OSTBAYERISCHE
TECHNISCHE HOCHSCHULE
REGENSBURG

FMS-BERICHTE
WINTERSEMESTER 2023

Hans Meier, Michael Niemetz, Thomas Fuhrmann
und Andrea Reindl (Hrsg.)

Seminar zu aktuellen Themen der Elektro- und Informationstechnik

Forschungsmethoden-Seminar Ausgabe 5,
Winter 2022/23

26. April 2023

Inhaltsverzeichnis

Vorwort

Dieser Bericht entstand auf Initiative der Studierenden des Masterstudiengangs „Elektro- und Informationstechnik (MEI)“, die an der Pflichtlehrveranstaltung „Forschungsmethoden und Seminar (FMS)“ im Wintersemester 2022/23 teilnahmen.

Diese Lehrveranstaltung hat das Ziel, systematisch an das wissenschaftliche Arbeiten, speziell die Wissenschaftskommunikation, heranzuführen. Daher war geeignete Literatur zu einem individuellen Thema zu recherchieren, Veröffentlichungen auf ihre Relevanz hin zu beurteilen und letztendlich eine eigene Ausarbeitung basierend auf der Literaturrecherche zu erarbeiten und diese in einem Vortrag zu präsentieren.

Parallel dazu erfolgte im Theorieteil die entsprechende Hinführung zu den verschiedenen Elementen der Wissenschaftskommunikation:

- Bedeutung der Wissenschaftskommunikation für die Arbeit der Ingenieure in Forschung und Entwicklung
- Literaturrecherche, Suchmaschinen, Sichtung und Analyse vorhandener Publikationen, Bewertung der Qualität aufgefundener Fachliteratur, Auswahl geeigneter Materialien für die eigene Arbeit
- Aufbereitung und Darstellung der recherchierten technischer Inhalte in Form einer seitenanzahlbegrenzten wissenschaftlichen Ausarbeitung
- Einhalten formaler Randbedingungen bzgl. Strukturierung, einschl. Bildnachweise und Zitationsstile
- Peer-review-Prozess bei wertschätzender Beurteilung der Leistung anderer
- Publikumsangepasstes Aufbereiten komplexer fachlicher Inhalte mit hochschulöffentlicher Präsentation der Ergebnisse
- Führen mündlicher wissenschaftlicher Diskurse

Nachdem die Masterstudierenden in der Regel über noch keine eigene wissenschaftliche Forschungserfahrung bzw. -inhalte verfügen, lag der wählbare Schwerpunkt der Literatursuche auf der Bearbeitung von vorgegeben aktuellen technischen oder gesellschaftspolitischen Forschungsthemen.

An dieser Stelle noch vielen Dank an Florian Geitner, Jakob Bindl, Christian Syländer und Franz Kerschreiter für die Übernahme der Session Chair Rolle.

Sensory Approaches for Preventive Damage Monitoring in Wind Turbines

Florian Geiss
Faculty of Electrical Engineering and Information Technology
Technical University of Applied Sciences
Regensburg, Germany
florian.geiss@st.oth-regensburg.de

***Abstract*—Maintenance and repair of Wind Turbines (WT) have become more challenging and essential due to larger dimensions or locations with limited access. Repair and downtime costs can potentially be reduced with the continuous monitoring of structure and environment. Especially, maintenance of Wind Turbine Rotor Blades (WTRBs) is important, due to high component costs and failure rates. This paper gives an overview of the different sensory approaches that are suitable for the preventive damage monitoring in WTRBs. While the basic principles and functionality behind the methods will be explained, detailed technical and mathematical background about the sensors will not be given. Focus is instead put on the sensor systems and their methods to detect damages in WTRBs. The overview includes acoustic emission sensors, fiber optic sensors for strain measurement in the WTRBs and blade tip movement sensors for structural monitoring. Each sensor can detect specific kinds of damages, and the more approaches are used in a wind turbine, the more effective preventive damage monitoring becomes. Commercially available sensor systems, based on the mentioned sensor principles, are presented to show the practical use of Structural Health Monitoring (SHM) in WTRBs. With the growing need of renewable energies like wind energy, the presented SHM systems can potentially reduce the maintenance costs and therefore also the cost of environmental friendly energy.**

