction) forces come with the challenge of flexible system behaviour in which control could also play a crucial role. The control of flexible structures exploiting a large number of sensors and/or actuators poses challenges not only on the (modal) control design and optimization, but also on the system design, computation and (wireless) communication. Other challenges include further reduction of vibration levels by using multiple sensor types in parallel without introducing extra noise and minimizing the effect of noise sources and pressure pulses. Control challenges mainly lie in coupling distur-

Fig. 22. TWINSCAN NXT:1980Di lithographic system used for IC manufacturing.

bance measurements to vibration control systems and methods to deal with uncertain characteristics of, e.g. the floor on which the machine is placed.

- **High-bandwidth requirements for positioning performance can only be met using simultaneous optimization of mechanical and electronic components together with control design.** The time when individual design teams created separate components based on external specifications has long gone, hence the birth of mechatronics. Today, structure optimization, actuator and sensor design and layout, and controller design form an iterative cycle towards an optimal overall design. Since most of the design cycle time is spent creating working prototypes, the first actual hardware must be very close to the final system. Improved accuracy in modelling techniques has contributed considerably to this goal. The long-term trend here is toward integral hardware and control design. The ultimate goal is to find, in an optimization context, the optimal combination of controller and structure (including sensors and actuators) that minimizes some performance criterion subject to constraints. A short-term trend is typically given by input shaping and synthesis in which the hardware properties (possibly nonlinear and time-varying) are taken into account in the reference generation as a means to not excite structural modes or dynamic links.
- **Imaging accuracy is not only determined by mechanical positioning accuracy but requires a multi-physics approach including thermal behaviour and control.** The point-of-interest of the imaging process is not only determined by (position) sensor output, but also by mechanical and thermal modes. Modelling accuracy of these modes determines the correction potential that can be achieved by observer-like structures. A main challenge lies in the prediction accuracy of these phenomena in real time. Trends in model updating and observer design support the need for accurate models in model-based control design. This includes, for example, the characterization of visco-elastic materials and behaviour that through the introduction of rubber is expected to dominate future high-tech mechanics. Traditionally, these mechanics were designed to be lightly-damped. Also, trends in both analytical and experimental Linear Parameter Varying (LPV) modelling are important as being an enabler for control design techniques like modal decoupling and norm-based control. In the area of light source control, nonlinear modelling the physics becomes increasingly important. Another aspect in this regard is the fact that the point-of-interest, which is usually defined as the performance location where the actual process takes place, for example the exposure in the case of lithographic tools, is usually time-varying. Feed-forward control designs capturing the time-varying and position dependent characteristics associated with the point-of-interest therefore form a key challenge toward improved scanning performance and throughput. Approaches based on either distributed parameter systems or finite element models are expected to better deal with the position-dependent compliance and/or resonance effects occurring in flexible structures. This includes the time-varying aspect in the presence of right-half plane zeros being the result of varying sensor locations.
- **Coordinated cooperation between subsystems is increasingly important.** Individual control of multiple systems, e.g. stages and lenses, is no longer sufficient to obtain the required performance. More and more subsystems such as stages, lenses and metrology systems (Fig. 23) become actively coupled by control. Managing the resulting complexity is a major challenge in terms of stability, robustness, as well as integration and qualification of the individual components. Trends toward multi-

Fig. 23. YIELDSTAR Metrology system allows real-time chip property measurements enabling a closed-loop rejection of process disturbances.

agent and multi-level control are reasonable to expect from the increasing complexity in dealing with coordinated cooperation. And trends in multi-rate control, e.g. supporting different co-existing sampling frequencies, may prove necessary in this regard and offer significant challenges. Also, physical interaction between systems plays a larger and larger role when going to higher accuracies. For example, thermal, magnetic and acoustic coupling between subsystems can no longer be neglected. Compensation of these effects is in principle possible when accurate models are available that can be used for on-line prediction. In the absence of such models, data-driven calibration is considered as one of the key alternatives. Challenges in data-driven optimization and calibration are typically robustness to disturbance variation, time-varying aspects and bias in the parameter estimates, and convergence (speed) of the optimization algorithm. In this regard, an important trend is seen in machine-dedicated controller tuning as a means to achieve machine-specific performance by acting on machine-specific disturbances, coupling and interactions. A key challenge here lies in combining controller specifications, which are often posed in the frequency-domain, with optimization criteria, which are generally posed in the time-domain.

- **Cost reduction and fault monitoring.** Next-generation high-tech systems tend to become increasingly complex and costly thereby lowering the return-on-investment and resulting in downtime to become practically unaffordable. Trends toward advanced control designs as to maintain performance but with less-expensive hardware or automated fault monitoring, and diagnosis as to reduce downtime, are expected to grow in importance. Typical topics involve nonlinear control design for linear systems as a means to avoid bandwidth limitations, waterbed effects, and/or time-domain trade-offs like the necessary increase of overshoot by introducing an extra integrator in the control design. Automated fault monitoring and diagnosis typically involves detecting faults in large-scale data sets and pinpointing the source, state and/or trend monitoring, and model-based prediction combined with some form of supervisory control, i.e. an increased level of automation and control.
- **Machine functioning in a chip manufacturing fabrication facilities (fab) is able to mitigate process disturbances by applying large-scale feedback loops.** Disturbances induced in the factory's material flow, when measured, can be mitigated by the lithographic tool, e.g. by on-line changing the exposure dose. Measuring the on-product effects of the disturbances and creating an overall optimizing feedback loop, allows a further improvement of chip manufacturing quality. Trends toward holistic lithography in which control exceeds the system boundaries of the lithographic tool itself seem both promising and challenging at the same time. Measuring key performance measures like on-product overlay and using this information to modify process control parameters as to control product quality, may also evolve toward adaptation of the servo control loops of the lithography tool itself. Servo performance of the (coordinated) subsystems of the lithography tool is (partly) responsible for the on-product overlay in the first place.

In summary, tremendous progress in the density of features in integrated circuits would not have been possible without advances in the area of control, together with a large increase in modelling capabilities and structure design. To support the continuation of Moore's law while keeping manufacturing equipment affordable, many control-related challenges are identified. Further accuracy improvements force us to consider new physical phenomena that were not relevant before. Subsystems are coupled by control algorithms to optimize overall performance. In the end, IC manufacturing will create smaller features on chips, supporting lower power consumption while simultaneously increasing IC functionality, memory storage space and operating speeds. In this vision, control will again be indispensable.

5.15. Mechatronics and control co-design and automation

As stated in the previous section, modern mechatronic systems demand a high integration of functional and control design. This is not only true for the aforementioned high-tech industry but also for the development of general mechatronic systems. The main driver for this is the stronger differentiation by software functionalities in modern mechatronics. A key component, resulting in either better performance, efficiency or robustness, is the underlying control architecture and the implemented algorithms. As a consequence, control evolved from a fundamental to a critical technology, necessary to keep modern mechatronic systems alive, underlining its importance for future developments.

In Europe, many traditional mechanical industries have moved quickly towards intelligent machines and appliances, with the aeronautical and automotive industries spearheading this revolution. The added value of new products comes in part from more intelligent sensing and control. The traditional approach of sequentially and separately developing the mechanical design and control of components, pieces of equipment and machines is no longer sustainable, in view of the growing complexity and need for the optimal use of resources. A similar situation exists in chemical engineering where the design of the plant and its control structure must be integrated to enable flexible, efficient and safe operations. The modern approach in co-design requires interdisciplinary competence to merge the traditional strength of mechanical, industrial and chemical engineering education in rigorous modelling. Mechatronic research projects have so far only partially addressed the need for integrated system design. Only individual components and low-level integration used to be considered, lacking a wide system view. Research initiatives in the direction of integrated mechanical and control co-design as well as process and control co-design will boost the integration of competences and strongly impact future industrial competitiveness.

Model-based design. One way to improve model-based design is to merge development processes and combine formerly sequential steps in the generation of a mechatronic system. Given the complexity of today's applications, this is indeed necessary to keep up with shortened development cycles. Fortunately, the available tools allow us to move reality into a simulation model and get first-hand experience of the system's behaviour – from the complicated interaction with the environment to the cycle time of the control code, it all can be tested under realistic conditions. This shows the paradigm shift in the 21st century: while formerly general purpose controllers did the job fairly well (and some still do), highly complex devices need control co-design from the beginning since control is the enabler of a large portion of the functionality of the device. In addition, to make best use of the available resources, general purpose controllers are out of scope. The age of model-based design is blooming, where the control algorithms are defined to fit model/system needs without compromise. Consequently, engineers can start early on to optimize and improve the system behaviour and do not have to wait until the commissioning phase any longer.

Advanced control algorithms. This shift of paradigms is pushed mainly by advancements in processing power of computational devices. Only with these advances can we solve multi-stage optimization problems online, like in a Model Predictive Control (MPC) or bi-simulation of mechatronic systems. This allows us to know what caused the system to behave in a special way, deduce proper root causes for it and provide corrective actions.

Reduced modelling effort. However, now that controllers can be dedicated to a specific purpose based on sophisticated models of the controlled devices, there is one open point remaining: the lion's share of the work is actually not the design of the control algorithm, but it is the generation of a proper model. With automation technologies advancing towards the next industrial revolution, devices will have virtual counterparts and all data (e.g. dynamics) will be shared between both the device and the model, using internet and cloud technologies (Fig. 24). This way, devices will become self-aware, and so might simply be equipped with a detailed model of their dynamic behaviour. The future tools will be able to interpret such models of devices, and by combining devices to a system, their models will form a dynamic model of their own, available for detailed algorithm design and testing. Moreover, the tools will communicate with each other and propagate changes along the toolchain automatically. Adaptation of the hardware design hence becomes adaptation of the software design and vice versa.

Tailored control design. Finally, this will cut the effort needed to generate the model, and control engineers will be able to focus on their core competence – control design. Then we will be able to decide if we want performance, accuracy or robustness, and ask should the system suppress disturbances fairly well or do we have to track a perfectly shaped reference trajectory? These are the questions control engineers need to focus upon to squeeze the most out of a given approach and to dedicate their design to a purpose rather than vice versa. The consequences for mechatronic systems are tremendous: instead of perfectly designing a major part of it, only the most crucial elements need care, and control can take over ever more tasks to enable modularization, reusability and interoperability. Once there is a model available automatically, all optimization-based methods can finally prove their potential. None of the designed controllers will be set up without minimizing a well-defined cost function. This will allow purpose-based installations that directly fit the needs of the plant owners. Control will be a major driver for sustainability in an ever-more resource constrained world.

Reliable, robust and certified solutions. A second major point relevant for industrial applications is reliability. Given the details of today's modelling tools and the future self-awareness of the devices, the generated models become more and more sophisticated and detailed. Finally, this will help us to also respect failure rates of objects (e.g. communication channels) into the simulation, and more importantly, into the design of mechatronic systems. We can provide guarantees that might lead to software-based certification only, in the future. Moreover, we do not have to guess anymore what the problem is with the reliability of the current set-up. We know if we have to redesign the controller or if we need a better / different communication protocol to meet the standards of functional safety and other higher-level requirements. Again, control (by co-design), will be a main source and driver for increasing the quality of future mechatronic systems.

5.16. New dimensions of robotics

Previously robots were mainly used by big companies in high volume manufacturing (i.e. car industry). Industrial robots were perfect to execute procedures simplified to tiny (angular) movements. It takes approximately 400 times longer to program a classical industrial robot in complex operations than to execute the actual task. The robots are programmed by robot specialists.

Today, the classical industrial robots have appeared in small and medium sized enterprises (SME), thus the efficiency of robot

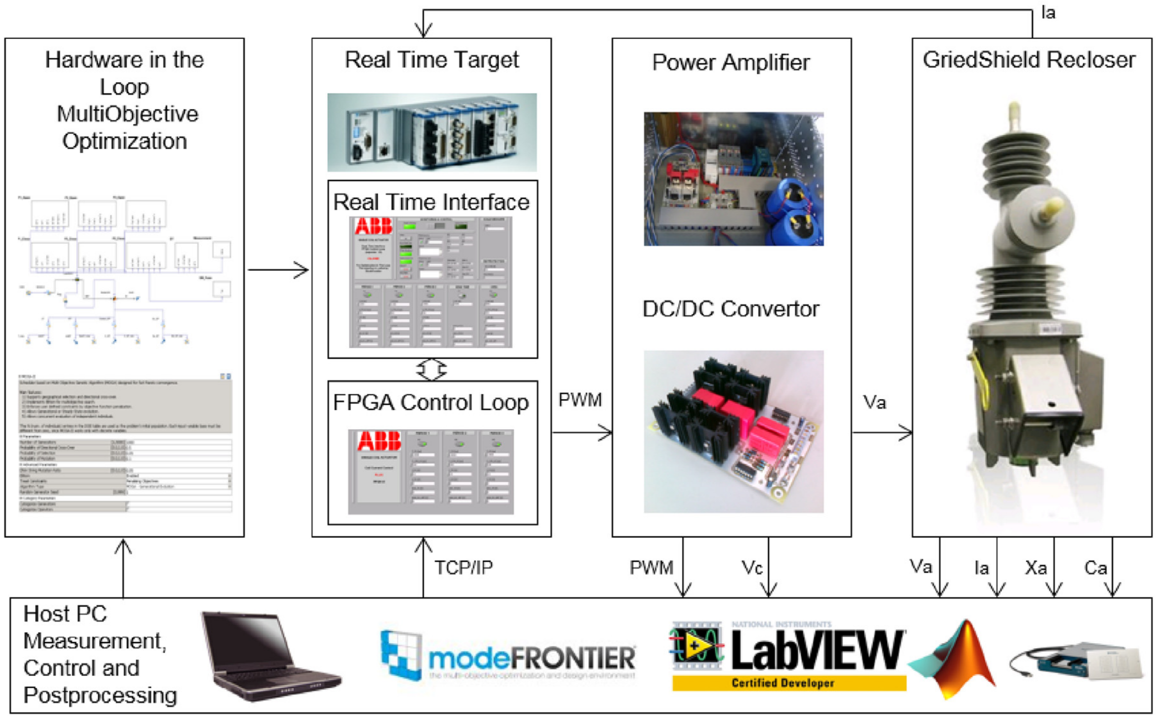

Fig. 24. Modern hardware-in-the-loop multi-objective mechatronic design optimization including mechanical and electrical component design and control functionality for the example of a medium-voltage recloser operation optimization.

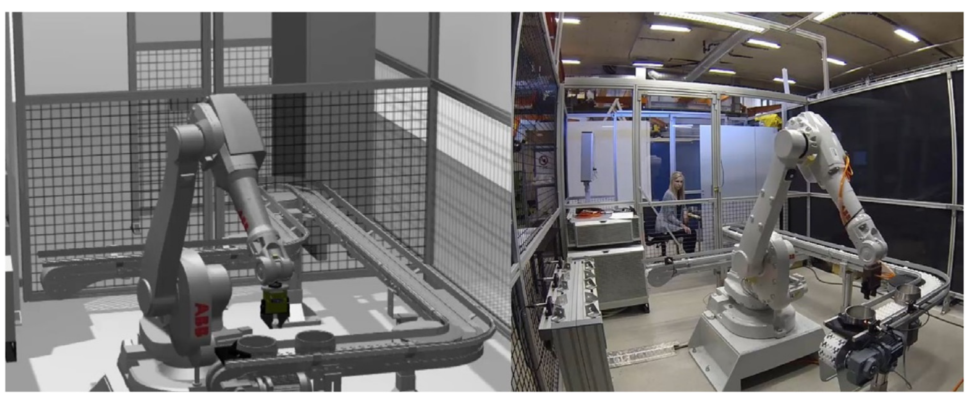

Fig. 25. Use of a virtual prototype enables faster commissioning at a lower risk. This concept will be increasingly important in the future.

programming methods must be improved in order to avoid losses caused by frequent switches in small scale production. Since an SME cannot employ a robot specialist, the robot is commissioned by third party system integrator and programmed by an engineer, who is not a robot specialist. One possible solution is virtual commissioning (see Fig. 25) and remote operation via augmented reality, where a robot specialist operates several robots working at different SMEs in real-time. Telepresence is a big challenge during remote operation. Telepresence is the psychological feeling of being in an environment based on a technologically founded immersion environment. It should provide the ideal sensation, i.e. we get the necessary information fed back from the remote environment with no delay. Another solution can be the super flexible programing or supervisory system. The easiest and most natural way of programming a robot is to show the task, or just interactively instruct the robot on the task. The robot needs special visual and tactile sensors and special cognitive abilities to understand and learn the situation. The robot is considered as an apprentice with the physical strength, manipulation ability and learning capacity required for precise manufacturing. It has a special kind of intelligence but is handicapped in some senses. Therefore it needs special treatment. We have to command it clearly in a special way, and we have to supervise its work. If we can learn how to communicate with and train this "new worker," we can gain a new capable "colleague." The long-term goal is that the plant manager would be able to assign daily tasks to a robot as naturally as to the human workers. For example, using CAD documentation and some verbal explanations. If we have an unskilled worker, the next step (challenge) is skill acquisition. Most of the manipulation skills can be learned in a non-model based trial-and-error method. The robots can learn tasks by mimicking the actions of a human operator, but a skill transfer is also necessary since the end-effectors of the robots are different from the human hand. A skilled robot can identify the problem and select the necessary skill to perform the task, thus becoming an autonomous agent, a manufacturing robot assistant. The next step is when the robot can work together with a skilled human worker, or more workers, and more robots (robot operation in shared space).

Robotics is facing a paradigm shift. From an organized industrial environment, robots are soon to step into our complex daily life. To counterpoint the new use, these non-industrial robots are called service robots in the corresponding ISO and IEC standards. Service robots need to have totally new functions and behaviours, which means that new problems of robot control are emerging and must be solved. For service robots, it is not enough to execute a pre-programmed action line. They must be able to adapt to changing environments, make their own decisions and in addition, they have to socially fit into the human environment. But when can we say that a porter robot was polite? Or a police robot decided and professional? How can we describe politeness, attachment, affordance and other social behaviours in a mathematical way? These questions might seem remote, however we need to think ahead. Along with the technical development of robots, we also have to address their social integration. Similarly, we start the education of a child in childhood and not when she/he's already grown up.

With the integration of service robots into our everyday environment and the expansion of roles they can potentially fulfil, new types of users will be introduced to robots. Alongside the classic users of robots, people can be divided to four main groups:

- Robot specialist engineer
- Engineer, but not robot specialist
- Non-technical but technology-literate end user like caregiver, employee of courier company, etc.
- Non-technical, not technology-literate end user like elderly people

Therefore, there is an increasing need to make the training of robots more automated, while at the same time, to enable robots to fulfil more and more sophisticated tasks. This requires a more sophisticated robot control.

In the future, robots will become part of our daily lives. We can also state it this way: robots are already part of our daily lives, but not in a coexistent manner. Industrial and service robots are working for humans every day. However, as long as we think that robots only exist to follow human orders, we have the wrong understanding of next-generation robotics. Today a communicational barrier exists between “robots” and humans. In order to fill this gap, we have to start in the beginning: how to create a system which has actuators, sensors and intelligence, in such manner that people accept and interact with it without having any kind of challenge communicating/interacting with them.

Today's robot applications are mainly classified in the following categories:

- Manufacturing (industrial and service robots)
- Healthcare (medical robots)
- Outdoor (disaster management, agricultural, construction robots)

All above-mentioned sectors have their own specialties related to communication/interaction with robots. Up to now, most of the effort was to achieve higher user friendliness: to simplify communication interfaces, to ease programming and to develop artificial intelligence, which could react to sudden changes of the environment, etc. However, the main challenge is real autonomous operation in the non-industrial (non-standardised: e.g. outdoor, medical) environment and in the social behaviour these robots display for example in households and offices, in caring for the elderly, in developing the abilities of autistic children or in rehabilitation. If we want to coexist and cooperate on a daily basis with robots, we have to first think about robots not simply as human replacements, but as co-workers that aid their partners. To achieve this, we may have to rethink Asimov's robot law statements, how it could be implemented in a non-industrial environment, where even humans injure themselves unwantedly.

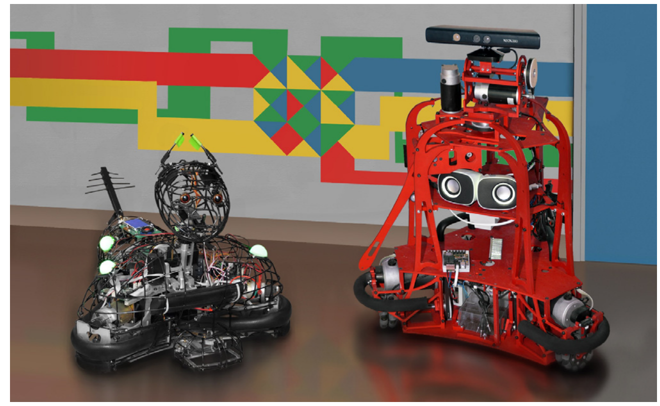

Fig. 26. Example of an ethorobot.

5.16.1. Ethorobotics

As in other fields of technology, the standardization of safety, sensors, actuators, intelligence descriptors, operating systems, middleware is the first thing to achieve. Without proper advancement in standardization, this is not achievable. As a next step, acceptance of robots in early ages (as early as in kindergarten) is desirable. This will result in a native understanding of robots existing alongside with human beings. Along with this, industrial robot revolution will take place, and they will be more autonomous, reconfigurable and understand more of the process they do, than today's sequential program execution. Both the human and animal-like legged locomotion for robots are challenging. From the point of view of technology, we need new locomotion designs, since the mass of existing electric motors are relatively bigger than that of the muscles found in the bodies of humans or animals. The classical model-based control methods are not applicable for sophisticated motion.

Most assistive robots were designed with human-like attributes in mind, but due to present-day technologies, they fail to fulfil the expectations of the users that their human-like appearance and behaviour incites. The solution can be ethorobotics (Korondi et al., 2015).

Ethorobotics (Fig. 26) is a new emerging interdisciplinary field which aims to bring together engineers who are building and programming robots and biologists who are interested in behavioural discipline. Ethology is the biological science of investigating animal and human behaviour in the natural environment. Robotics is slowly reaching a stage where autonomous behaviour and interaction with other robots or humans becomes a reality. Having “behaving” robots means that ethologist are needed both for studying human-robot interaction but also for cooperating in the design and modelling of robot behaviour.

Ethorobotics claims that robots must not be built on any pre-concept of being either human or animal-like but both the embodiment and the behaviour should be derived from the functional demand. This means that the engineers and the ethologists have to determine together (1) the actual environment in which the robot “lives,” (2) the performance which is expected from the robot, (3) the optimal (and most simple) embodiment and behaviour skill which is needed for successful working, and (4) complexity to minimum social behaviour if the robot is working in a human (anthropogenic) environment. The ethorobotic approach also stresses the strong functional relationship between embodiment and behaviour. Embodiment should not elude a capacity which is actually not functional, and behaviour should not go beyond the actual skills needed for good performance. Implying higher capabilities through embodiment or behaviour than what the robot is actually capable of can result in disappointment and a decrease in believability. The ethorobotic approach ensures that the robots are functional, show an acceptable performance, fit in the complex social human environment and are cost effective.