***Index Terms*—Wind turbine, sensors, preventive maintenance, condition monitoring, structural health monitoring, acoustic emission, fiber-optic sensor**

I. Introduction

In 2022, the German government released a Renewable Energy Sources Act, in which they set the goal to gain at least 80 percent of gross electricity from renewable energy sources by 2030 [1]. Like Germany, most other countries try to reduce their dependence on fossil fuels, so the demand for renewable energy plants is rapidly growing. Beneath solar and hydraulic energy production, wind energy plays an important role in the renewable energy sector, as Wind Turbines (WT) have produced 15.8% of the gross electricity in Germany in the year 2021 [2]. With a proportion of 38.7% that is the highest contribution from a renewable energy technology [3]. While the number of WTs has shown a rising trend during the last decades worldwide, those systems still suffer from many reliability issues, especially when they are placed in a harsh offshore environment [4]. Due to site-accessing difficulties and the challenges of conducting blade repair or replacement over sea, unforeseen malfunctions of Wind Turbine Rotor Blades (WTRBs) are an even bigger problem for offshore than for land-based installations. In order to keep the number of physical inspections low without increasing the risk of structural failure, a reliable remote monitoring system for damage identification is necessary [5]. Condition-based maintenance along with structural health monitoring SHM gives an opportunity for that and therefore is used in various fields [6]. According to a survey of 1500 onshore WTs over 15 years, 7% of WT failures have been produced by WTRBs, ahead of gearboxes (4%) and generators (4%) [7]. Figure 1 shows a diagram with the different WT components sorted by their annual failure rate. With the average downtime per failure of 4 days and an annual failure rate of 0.2, WTRBs are among the most vulnerable structural components in a WT. Additionally, blade failures can be associated with significant financial loss, as they make up 20% of the WTs total cost [5]. Because of that, the focus of this paper will lie on the sensory approaches for SHM of WTRBs. SHM is defined in our case as the condition monitoring of the WTRB structure. In chapter II different technologies are introduced, that are suitable for usage in real WTs. The systems based on those technologies are presented in chapter III, followed by a conclusion in chapter IV.

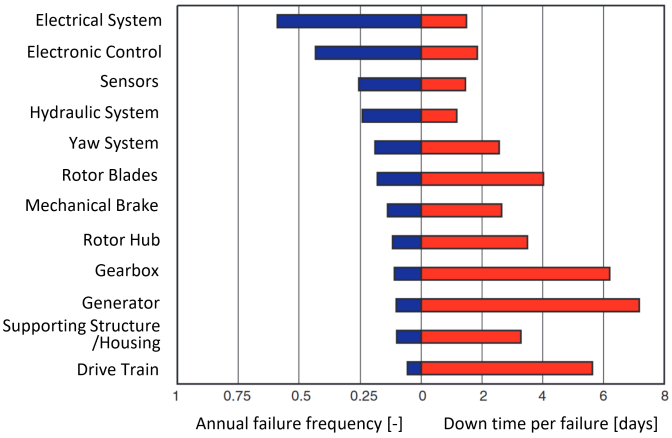

Fig. 1. Annual failure rates of WT components and corresponding down time in days [7]

II. Technologies

Numerous wind turbine damage detection and health monitoring techniques exist. Those can be divided in active and passive methods [8]. Passive sensors do not interact with the structure they are monitoring. Modal approaches, Acoustic Emission (AE), strain measurement methods and impedance tomography are some condition monitoring methods in this category. Active methods, like ultrasonic wave propagation, interact with the structure. [9]

The technologies used for SHM can be grouped according to the methods they are based on, each with their individual benefits and downsides. Methods like ultrasonic wave propagation [10], but also X-radioscopy method [11] and thermal imaging method [12] show precise damage detection, but are difficult to implement in operating WT with continuous SHM and are therefore not covered in this paper. An effective SHM approach actually used in the field is AE.