The bio-inspiration for ethorobotics comes from detailed observation of inter-specific interactions in nature. Human-dog interaction provides the most dominant model for ethorobotics because robots in human environment and dogs share many functions. Dogs acquired various social skills during domestication that helps them fit in the human social environment. Dogs excel in helping humans who are sightless, are disabled or suffer from mental disorders, but dogs also assist people living with diabetes or epilepsy. Although it is unrealistic in the short-term that robots could replace dogs in these roles, dogs serve as a very useful model for planning and designing social and assistive robots. Dogs inspire important behaviours with functional consequences for human-robot interaction. These include attachment, social monitoring, gaze contact, simple (non-linguistic) vocal communication, etc. The interaction between humans and dogs presents a different approach from the widely researched human-human interaction centric view. Today social robots are far from reaching the cognitive capabilities of humans, therefore communication between them and the humans should be developed accordingly and be mainly based on simple social behavioural elements. As the research on human-robot verbal communication is still in its early stages, the development of communication should focus on the robot's capabilities to understand simple verbal commands and on the non-verbal aspects of communication. Similarly to dogs, the understanding of simple verbal commands supplemented with contextual and gestural information should be sufficient for communication.

The other role of ethorobotics is to design test beds for detailed quantitative evaluation of human-robot interaction, i.e., benchmarking the service outcome. This approach goes well beyond present day methods that are based on short human-robot interactions and mainly use questionnaires for collecting data. The ethorobotic study of human-robot interaction aims for long (hours, days) interaction, automated data collection on human and robot behaviour and the use of appropriate control observations and benchmarks for performance. This aspect of research is essential in providing more direct feedback to the engineers and the ethologists for improving the robot.

5.16.2. Autopilots and hardware in the loop simulations for fixed-wing UAVs

Autonomous small Unmanned Aerial Vehicles (UAVs) are used frequently in military and civilian applications due to recent advances in communications, battery technology and Micro Electro-Mechanical Systems (MEMS) electronic devices. The core technology is the *autopilot system*.

Successful deployment of small UAVs requires *powerful and lightweight autopilots* with increased autonomy level including path planning, trajectory generation and tracking algorithms. An autopilot is a device able to define and impose the commands that an aircraft has to implement in order to follow a desired flight condition, determined according to the mission requirements. While autopilots are present on commercial airliners to relieve the crew workload, on an unmanned aircraft, they are an essential part of the aircraft Flight Control System. Guided by an autopilot system, an unmanned aircraft is required to follow several waypoints, predetermined or updated in real time. This process is called Navigation. The definition of the aircraft flight parameters needed to approach these waypoints is the Guidance, which varies according to the mission requirements, the aircraft properties and the payload features. The Control process is responsible for maintaining an aircraft attitude that guarantees the predefined flight conditions. Based on the hardware features and limits, the software segment

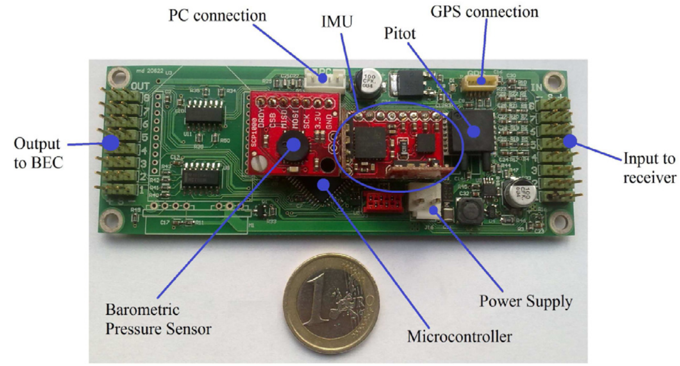

Fig. 27. DIMEAS open-source autopilot.

of the autopilot performs the functions of data analysis, state estimator, path follower and controller which should guide the UAV in flight *without human assistance*.

However, even if the MEMS technology has been significantly improved in the last decades, currently UAVs tend to be complex, expensive and require significant time for completing a task. This observation has motivated the development of a *collaborative control of UAVs*. In particular, the control of multiple UAVs in a swarm, or in a cooperative team scenario, has been a topic of great interest for more than a decade, and it is still growing with the advancements in UAV technologies. A collaborative team of autonomous UAVs may provide more effective operational capabilities to accomplish difficult and complex tasks, compared to the control of a single UAV.

For a formation flight, the fundamental challenges are associated with (i) planning a large team in real time, (ii) developing controllers that are *robust to uncertainties* and are flexible enough to quickly respond to dynamic changes, and (iii) using communication networks to develop cooperative plans. Recent research activities are mainly focused on the second objective, i.e. the development of a robust controller in the presence of model uncertainties arising during platform manufacturing/modelling process, or due to platform geometric and weight variations (such as large variations due to varying payload mass or to geometric inaccuracies) which may occur during the flight tests.

Another critical issue in the study of formation flight is that the aerodynamic coupling may reduce aircraft handling qualities. Thus, the control system designed for a specific flight condition could be inadequate for different operating points.

The *drawback of commercial autopilots* is that they are *not reconfigurable*, and changes of the on-board software are not allowed. In many cases, and especially for small UAVs, automated and systematic approaches for autopilot design are lacking. Autopilots are based on simple single-variable PID controllers. In this case, extensive manual tuning is required for obtaining adequate performance. More recently, customized autopilot systems that enable autonomous flight using an on-board microcontroller and customized on-board algorithms, have been developed. In this way, autopilots can be designed in hours instead of weeks, with great cost reductions, and furthermore the UAV performance is significantly improved.

As an example, a *custom-made autopilot*, designed and produced at the Department of Mechanical and Aerospace Engineering (DIMEAS) of Politecnico di Torino (Fig. 27), is now briefly described. Its main features include an open architecture, the option to be re-programmed during flight and real-time telemetry. Sensors include a GPS, a barometric sensor, a differential pressure sensor and three-axis gyros and accelerometers. The CPU is the AtMega 64A3 model with 64 Kb flash memory and 8 Kb of RAM.

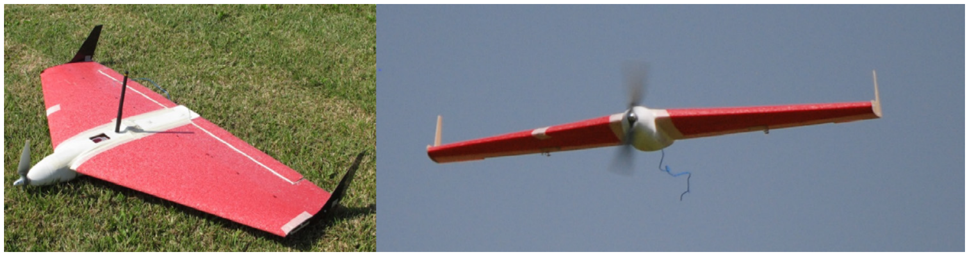

Fig. 28. MicroHawk fixed-wing UAVs of DIMEAS group.

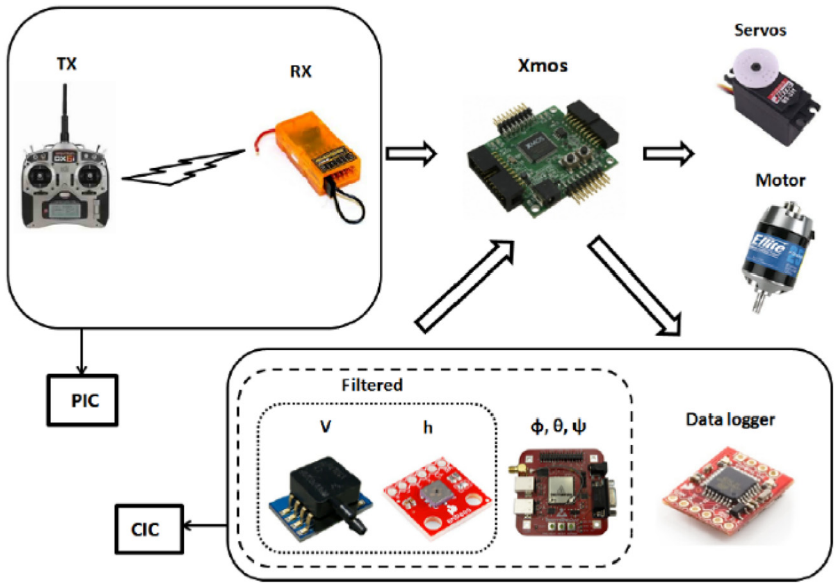

Fig. 29. HIL controller integration scheme.

This autopilot is installed on-board of a fixed-wing UAV (*Micro-Hawk UAV*), developed at the DIMEAS (Fig. 28), to promote innovative scientific techniques for Antarctica exploration and demonstrations within the project ITHACA (Information Technology for Humanitarian Assistance and Cooperation Actions - in cooperation with the UN World Food Program). In particular, this project aims at conducting an intense operational and research activity in the area of geomatics for the analysis, evaluation and mitigation of damages caused by natural or anthropoid hazards.

In the academic community, most of the studies rely on simulations to test the proposed control strategies, and only a few institutions have the infrastructure required to carry on the *experimental testing* of multiple UAVs. The aircraft needs to be piloted during non-controlled manoeuvres (take off and landing, for instance), or as a safety feature to recover from undesired behaviours. In order to reduce the economic costs of a complex multi UAV system, and to avoid crashes of the system during flight tests, *Real-Time Hardware in the Loop* (HIL) simulations are very effective methods for testing the overall control performance and the safety of the systems before conducting actual flight tests. The challenge lies in testing the autopilot features and the closed-loop performance without crashing and damaging the real hardware (Figs. 29 and 30).

If the HIL simulation is embedded in a model-based design process, it can be used during the early aircraft design stages. The HIL simulator is jointly designed with the real plant and can be used by control engineers to test the performance of the control systems. These tests could reveal problems and errors that would otherwise have been detected at the final stages of the design process, when the control system and plant are already integrated. HIL simulation may be embedded in the design process by test automation. The process can be fully automated by including it in the system of the controller design.

When a HIL simulator is equipped with 3D visualization device, which represents the plant very well, it can be also used for training. These HIL simulators are called *training simulators*. Training simulators are very useful when operating on hazardous or very costly UAVs. This allows operators to be trained in a secure environment in both standard and non-standard conditions.

UAVs are difficult to control because their size drives the envelope of aircraft technology, and their aerodynamics data are more difficult to be set. Moreover, they are generally more sensitive to wind gust disturbance than the full-size aircraft. In addition, low cost onboard sensors produce significant sensor data errors and measurement noise. These are some practical observations which suggest to *incorporate uncertainties* in the UAV model. Mini-UAVs need a controller that is robust even under unknown situations. Most of the classical controllers need to be tuned and their performance degrades if the UAV moves away from the predefined trim conditions.

Future trends in autopilot design require to develop controllers that are robust to uncertainties that may occur during the flight tests. The implementation of a systematic procedure and framework of uncertainty modelling, simulation and analysis of UAV flight dynamics for robust flight control design, should be analysed to deal with *real-world challenges* of single and multi-UAV control systems.

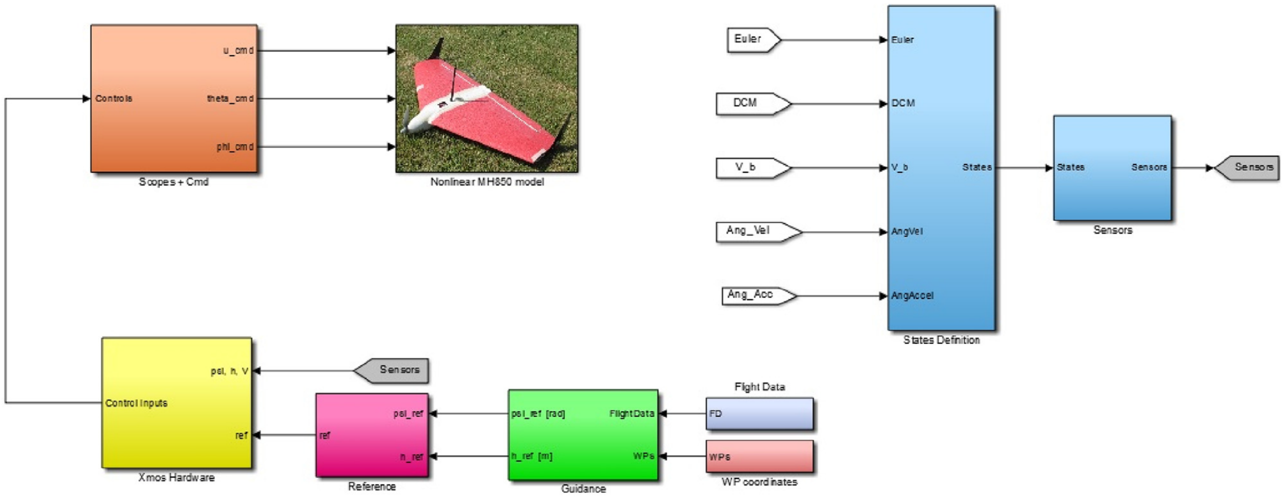

Fig. 30. Example of HIL scheme on Matlab/Simulink.

5.16.3. Robot cooperation

Many robot applications require multiple agents (software and hardware) to provide adequate coverage or a timely response to events. Example multi-UAV applications include environmental disaster relief, urban search and rescue in the aftermath of an attack (chemical, biological, radiological or nuclear), precision agriculture, surveillance and reconnaissance operations, fighting forest fires, and environmental monitoring. There are numerous other applications – in fact any system, such as one that utilizes the service robots in Section 5.15 – for which the timescales of the robot motion are comparable to the timescales of the system events (i.e., the desired response time of a service robot), then using multiple robots could be beneficial.

Depending on the application, these robot teams might consist of homogeneous (all can do every task) or heterogeneous (sub-groups distinguished by different sensors, capabilities, or dynamics) agents. In the latter case, team-sizing and coalition-forming are key parts of the overall problem, especially if the groupings can change dynamically during the mission. The following lists other key challenges in the development of robot cooperation (A Roadmap for U.S. Robotics: from Internet to Robotics 2013; Autonomy Research for Civil Aviation: Toward a New Era of Flight 2017).

Computation and algorithms: The overall planning problem is to obtain proper coordination between the robots to achieve an efficient execution of the mission. This process typically requires ensuring spatial and temporal de-confliction and synchronization of the team while considering mission costs, available resources, and network constraints. The problem usually combines task assignment, trajectory optimization, and obstacle/collision avoidance. The complexity of this combined planning problem scales up dramatically with the number of agents, the number of tasks, the sophistication of the dynamic constraints modelled in the problem, and the degree of coupling that exists in the tasks themselves. Common solution approaches to this planning problem include integer programming, Markov decision processes, game theory and biologically inspired approaches. The primary challenge in all of these approaches is to develop tractable algorithms that provide both performance guarantees and good mission performance on the actual system, which are classical questions addressed by Systems & Control research.

Network connectivity: Multi-robot systems typically rely on communications to operate. Failure to communicate remotely sensed mission data to the base may render the system ineffective, and the inability to exchange command/control messages could lead to inefficiency and/or system failures. The challenge for the Systems & Control field is to develop algorithms that either ensure a sufficient degree of network connectivity is maintained or develop alternative mechanisms for moving the data around in the network. The former requires developing detailed models of the communication environment that can be used to decide when and where to locate the communication nodes in the network. The latter requires a systematic analysis of the trade-offs of utilizing agents for the mission tasks versus using them to “ferry” the data. A further Systems & Control challenge is that the planning, estimation and control algorithms for these networked robot teams must be designed to operate asynchronously and be robust to potentially inconsistent information across the team. Experience has shown that incorrectly addressing these communication issues can have a fundamental impact on the ability of robots to cooperate.

Trust: To collaborate, robots must be able to trust the other agents in team – both robotic and human. This requires development of new levels of health awareness (fault detection, isolation, and recovery) for the robots so that this information can be factored into the planning system. To efficiently cooperate with operators, the humans will need to be able to develop a high level of trust in the robotic system. Achieving human trust in the autonomy will require that the algorithms and software achieve new levels of transparency and system self-health assessment, and that joint operator-autonomy training be performed. In these mixed-initiative human-autonomy teams, the software may also need to develop a level of trust in the operator. This leads to important implications in the design of future two-way communications between the autonomy software and the operator. This must include communication of the current status and future projections. Achieving this goal requires much more information be shared than the current state, and that the information be shared in both directions – and all of that must all be accomplished without overloading the operator.

Sensor fusion: Effective cooperation typically requires that the vehicles have the same situational awareness (i.e., they are “on the same page”). This leads to a challenging problem of developing algorithms that enable each robot to process and model the environment from the streaming data from the onboard sensors and then fuse both the information and models from the other networked agents. Algorithms are needed from researchers in the Systems & Control field to develop new filters that can handle (and then fuse from across the team) continuous variables (e.g., localization with nonlinear measurements and non-Gaussian noise) and discrete variables (e.g., semantic labels and *intents* of the objects that are being mapped). These efforts would improve the mapping capabilities of the team, leading to better navigation capabilities that enhance the degree of cooperation. Further effort is also required to enable the robots to collaborate on active sampling the world to efficiently reduce any uncertainty about the environment. Information-based planning represents a very tight level of coupling in robot cooperation, and is a significant challenge for the control systems community.

Learning and adaptation: Many of these planning systems are model-based and thus can be susceptible to errors in the model (e.g., in the system dynamics, in the noise levels of the perception system) or the problem specification (e.g., the task specifications). Learning and adaptation can be used to improve the system performance, especially in non-stationary environments. The challenges in this case for multi-robot teams is to ensure that the agents can learn efficiently from each other (transfer learning) and to ensure that the learning algorithms scale well with the team size. Significant future effort is also required to develop techniques that can bound the performance of learning-based multi-robot teams.

5.16.4. Soft robot

Soft robots are primarily composed of easily deformable matters such as fluids, gels, and elastomers that match the elastic and rheological properties of biological tissue and organs. Like an octopus squeezing through a narrow opening or a caterpillar rolling through uneven terrain, a soft robot must adapt its shape for a broad range of tasks, obstacles and environmental conditions. Soft robotics is bringing a renewal of robot design: future robots, made of such complex deformable structures composed of stiff and soft regions, open attractive perspectives in terms of new applications, reduction of manufacturing costs, robustness, efficiency and security (due to their compliance, they are much less dangerous than rigid ones when interacting with humans). Controlling such robots in a safe and accurate way would constitute a great jump in robotics in the following years, with applications in surgery, medicine, domestic robotics, game, arts, etc. The lack of control and modelling methods for soft-robots is one of the main obstacles (identified by all surveys in the field), particularly when interacting with a complex environment. Among the main control issues, the underlying physics of soft materials are known to be infinite dimensional (distributed parameter effects) and highly nonlinear when considering composite structures made of stiff and soft elements.

5.17. Control for smart cities

As of 2014, 54% of the earth's population resided in urban environments, with a continuing increase estimated at 1–2% per year. This has motivated cities to look for ways to ensure a sustainable, comfortable, economically viable future for their citizens by becoming "smart." The emerging prototype for a smart city is one of an urban environment with a new generation of innovative services for transportation, energy distribution, health care, environmental monitoring, business, commerce, emergency response and social activities. The technological infrastructure for a smart city is based on a network of sensors and actuators embedded throughout the urban terrain interacting with wireless mobile devices (e.g., smartphones) and with an internet-based backbone with cloud service. The data collected and flowing through such a Cyber-Physical System (CPS) may involve traffic conditions, occupancy of parking spaces, air/water quality information, the structural health of bridges, roads or buildings, as well as the location and status of city resources including transportation vehicles, police officers and health care facilities.

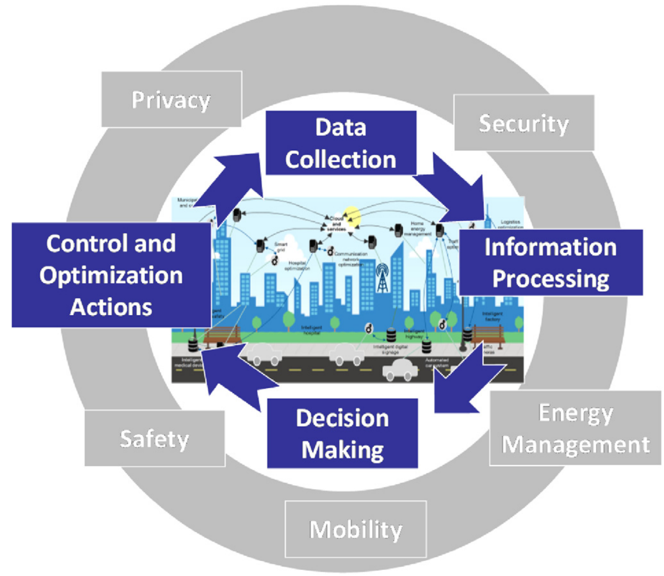

Fig. 31. Smart city cyber-physical infrastructure.

Enabling such a smart city setting requires a cyber-physical infrastructure combined with new software platforms and strict requirements for mobility, security, safety, privacy and the processing of massive amounts of information (so called "big data"). It is important to stress that the ultimate value of a smart city lies in "closing the loop" (see Fig. 31) that consists of sensing, communicating, decision making and actuating – rather than simply collecting and sharing data. This requires a balanced understanding of both "physical" and "cyber" components and the development of new control and optimization methods for this environment. Key components of the research agenda for making smart cities a reality include the following:

- **Sensing and cooperative data collection.** A key challenge is the highly inhomogeneous and distributed nature of the data sources and the sensing devices charged to interact with them and with each other in a cooperative manner. A sensor network may be viewed as a control system encompassing three main tasks: coverage control, data source detection and data collection. The interactions among these three tasks are important and define significant trade-offs.
- **Security, safety, privacy, energy management in the collection and processing of data.** The entire process of data collection and processing is subject to constraints, some strictly physical (such as limited energy) and some imposed by legal, cultural, and economic principles and regulations.
- **Dynamic resource allocation.** Most of the smart city functionality involves managing limited sharable resources, such as transportation capacity, services, water or energy. The highly dynamic nature of the urban environment calls for novel resource allocation mechanisms beyond conventional algorithms.
- **Data-driven control and optimization.** In part due to the availability of enormous amounts of data in a smart city setting, there is an opportunity to develop new schemes for control and optimization which are driven by real-time data as much as sophisticated off line models. In addition, the complexity of the stochastic processes involved in traffic or in the demand for certain resources arguably makes such data-driven mechanisms an absolute necessity.
- **Interdisciplinary research.** Technology alone cannot transform a city without the participation and cooperation of its citizens. A smart city is in fact a socio-technical ecosystem of people,

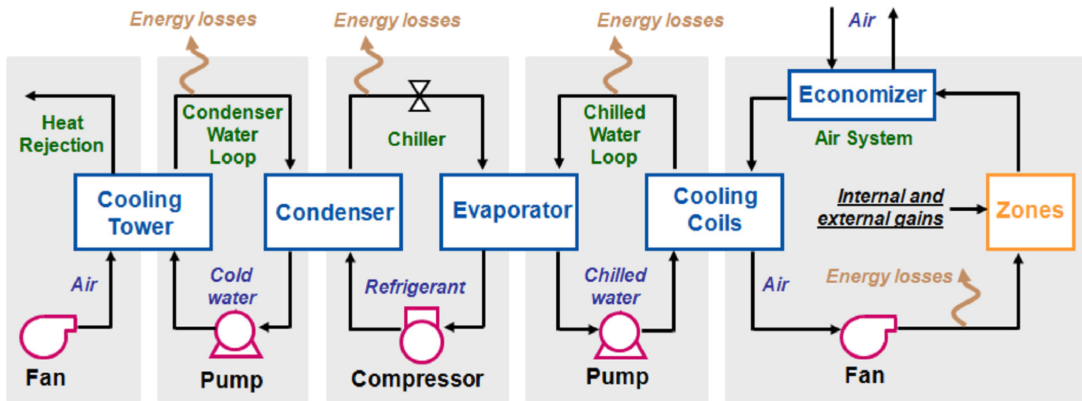

Fig. 32. System view of building HVAC.

technology, organizations and information. As such, the proper design and control of this ecosystem needs to bring together engineers, ecologists, economists, urban planners and social scientists providing a wealth of interdisciplinary research opportunities based on fundamental principles of dynamical systems, control theory, game theory and optimization.