A. Acoustic Emission (AE) Sensors

AE is based on the detection of propagating elastic waves released by sudden stress-strain change in materials. Processes such as cracking, deformation, debonding, delamination, rubbing, impacts, crushing, turbulence and others, all produce localized transient changes in stored elastic energy with a broad spectral content. Acoustic emission sensors detect the high frequency component of the elastic waves (or stress waves) naturally generated by these energy loss processes within materials and structures [13]. The energy released by an AE source propagates through the medium and reaches the AE sensor, which converts the mechanical signal into an electrical signal. AE sensing is an effective technique used to immediately indicate the location and onset of damage [13]. This method is able to detect developing damages even at microscopic level and spreading cracks before they are visible. References [14], [15] reported that the characteristics of an AE event and the increase of intensity as damage becomes more critical, can be utilized by pattern recognition software to evaluate the damage. This pattern recognition software has the potential of being applied to various similar WT blade designs [16]. When high accuracy of damage evaluation is needed, the number of sensors must be increased. To reduce the number of data output to the signal processing system, a Structural Neural System (SNS) for SHM can be used [17].

The SNS contains a highly distributed continuous sensor concept, that mimics the signal processing in the biological neural system. Figure 2 illustrates a basic SNS structure. The row and column continuous sensors are individual Piezoelectric Sensors (PZT), while two adjacent PZTs (one row and one column) are combined in one node. Rows and columns are connected to separate analog processors. The processors will simplify the data and send only two outputs each to the SHM computer, which greatly reduces both the number of data to be processed and the computer power [16]. It was reported that this method can detect damage early and track the AE event during the damage growth in a WT blade [18]. While PZT sensors are used successfully in many applications, there

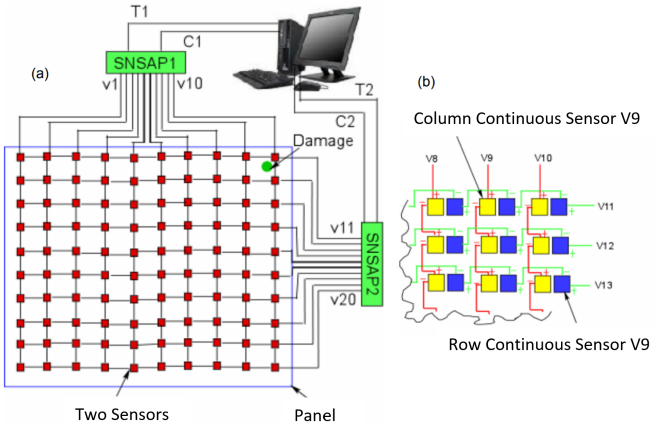

Fig. 2. Architecture of the Structural Neural System (SNS): (a) each small square indicates two adjacent sensor nodes; (b) the magnified view shows the detail of sensors arrangement to form row continuous sensors and column continuous sensors [18]

is a higher risk of damage from lightning strikes caused by the electrical conductive wires. Wind turbine blades manufactured of non-conducting fiber composite materials without any conducting components are frequently struck by lightning, particularly at the outermost part of the blade [19]. In contrast to other AE approaches, the system proposed in [20] detects damages using airborne sound in the lower frequencies from about 200 Hz to 20000 Hz. The benefit of using airborne sound in these frequencies is, that a much smaller amount of sensors is needed. As sensors, fiber optic microphones are used. Their cords are optical fibers which are non-metal, so they do not increase the risk of damage from lightning strikes. It is shown, that airborne sounds are suitable for detection of WTRB damages [20].