5.18. Advanced building control

Globally, the building sector is responsible for 40% of annual energy consumption and over 30% of all energy-related greenhouse gas emissions; hence, the interest in increasing energy efficiency in buildings remains one of the key drivers in this sector. In addition to innovative building construction materials and design techniques, much can be accomplished by the advanced control and optimization of major building systems, such as the heating, ventilation and air conditioning (HVAC), lighting, renewable generation sources, storage, active façade systems or others (Fig. 32).

Control and optimization can help building owners minimize energy consumption and reduce their utility bills, while maintaining comfort conditions for occupants. But it is important to overcome some of the traditional challenges in the building sector. Sometimes the prevailing focus on short-term reduction of upfront engineering costs prevents implementing solutions that will ensure cost savings in long-term. Consequently, a significant portion of buildings may be equipped with disparate, stand-alone and siloed systems, which are prone to faster performance deterioration.

Building automation systems are continuously evolving to efficiently address these challenges and enable flawless and cost-effective operation of high performance buildings. The most recent trends are influenced by the following aspects.

- **The cloud and data analytics** is one of the most important technology advances over the past few years. The capability to collect data from multiple data sources and move them to a cloud repository enables implementation of powerful applications that may provide insights into building operations. Cloud connectivity enables the retention of more detailed data about the building and this then enables more powerful building analytics that can better inform facility managers about likely HVAC equipment faults, deviations from the expected energy consumption or underperforming controllers.
- **Internet of Things** (IoT) enables connecting building automation components to the IT network and generally improves the interoperability and connectivity of control devices. IoT can help overcome the issue of isolated building systems and support creation of a more cohesive environment. Also, with the advent of IoT, the delivery of comprehensive and cost-effective building automation solutions will potentially require fundamental changes to how systems are designed and installed. New types of more intelligent devices and systems will be required that collect and store data directly in the cloud, where they can be used by advanced applications.
- **User experience** aspects play an increasingly important role in the design of new applications that take advantage of connected equipment, devices and automation systems whose data can be shared with a wider audience. Two main categories of users are building occupants and people involved in the facility management. New types of applications and user interfaces are delivered via smart phones and tablets that can provide multiple real-time functions such as secure monitoring of equipment operation, changing setpoints, viewing and acknowledging alarms, or adjusting schedules.

Challenges and opportunities for advanced building control: Today's buildings are complex environments whose operation is affected by multiple dynamic factors. When considering systematic implementation of advanced control methods, one particular challenge consists in the continuous degradation of building systems. Mechanical malfunctions, sensor drifts and inappropriate configuration parameters should be corrected before introducing more sophisticated control techniques.

Some of the new opportunities for building control include:

- **Flat IoT control architecture**. Traditionally, the HVAC control systems are designed in a bottom-up manner that introduces several hierarchical layers. The bottom part is represented by individual single-input-single-output (SISO) control loops, which are then coupled by a higher level logic, residing in plant or supervisory controllers. However, this state-of-the-art is increasingly impacted by the proliferation of the cloud, open architectures and IoT technologies that support separation of typical functions into only two levels: intelligent edge devices and the cloud. This may imply that plant and supervisory controllers may not be further needed under this scenario. The base level closed-loop control functionality will be implemented through a flat architecture of multiple cooperating edge devices, while the supervisory functions will be pushed completely to the cloud environment. This concept can be significantly cheaper to deploy, but the overall impact on the performance of such control architecture still needs to be explored.
- **New ways for balancing comfort with energy costs**. With the increased emphasis on user experience and people's productivity, the thermal comfort in buildings should be maintained in a way that satisfies the maximum number of occupants. This can be achieved by allowing individuals to define their personal comfort preferences and provide immediate feedback on

the current comfort conditions. Then new algorithms will be needed to aggregate and properly process all such inputs from occupants and determine new crowd-sourced setpoints in the most cost-effective way.

- **Distributed approach to whole-building optimization**. Economic optimization of all building energy systems can be formulated at the whole-building level to integrate all subsystems such as HVAC, lighting, onsite generation, and storage. The implementation of this approach is complicated by disturbances, such as weather conditions and occupant behaviours, and potentially also by dynamic pricing of electricity. However, the fundamental issue lies with building-wide optimization models, which will always be hampered by significant inaccuracy, uncertainty and lack of data measurements. Distributed optimization approaches could be more viable; these would first divide the building into meaningful sub-systems and then optimize each sub-system locally but not independently of others.
- **Multivariable HVAC supervisory control**. The primary goal of HVAC control is to maintain occupants' thermal comfort and system energy efficiency. This requires adjustments of multiple setpoints—primarily temperatures and flow rates. Today these setpoints are either kept constant or manipulated by simple reset rules. An obvious opportunity exists for new robust multivariable supervisory control strategies that will leverage principles of model predictive control (MPC) to dynamically adapt key HVAC setpoints based on weather conditions, occupancy, and actual thermal comfort in zones. The challenge of developing reliable HVAC models for MPC might be addressed by moving the optimization engine to a cloud and coupling it with efficient analytics for identification of suitable models from the HVAC data.
- **Building-to-grid-integration**. Recently, demand response (DR) has been recognized as a promising approach for the electricity market and an essential element of smart grid implementations. By sending changing power-price signals to building automation systems, adjustments of temperature setpoints, cycling of HVAC equipment or other actions can be initiated, and consequently energy use and expenditure can be reduced. A fundamental challenge is to enable the building to participate in DR without violating thermal comfort. Advanced control strategies are needed that will manage building loads and use the building's thermal mass to implement pre-cooling or pre-heating strategies and adapt zone temperature trajectories. In addition to dynamic load management, in many cases, the scope of optimization could also encompass local generation and storage devices.

5.19. Nanoscience and quantum engineering

Systems & Control can lead to significant improvements in emerging quantum technologies ranging from magnetic resonance imaging (MRI), inertial navigation systems, optical communications to high precision metrology, quantum communications and circuits and prototypes of quantum computers. For example, the control of the motional state of trapped ions, the internal state of atoms or the quantum field in an optical cavity is commonly performed in the lab nowadays.

This is a result of the pioneering work of experimental teams such as S. Haroche's and D. J. Wineland's who were both awarded the Nobel Prize 2012 for ground-breaking experimental methods that enable the measurement and manipulation of individual quantum systems. Solid-state setups like the SQUID (Superconductor QUantum Interference Device) were also investigated, as well as hybrid systems which combine atoms, molecules or quantum dots with superconducting cavities (Fig. 33).

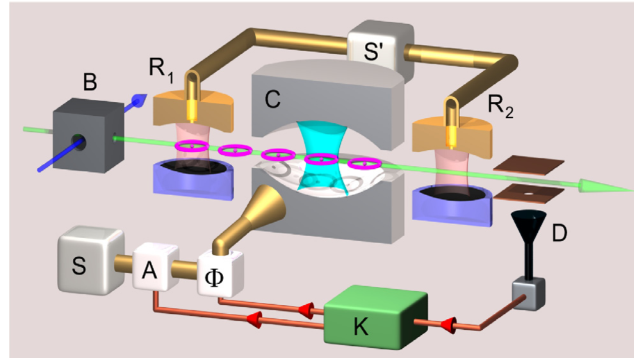

Fig. 33. Experimental methods enabling the measurement and manipulation of individual quantum systems.

Extensions of traditional control concepts developed for classical systems, such as optimality, feedback, stability, robustness, filtering and identifications to quantum systems are becoming key issues. As an example, feedback control admits two different quantum counterparts: measurement-based feedback where the controller is a classical object and coherent feedback (or autonomous feedback) where the controller is another quantum system coupled to the system of interest. This multiplicity results from a fundamental difficulty that must be overcome by any quantum extension of classical feedback, namely the random back-action on the system caused by sensor measurements. Similarly state estimations (quantum tomography) and system identifications (quantum process tomography) face related difficulties leading to limits in the precision imposed by the Heisenberg uncertainty principle. Future technological developments exploiting quantum features will have to include some robust stabilising mechanism to protect their fragile quantum states against decoherence due to environment coupling.

Most systems are composites built with several sub-systems. The quantum states of such composite systems live in the tensor product of the Hilbert spaces of each sub-system. This is a crucial difference with classical composite systems where the state space is built with Cartesian products. Such tensor products have important implications such as entanglement with existence of non-separable states that are crucial resources for quantum communication and cryptography. These quantum technologies will rely on control concepts redesigned for the quantum world, in particular for composite systems and networks.

Scheme of the first experimental quantum feedback loop. The system consists in micro-wave photons trapped between two super-conducting mirrors forming the cavity C (blue cone). The measurement process (sensor) relies on probe atoms (pink torus flying from left to right). They are prepared in B, interact with the system during their passages between the two mirrors of C, and are detected in D. The actuator is the classical source of photons S with tunable amplitude A and phase Φ. The controller is implemented in the classical computer K. It provides in real-time the quantum state ρ (density operator of the micro-wave photons) via a quantum filter and defines the actuator values A and Φ via a quantum-state feedback. The feedback law, A(ρ) and Φ(ρ), stabilizes a prescribed photon number state (the set-point).

In the September 2011 issue of Nature, the first experimental implementation of a quantum feedback loop was reported. It has been realized in the group of Serge Haroche (see Fig. 34). It was the first time that a full quantum state, involving not only the populations but also the coherences, was computed in real-time and exploited in a feedback loop. The control goal consists in stabilizing light around photon-number states (Fock states). These quan-

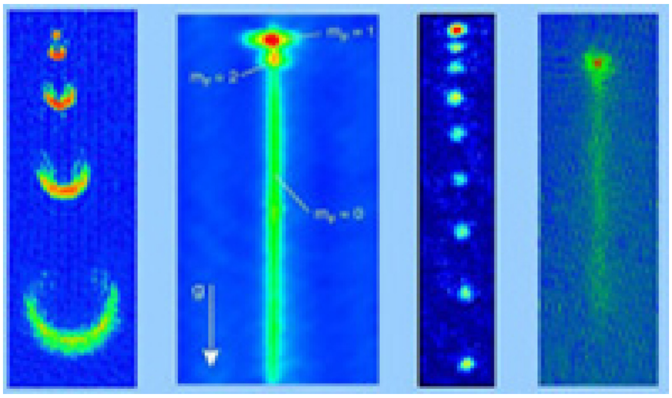

Fig. 34. Experimental implementation of a quantum feedback.

tum states are very different from classical ones describing usual light: they are fragile and difficult to generate and stabilize. The interest, but also the difficulty of this experiment, lies in the fact that the measurement process, present in any feedback loop, induces unavoidable random perturbations on the system to be controlled. The feedback algorithm used for this experiment relies on a quantum adaptation of Lyapunov control techniques. This kind of feedback where the controller is a classical system corresponds to measurement-based feedback.

5.20. Social and techno-social networks

The systematic study of social networks, or Social Network Analysis (SNA), takes its origins in 1930s, but was, in fact, anticipated and inspired by earlier contributions on social philosophy and psychology, published at the turn of 19th and 20th centuries. SNA, which has now grown into a large interdisciplinary area, inaugurated a new paradigm in social and behavioural sciences. Unlike "individualistic" social theories, which consider individual choices of social actors, this new theory focused on examination of social relations, influences, structures and movements (i.e. group actions). Subsequently, the term *social network*, which is a central concept to SNA, was coined. Social network is a structure, consisting of social actors (either individuals or organizations) and *ties* between them, which stand for social relations or interactions.

Representing social networks by their interaction *graphs*, SNA efficiently employed many graph-theoretic tools. The symbiosis between mathematical (and later, algorithmic) graph theory and SNA provided many computationally efficient algorithms for analysis of large-scale systems (e.g. graph density and clustering coefficients, cliques, homophily and cohesion, measures of centrality, including closeness, betweenness and eigenvector). The statistical theory of complex networks and physical processes over them (Newman, 2003) was strongly inspired by the "small-world effects" in social networks that are prominently illustrated by the widely known "six degrees of separation' ' theory. We also notice that statistical properties of complex graphs, in particular the World Wide Web (WWW), lie at the heart of efficient algorithms for node ranking (e.g. the PageRank paradigm) and fast web search.

On a parallel path, more than 60 years ago, Norbert Wiener introduced the general science of cybernetics, with the objective to unify systems, control and information theories. In the context of social sciences, Wiener argued that "*society can only be understood through a study of the messages and communication facilities which belong to it*" (Wiener, 1954). The development of social and behavioural science in the 20th century confirmed the key ideas of Wiener, which more recently led to the increasing comprehension that "*the foundational problem of sociology is the coordination and control of social systems*" (Friedkin, 2015). Nevertheless, the well-developed theory of SNA, extensively adopting tools and ideas from applied mathematics, statistical physics and computer sciences, still has little intersection with Systems & Control. In fact, the realm of social systems remains a key challenge for modern Systems & Control science despite rapid developments and impressive achievements (Murray, 2003; Samad and Annaswamy, 2014).

The most important reason for the gap between systems and social sciences was perhaps the lack of mathematical models representing a social group as a *dynamical system*. Focusing on the topological properties of social networks, SNA paid much less attention to dynamics over them, confining to special processes, fully determined by the topology, such as random walks, epidemic spread or percolation phenomena (Newman, 2003). On the other hand, dynamics were very limited in the tools for mathematical analysis and numerical simulation of large-scale social groups. Recent years, however, have witnessed substantial activities towards the creation of *dynamic* social network analysis, which were opened up by the general rapid progress in the study of complex systems and the development of algorithms and software for their analysis. A closer examination of complex networks, including those arising in nature, economics and industry, has revealed some common principles of their coordination and self-organization (e.g. consensus protocols for distributed decision making, synchronization of coupled oscillators and coordinated motion of birds in a flock are governed by similar models). The discovery of these analogies, on one hand, attracted considerable attention of many research communities to models of social groups evolution. Even for groups of small size these models exhibit rich and non-trivial dynamics, which may give a clue to the solution of hard algorithmic problems (such as clustering) and evolution of complex natural systems, which exhibit persistent disagreement and other "irregular" behaviours. On the other hand, many models of social dynamics were *inspired* by complex processes studied in various applications (Newman, 2003).

The number of models which describe social groups dynamics and primarily deal with the evolution of individuals' opinions is currently growing (Friedkin, 2015). Even commonly accepted and experimentally validated models, such as bounded confidence models by Hegselmann–Krause and Deffuant–Weisbuch or the Friedkin–Johnsen model, are still far from being deeply investigated from the Systems & Control-theoretic viewpoint. In particular, some basic questions concerning their identification, controllability and robustness still remain unsolved. Furthermore, how useful these models are in describing the behaviour of large groups in real social networks, is a widely open problem. Its solution may bridge the gap between the "dynamic model-based approach" provided by the models previously discussed and the "data analysis approach" proposed by other scientific communities such as computer science or physics.

Much more challenging, however, are the properties of Techno-Social Networks, emerging from the interplay between technological and social networks (Fig. 35). Besides usual personal interactions among social actors (such as face-to-face meetings), the recent progress in communications and networking opened up the possibilities of virtual (online) interactions via instant messages, chats and other microblogging tools. Virtual interactions enable highly asynchronous, flexible and heterogeneous opinion formation processes, where individuals may discuss simultaneously several interrelated topics in different forums, getting immediate response and *feedback*. The downside of virtual interactions are new threats, such as malicious attacks, spamming and dissemination of extremist doctrines. The relevant effects and dynamics are uncovered by the existing models, and require the creation of a new theory at the "crossroad" of engineering, computer science and social sci-

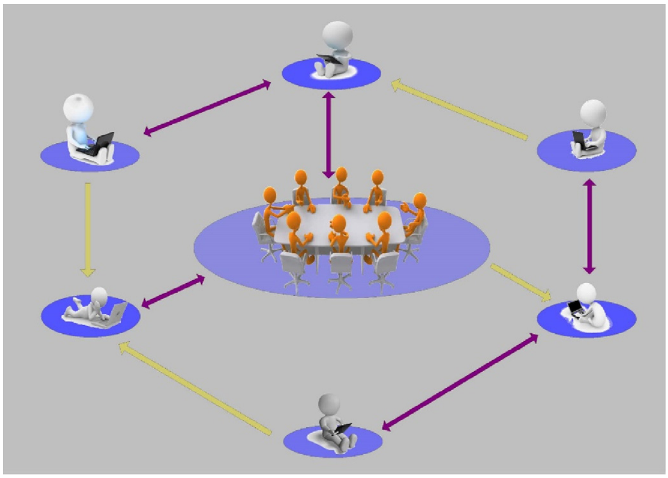

Fig. 35. Face-to-face meeting and virtual interactions in a techno-social network.

ences, where mathematical methods of Systems & Control theories will undoubtedly play a major role.

The challenges of Techno-Social Networks are numerous, and many of them are now briefly outlined. It seems that these challenges may be labelled as *Dynamics over Techno-Social Networks* (DTSN). The theory of DTSN will focus on the study of both *topological* and *dynamical* properties of the network, especially those related to its resilience (robustness against failures of nodes, links and malicious attacks at them). Unlike the classical SNA and existing results of complex networks, DTSN has to deal with graphs that are not only large scale, but also evolve over time, so that nodes and links are constantly created and deleted. An important issue is elaboration and empirical validation of opinion formation models over such networks, allowing the social actors to interact *asynchronously* and simultaneously communicate on interdependent topics in several forums.

A list of some future challenge candidates includes:

- analysis of belief system dynamics with interrelated topics and logic constraints
- study of the "temporal" communication graphs whose nodes and links may continuously appear and disappear;
- new concepts of centrality measures in temporal graphs and efficient distributed algorithms for centralities computation;
- new models of node and link ranking (e.g. extension of the PageRank paradigm), and relevant algorithms (these algorithms should ideally detect "techno-social spamming" when individuals, or machines, are fictitiously broadcasting artificial connections);
- analysis of the mathematical relations between the models of opinion dynamics and the models for node and link ranking;
- algorithms for counteracting malicious individuals, sending erroneous messages (extensions of the Byzantine agreement approach studied in computer science to deal with resilient quantized consensus);
- analysis of the effects of combined communication failures and malicious attacks;
- asynchronous and spontaneous interactions among individuals connected on the web.

The theory of complex networks has recently attracted the attention of many research communities in the fields of applied mathematics, engineering, natural, social and behavioural sciences, economics and finance. The key to robust and sustainable functioning of many networked systems arising in these areas, is the synergy between dynamical models and feedback, which is a very distinctive feature of the Systems & Control theoretic approach. This approach is now proving its success and impact in many real world applications. These major achievements give a hope to the future success of Systems & Control theoretic methods in analysis and control of social and techno-social networked systems. This progress, however, requires further convergence between Systems & Control, engineering and social sciences, to facilitate the creation of new dynamic models, and the development of efficient data analysis algorithms for their testing.

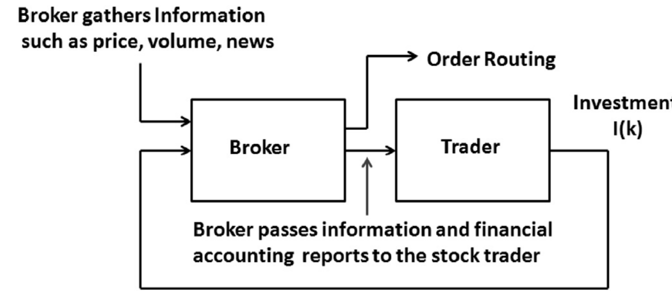

Fig. 36. Feedback loop for trading.

5.21. Control and model-free stock trading in financial markets

In this sub-section, an overview of a new application area involving the use of tools from classical control theory in the context of stock trading in financial markets is provided. For a detailed survey of this research area, the reader is referred to the March 2016 issue of the IEEE Transactions on Automatic Control where the relationship between this application and results in the literature, both in finance and control, are discussed. In Fig. 36 below, a high-level feedback system for stock trading is depicted. The motivation for the research area underlying this application comes from the fact that dynamic models for stock prices are not notoriously unreliable. This is epitomized by the two market crashes over the decade covering 2000–2010. These clearly illustrated deficiencies in existing theory.

During these crashes, dramatic and sudden changes in market volatility and asset correlations rendered classical stock-trading price models rather useless. The new applications and associated theory being reviewed here are based on fundamental ideas about adaptation in feedback loops. The work concentrates on regulation of a portfolio's gains and losses with the attainment of «robust performance» as the goal. This means that the portfolio's value is increasing over time without large «drawdowns».

Model-free aspects: Based on profits and losses, investment levels are adaptively adjusted by a controller without using any type of stock-price model. This is made possible by having the amount held in various assets adaptively modified based on observed performance. We call such a scheme *performance-driven model-free asset management*. Simply put, we use a set of «rules» for investment level updates rather than a predictive model. To provide an analogy for the reader uninitiated in finance, imagine the problem of bringing bathtub water to a desired temperature set point. For a feedback-control theorist, the fact that a thermodynamic fluid model may not be needed is second nature. With adequate sensing and actuation, the following simple feedback rule will often suffice: «If the bathtub water is too hot, add cold water. If the bathtub water is too cold, add hot water». For the case of stock trading, our view of asset management is much the same as the one given for this bathtub scenario. Whereas the hot-cold water inflow for the bathtub is regulated in a reactive manner without a predictive model, for the case of investment, the inflow or outflow of investment dollars into portfolio components is regu-

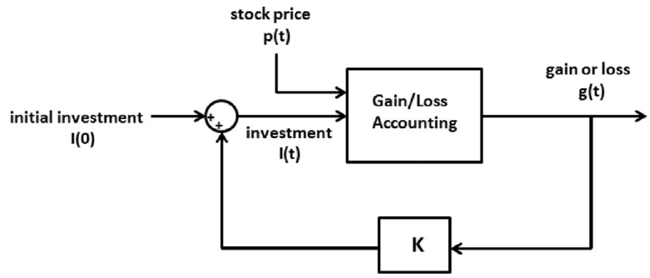

Fig. 37. Stock trading via linear feedback.

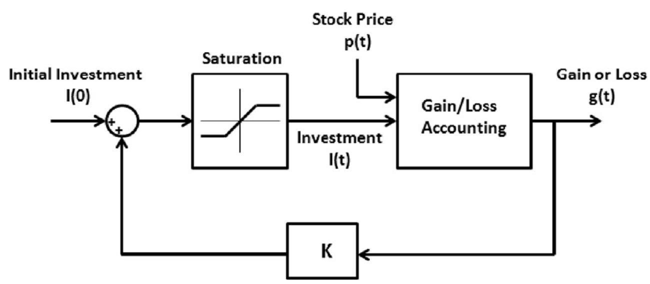

Fig. 38. Stock trading with saturation.

lated in a reactive manner without regard for prediction of future prices.

Simple linear feedback example: A simple closed-loop system is obtained by allowing the instantaneous investment amount to be a feedback on the gains or losses. The implementation for the special case of a simple linear static output feedback is shown in Fig. 37.

For such a controller, an upward price trend increases the profit which in turn results in an increased investment. Similarly, a price decline reduces the profit leading to a reduced investment. This strategy might be appropriately be called a «trend follower» because it will adaptively grow the investment amount to «catch» an upward trend and attenuate losses on a stock price decline. Similarly, by using negative feedback gains and «shorting» a stock, one captures a downward trend and reduces losses on a price rise.

More general controllers: The use of a pure gain above is just one of many possible control structures which one can use to determine the investment level. For example, one can add memory to the controller via a PI configuration and considerations for broker-imposed leverage constraints can be handled via inclusion of a saturation device in the forward path; e.g., see Fig. 38.

Simulation and back-testing with historical data: Ongoing research in this application area includes a significant computational component. To benchmark the controller, we typically carry out two types of simulation. The first type of simulation involves synthetic price data. In this regard, we often use the most famous class of prices in finance: those which are obtained as sample paths of a Geometric Brownian Motion. This class serves as a «proving ground» within which performance can be assessed. The second type of simulation involves a back-test using historical data. To illustrate what is meant by this, we begin with Fig. 39 where closing stock prices for Facebook (FB) are given covering the two years period 2013 and 2014.

Using one of our feedback controllers called «Simultaneous long-short» we simulate daily trading over the two-year period and generate a plot of the account value as a function of time. Beginning with an account value of $10,000, and investment limited to $20,000, using forward, following the initial period of decline corresponding to the time period immediately after the initial public offering, we largely encounter an upward trending market. For the

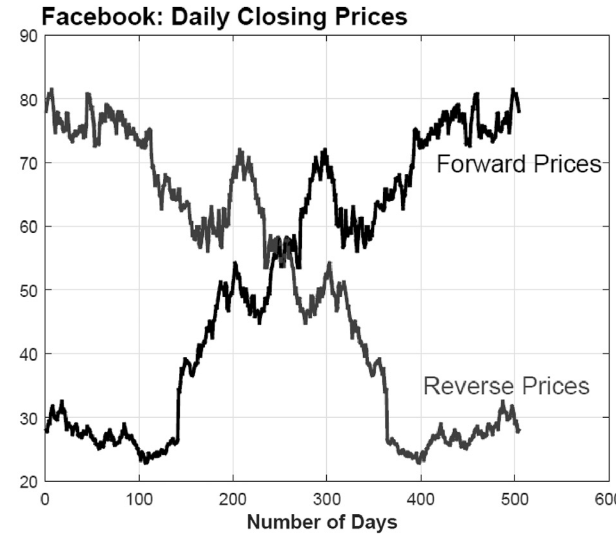

Fig. 39. Facebook daily prices in 2013 and 2014.

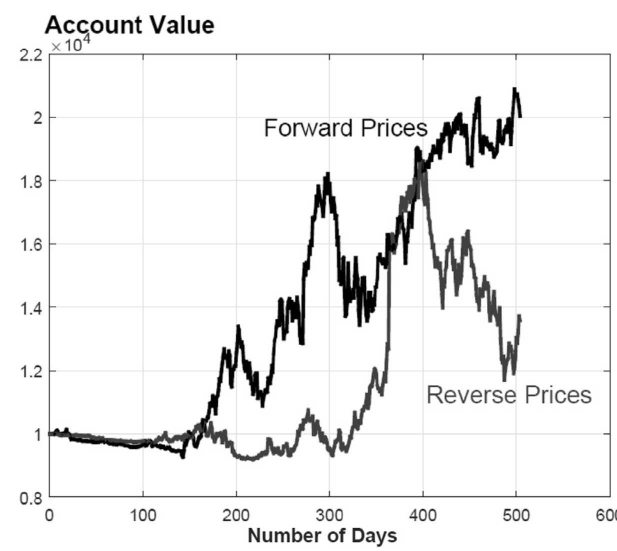

Fig. 40. Account value evolution for Facebook.

second test simulation, we use the reverse-order prices which are seen to be downward trending. One of our objectives here is to demonstrate that the controller is «smart enough» to adapt to either market direction. For these two cases, Fig. 40 shows evolution of the account value.

Directions for further research: The discussion above involves trading a single stock. Ongoing work in this area involves extending this feedback control framework to trade a portfolio consisting of many stocks. In this setting, there are many exciting adaptive control problems which would be of interest to consider. For a portfolio, we envision dynamic adaptation of a set of feedback gains which essentially serve as the weights for portfolio components. That is, the *i*th feedback gain tells the investor what percentage of the amount invested should be allocated to the *i*th stock. The formulation of this «gain selection problem» can provided in terms of optimization theory. That is, given a set of performance and risk metrics, one formulates an optimization problem, perhaps even convex, whose solution provides the desired portfolio weights above.

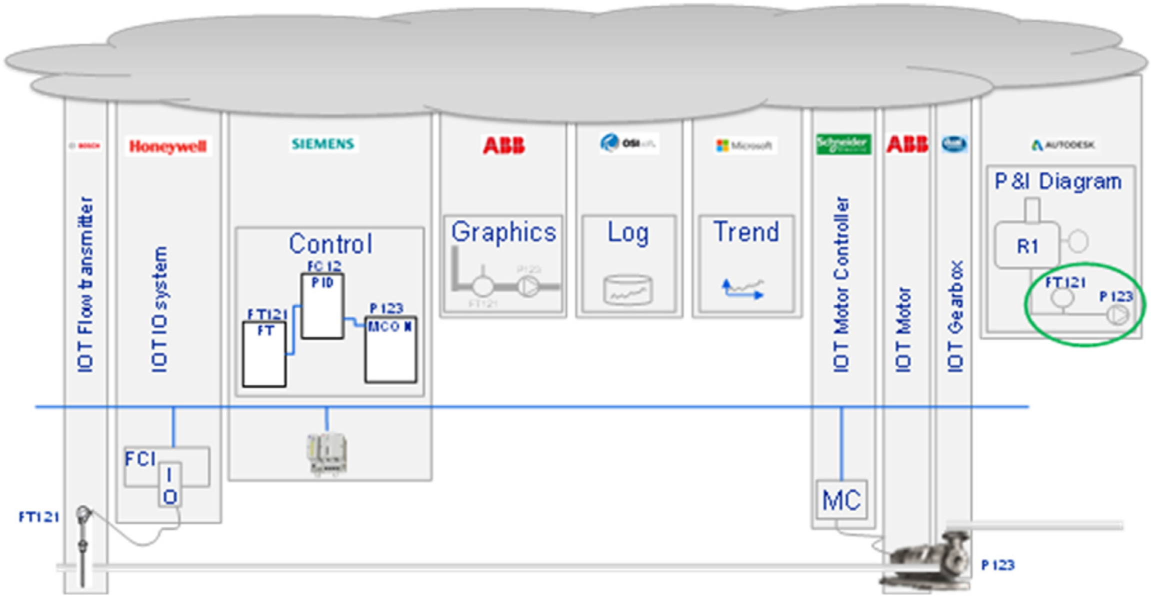

Fig. 41. IoT for process automation will require interoperability between different vendors and a trade-off between fog/edge and cloud computing.

5.22. The role of control for IoT[1]

In this “first wave” of Internet of Things (IoT), attention has concentrated on wireless sensors, cloud connectivity, big data analytics and mobile apps. The concept of IoT, however, extends beyond these components and capabilities (see Fig. 41). For example, a white paper by the IEEE Internet of Things Initiative defines a much more expansive vision (Minerva, Biru, & Rotondi, 2015):

“Internet of Things envisions a self-configuring, adaptive, complex network that interconnects ‘things’ to the Internet through the use of standard communication protocols. …The things offer services, with or without human intervention, through the exploitation of unique identification, data capture and communication, and actuation capability.”

This “actuation capability,” especially “without human intervention,” needs additional tools and research. Indeed, most definitions of IoT are from the perspective of information and communication technologies (ICT), but closed-loop control in any context is not just, or primarily, an ICT challenge. Deep understanding of dynamics and control is essential. Feedback can qualitatively change the behaviour of a dynamical system, for better or worse. A seemingly benign system can become unstable if feedback is inappropriately applied, and, on the other hand, automatic feedback control can enable unstable systems to reach levels of performance unattainable by stable systems. The closed-loop integration of physical systems with the internet will require close collaboration between control experts and ICT experts.

Topics for control research. IoT also promises new vistas for the control research community. The fact that aircraft, cars, refineries, buildings and medical devices function as well as they do is testament to the power and maturity of control science and engineering. But it’s worth noting a few assumptions on which this success rests. The communication networks in control systems are generally assumed to be deterministic and reliable. Real-time operating system platforms rely on predetermined, static schedules for computation and communication. Some control is now occurring over the internet, but at a supervisory level—for power grid distribution stations, wastewater treatment plants, some commercial buildings and other applications. Closed-loop automation, more often than not, requires a dedicated, on-site, end-to-end control system.

Control in the Internet of Things imposes control-theoretic challenges that we are unlikely to encounter in our usual application domains. More research is needed in a number of areas, including the following:

- **Control over nondeterministic networks**. Today’s control systems assume deterministic communication and computation—in fact the execution and communication infrastructure is rigorously designed to ensure determinism. Nondeterminism—e.g., unpredictability in sensor reading, packet delivery, or processing time—complicates closed-loop performance and stability.
- **Latency and jitter.** Control over the internet and clouds will require much greater attention to latency (the end-to-end delay from sensor reading to actuation) and jitter (the variance in the inter-sampling interval). The techniques used in control applications today to deal with these phenomena are unlikely to suffice.
- **Bandwidth**. Many control applications are not demanding of communication bandwidth—a few sensor reads and actuator outputs per second can suffice. But even this level of network performance may not be assured with mobile and/or internet connectivity. Furthermore, in the Internet of Things, closed-loop control with feedback of video and other high-dimensional data is envisaged. The sophisticated signal and image processing algorithms involved will best be run on cloud platforms and will stress available bandwidth.
- **Cyber- and physical security, and resilience.** The physics of the “things” in IoT, if appropriately incorporated, can enhance detection and protection approaches for both cyber and physical security. Conversely, physics and feedback can open the door to new attack scenarios: e.g., a well-performing control system may be rendered unstable by introducing small delays in communication pathways.
- **Interoperable and plug-and-play sensors, models, and algorithms**. With our digital devices and platforms we have become accustomed to features such as auto-discovery, search, composition of services, and plug-and-play integration. These are not as yet available for control applications. To get there, interop-

[1] Parts of this section appeared previously in (Samad, 2016, Samad, 2016); copyright IEEE.

erability will need to extend beyond the interface specification; "dynamic" compatibilities will also be critical.

Related IoT enhancements. Today's IoT infrastructure places limitations that new theoretical and algorithmic developments by the control community can only partially overcome. Available component technologies and the IoT stack as often envisioned fundamentally limit the potential for advanced control. Research in IoT technologies is targeting these limitations and will also open the door for closed-loop control, especially for high-bandwidth, highly reliable, and high-performance applications:

- **5 G networks.** Cellular communication technology has progressed by "generations" but the next advance is seen as a "paradigm shift." Dramatic enhancements in bandwidth, flexibility, and intelligence are foreseen, with data rates two to three orders of magnitude greater than 4 G systems.
- **The tactile internet.** Round-trip latencies with wireless communication are not currently low enough for many real-time control applications. A key threshold is seen as 1 ms, at which point human-in-the-loop wireless control becomes feasible. This era of the "tactile internet" is expected to open up a vast space of new closed-loop applications.
- **Fog computing.** The cloud is a central element of today's IoT stack, but a critical bottleneck for reliable real-time control. "Fog" or edge computing architectures enable processing to occur closer to the sensors and actuators, with advantages of speed, security, reliability and efficiency .

Conclusion. Control expertise will be required to realize the visions of IoT that we, its proponents, are promising. At the same time, IoT brings new and exciting opportunities for research and development in control. Research in IoT platforms and technologies is also targeting enhancements that will provide the infrastructure necessary for supporting advanced real-time closed-loop applications.

To illustrate, here are some prospects that can motivate collaborative research:

- Systems that are not physically connected or co-located could be coordinated in real time;
- Optimized performance (e.g., energy efficiency) could be achieved for small-scale systems that cannot afford dedicated control systems;
- High-fidelity models could be widely applied for real-time control via IoT implementations;
- Global networks of sensors and actuators could be implemented and coupled with sophisticated control and optimization algorithms;
- Greater redundancy and fault-tolerance could be achieved across critical infrastructures.

There is much to be done before the full IoT vision can be realized and control engineers and scientists have a critical role to play.

6. Operational recommendations

Funding agencies are a key tool for fostering these new research and development. However there seems to be a trend of government and funding agencies moving towards favouring funding areas that can show a direct and more-or-less immediate effect on technology development and successful applications with direct economic impacts. While Systems & Control science is at the heart of these new multidisciplinary developments, it is often hard to explicitly recognize and show the importance of our field outside our community. Therefore, to remain successful in the quest for funding support, the community will need to position itself in a strategic way and may need to seek stronger connections with neighbouring societies and disciplines, while at the same time pursuing development of tools and methods for addressing these next-generation Systems & Control problems. In the following some recommendations are listed in order to provide the means to develop this extremely important scientific and technological discipline whose critical role in ICT is essential in the future:

➢ **Pursue all recommendations from Murray (2003)**

- Substantially increase research aimed at integrating control, computer science, communications and networking. This includes principles, methods and tools for modelling and control of high-level, networked, distributed systems, and rigorous techniques for reliable, embedded, real-time software.
- Substantially increase research in control at higher levels of decision making, moving toward enterprise-scale systems. This includes work on dynamic resource allocation in the presence of uncertainty, learning and adaptation, and artificial intelligence for dynamic systems.
- Explore high-risk, long-range applications of control to new domains such as nanotechnology, quantum mechanics, electromagnetics, biology and environmental science. Dual investigator, interdisciplinary funding was suggested as a particularly useful mechanism in this context.
- Maintain support for theory and interaction with mathematics, broadly interpreted. The strength of the field relies on its close contact with rigorous mathematics, and this was felt to be increasingly important in the future.
- Invest in new approaches to education and outreach for the dissemination of control concepts and tools to non-traditional audiences. The community should do a better job of educating a broader range of scientists and engineers on the principles of feedback and the use of control to alter the dynamics of systems and manage uncertainty.

➢ **Substantially support the seven key research and innovation challenges:**

1. Distributed networked control systems
2. Data-driven dynamic modelling and control
3. Complexity and control
4. Critical Infrastructure Systems
5. Cyber-Physical System of Systems
6. Autonomy, cognition and control
7. Cyber-Physical and Human Systems

➢ **Support basic research and research focused on application domains**

Systems & Control has two facets: research in basic principles, theories, and tools, and research related to specific application domains. The strength of the discipline is the interplay between these two sides: while research related to application domains provides new solutions to pressing problems in these areas, it also generates new approaches, theoretical results and tools that can be transferred to other domains, as well as challenges to fundamental research. This fundamental research provides the sound basis on which technical solutions with guaranteed properties can be developed, independent of specific applications. Funding is necessary both for fundamental and applications-oriented research, and both in sector-specific programs and in ICT as a program that provides enabling technologies for all sectors.

➢ **Overcome the barriers between the traditional disciplines.**

The application of Systems & Control requires knowledge of application domains as well as its theoretical foundations. The full strength of the discipline is made effective through close collaborations with researchers and developers in the application domains. While this has already been achieved in many application areas, e.g., mechatronics and process control, a qualitative leap is necessary in the application of Systems & Control to sectors and phenomena where it has not yet been adopted, such as health care, biological systems, social systems and development of large infrastructures. Further development of Systems & Control also requires close collaboration with other disciplines that study complex systems – complexity science, cybernetics, synergetic, networks science, artificial intelligence – to mention a few. Collaborations with researchers in these areas would mutualise the forces and approaches used for analysis of complex systems, thus closing the gap between different disciplines in order to overcome existing compartmentalization of the science and use existing tools in more effective ways.

7. Appreciation of Systems & Control by industrials

"Controls is a cornerstone technology to make the new Industrial Internet era a reality. It provides the unique capability to change and optimize the behaviour of machines and systems at all levels of business and society. As the world becomes more interconnected, uncertain, and dynamic, we will increasingly turn to controls for robust solutions." *Brent Brunell(*), Technology Leader Connected Controls, GE Global Research.*

"The hidden capability of control system technology definitely settled a transformation having been moved from the status of "service" for a corresponding technological solution to the status of "product" being nowadays a fundamental brick for the optimization of investment indices like capital and operating expenditures. In the very next future, the control role is intended to further get transformed being now an unavoidable and transversal element in all the industrial sectors that, in turn, are presently subject to an exponentially increasing complexity." *Francesco Cuzzola(*), Executive Manager Danieli Automation SpA*

"Systems & control are the backbone of our operations environment, providing tools/technologies/insights enabling smarter and more economic operation while reducing the eco footprint." *Alex van Delft(*), Corporate Manager Process Control, Royal DSM NV, Corporate Operations, The Netherlands.*

"Real-time measurement, modelling & control platforms will drive a smarter planet through the broad Implementation of feedback control." *Darío Gil, Program Director, Energy Technology and Strategy IBM T.J. Watson Research Center.*

"The role of a control system is to ensure engineering innovation at the lowest cost with the highest quality." *Maryam Khanbaghi(*), former manager of advanced control system at Corning Incorporated, USA.*

"The systems, control and automation disciplines play a critical role in maximising value from the emerging and increasingly complex integrated cyber manufacturing systems." *Michael Lees(*), Process Control and Automation Manager, Carlton & United Breweries - Yatala plant (A subsidiary of SABMiller plc).*

"Systems & control provides the ability to understand, model, and often improve, everything. With today's megatrend of software in everything, it is more important, more powerful, and able to make a bigger impact than ever before on the future of humanity." *Jack Little, President and cofounder of MathWorks.*

"In the current age the extraordinary increase in complexity, a major challenge in many fields, calls for detailed system wide analysis, estimation and high performance control design in order to enable optimal system operation at the physical limits without increasing cost." *Silvia Mastellone(*), Principal Scientist, ABB Corporate Research Switzerland Ltd.*

"The transition towards a low carbon and sustainable economic will require a more efficient utilization of existing assets, raw materials and energy: advanced modelling, simulation and controls are definitively key enablers for this transition." *Patrick Panciatici, Scientific Advisor, RTE (French Transmission System Operator).*

"Control is the only discipline that provides a rigorous foundation for optimal decision making under uncertainty for dynamical systems, and it is thus a crucial discipline for managing all such systems—engineered devices of all kinds, complex infrastructures that are the platforms for our civilization, economic and social organizations, and indeed our planetary ecosystem." *Tariq Samad(*), Corporate Fellow, Honeywell (retired).*

(*) Member of the IFAC Pilot Industry Committee

References

A Roadmap for U.S. Robotics: From Internet to Robotics, (2013). Edition, https://robotics-vo.us/sites/default/files/2013%20Robotics%20Roadmap-rs.pdf

Albertos, P., & Mareels, I. (2010). *Feedback and control for everyone*. Springer.

Aström, K. J., & Kumar, P. R. (2014). Control: A perspective, survey paper. *Automatica, 50*, 3–43.

Aström, K. J., & Murray, R. M. (2008). *Feedback systems: An introduction for scientists and engineers*. Princeton University Press Available at *http://www.cds.caltech.edu/~murray/amwiki*.

Autonomy Research for Civil Aviation: Toward a New Era of Flight, By Committee on Autonomy Research for Civil Aviation, Aeronautics and Space Engineering Board, Division on Engineering and Physical Sciences, National Research Council (2017).

Baillieul, J., & Samad, T. (Eds.). (2017). *The encyclopedia of systems & control. Collects a broad range of short expository articles that describe the current state of the art in the central topics of control and systems engineering as well as in many of the related fields in which control is an enabling technology http://www.springer.com/engineering/control/book/978-1-4471-5057-2*.

Convergence, National Academy Press 2014. https://www.nae.edu/113283.aspx.

Convergence: The future of health. Cambridge, MA (2016). http://www.convergencerevolution.net/2016-report/.

Engell, S., Paulen, R., Sonntag, C., Thompson, H., Reniers, M. A., Klessova, S., et al. (2016). *Proposal of a European research and innovation agenda on cyber-physical systems of systems* Dortmund, Germany. Available at: http://www.cpsos.eu/roadmap.

Friedkin, N. E. (2015). The problem of social control and coordination of complex systems in sociology: A look at the community cleavage problem. *IEEE Control Systems Magazine, 35*(3), 40–51.

Hess, R. A. (2002). Pilot control. In P. S. Tang, & M. A. Vidulich (Eds.), *Principles and practice of aviation psychology*. NJ: Erlbaum, Mahwah Chap. 8.

Hodne, T. (Ed.). (2014). *DNV GL: The future of shipping*. Høvik.

HYCON2 Report. (2011). *Systems & control Recommendations for a European Research agenda towards Horizon 2020* Report prepared by the HYCON2 Network of Excellence and endorsed by major industries *http://cordis.europa.eu/fp7/ict/computing/documents/hycon2-report2011.pdf*.

Johansson, K. H., Pappas, G. J., Tabuada, P., & Tomlin, C. J. (Eds.) (2014). *Control of cyber-physical systems* (59). Special Issue IEEE Trans. Automatic Control.

Kim, K-D., & Kumar, P. R. (2012). Cyber-physical systems: A perspective at the centennial. *Proceedings of the IEEE: Centennial Issue, 100*(13), 1285–1286 May.

Korondi, P., Korcsok, B., Szilveszter, K., & Mihoko, N. (2015). Etho-robotics: What kind of behaviour can we learn from the animals? In *11th IFAC Symposium on Robot Control, IFAC-PAPERSONLINE: 48* (pp. 244–255).

Lloyd's Register, QinetiQ and University of Strathclyde. Global Marine Trends 2030, (2013).

Minerva, R., Biru, A., & Rotondi, D. (2015). *Towards a definition of the Internet of Things (IoT)* 27 May [accessed 22 Oct 2016] http://iot.ieee.org/images/files/pdf/IEEE_IoT_Towards_Definition_Internet_of_Things_Revision1_27MAY15.pdf.

Murray, R. M. (Ed.). (2003). *Control in an information rich world: Report of the panel on future directions in control, dynamics and systems*. SIAM Available at *http://www.cds.caltech.edu/~murray/cdspanel*.

Newman, M. E. J. (2003). The structure and function of complex networks. *SIAM Review, 45*(2), 167–256.

"PReVENT Project" official website. Available at http://www.transport-research.info/project/preventive-and-active-safety-application (2017).

Railway Technical Strategy, RSSB Dec (2012)., available from http://www.rssb.co.uk/Library/Future%20Railway/innovation-in-rail-rail-technical-strategy-2012.pdf

Samad, T., & Annaswamy, A (Eds.). (2014). *The impact of control technology: Overview, success stories, and research challenges* 2011. Available at https://s3-us-west-2.amazonaws.com/cc-ieeecss/IoCT-FullReport_v2.pdf, 2003 and an updated edition. available at www.ieeecss.org/general/IoCT2-report.

Samad, T. (2016a). *Closing the Loop with IoT—the Role of Control* IEEE IoT Newsletter, November [Accessed 3 March 2017] http://iot.ieee.org/newsletter/november-2016/closing-the-loop-with-iot-the-role-of-control.html.

Samad, T. (2016b). Control systems and the Internet of Things. *IEEE Control Systems Magazine, 36*(1), 13–16 Feb. doi:10.1109/MCS.2015.2495022.

Sharp, P. A., & Leshner, A. (2014). Meeting global challenges. *Science, 343*(6171), 579. doi:10.1126/science.1250725.

Shift2Rail Multi-Annual Action Plan, Shift2Rail JTI Nov (2015)., available from http://www.shift2rail.org/wp-content/uploads/2013/07/MAAP-final_final.pdf.

The United Nations World Water Development Reports. (2015). *"Water for a sustainable world" and "facing the challenges: Case Studies and indicators"* Available at: *http://unesdoc.unesco.org/images/0023/002318/231823E.pdf*.

Wellstead, P. (2017). *A new look at disease, Parkinson's as seen through the eyes of an engineer* Control Systems Principles http://www.control-systems-principles.co.uk.

Wiener, N. (1954). *The human use of human beings: Cybernetics and society*. Houghton Mifflin.

World Wildlife Fund, Water scarcity, https://www.worldwildlife.org/threats/water-scarcity (2017).