B. Fiber Optic (FO) Sensors

FO strain sensors are among the most widely used in WT structural monitoring [21]. This is due to the fact that strain is among the most important measurement variables for SHM in WTRBs [5]. One type of FO sensor are in core optical fiber gratings, which offer great possibilities to create devices for both sensing and telecommunication applications [22]. For optical fiber gratings, the short period of light (Bragg) is used to measure both the strain and temperature of a structure. The principle behind the most commonly used FO strain sensor, Fiber Bragg Grating (FBG), is an optical fiber with a periodic refractive index perturbation pattern inscribed in the core [23]. When an incident spectrum of light propagates through the grating, a specific wavelength, named the Bragg wavelength, is reflected, while the rest of the spectrum can pass unaffected. If an external axial strain is induced, the FBG reacts accordingly because of its length getting stretched or compressed. This causes a proportional shift in the reflected Bragg wavelength. With a one-time instrument calibration, strain and other derived parameters can be measured dynamically. The power in this mode is determined by integrating the scattered radiation

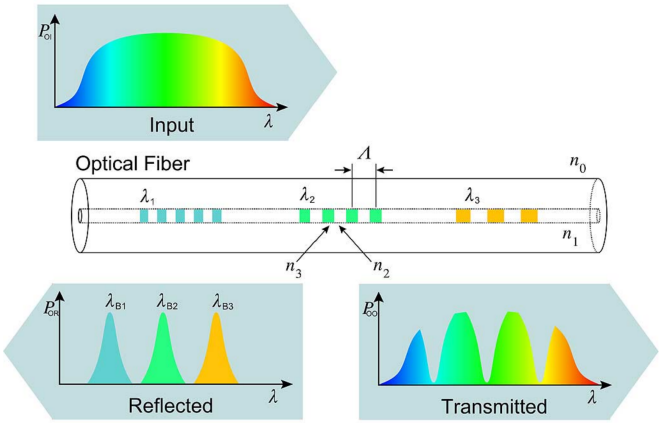

Fig. 3. Illustration of the fiber Bragg grating concept and its optical function [24]

at λ_B shown as P_{OR} (reflected optical power). The rest of the optical power P_{OO} is transmitted as a forward propagating mode. In this case, the peak of mode coupling in the reflected spectrum occurs at the resonant or Bragg wavelength λ_B given by (1).

$$\lambda_B = 2 \cdot n_{\text{eff}} \cdot \Lambda \tag{1}$$

In this expression, n_{eff} is the effective index of the mode in the grating. Measurement of the peak reflected wavelength results in the direct measurement of the optical product $n_{\text{eff}} \cdot \Lambda$ of the grating. Any perturbation that modifies n_{eff} or the grating period Λ will alter the measured Bragg wavelength. Thus, the fiber grating can be used as an intrinsic optical transducer which changes the spectrum of the reflected light. By means of an FBG, both the temperature and the strain can be measured [25]. This means, a specially intended temperature sensor has to be added to calculate effects on the Bragg wavelength caused by temperature changes [26]. A key advantage of FBG technology is its capability to multiplex many FBGs in the wavelength domain [25], enabling the very effective development of quasi-distributed sensor systems. Several FBGs can be written in the same fiber as an array of independent sensors with different period Λ_x or effective index $n_{\text{eff},x}$. This means, several different Bragg wavelengths can be clearly resolved using the same fiber line, as shown in Figure 3.
Their in-line optical connection property makes it feasible for FBGs to build up FO sensor networks using wavelength, time and both active and passive hybrid multiplexing schemes [27]. In addition to its wavelength encoded responses, FBG technology offers the possibility to build up sensors with linear output, high sensitivities, high dynamic ranges and resolutions that are insensitive to electromagnetic interference. The size of individual FBGs can vary in length from 0.1 mm to several centimeters [28]. A distribution of FBGs over the WTRB structure can be used to detect traverse crack evolution and also impact damage. The impact event can be detected in real time by measuring both the abrupt change of strain and the time delay in the strain changes between the sensors. By using the difference in arrival time of the strain response, the position of the impact event can be located [16]. Strain measurement is also used in the approach introduced in the next chapter.

C. Modal Approach

WTRBs can be approximated as slender beams in a cantilever configuration, accounting for the effect of rotational velocity and aerodynamic forces. Describing such a structure with the exact mathematical equations is in fact difficult, because the geometrical and material properties can not be factored out. This is because of inhomogeneities in the structure and its complex form [29]. Instead, the Institute for Statics and Dynamics (ISD) from the University of Hanover created a system that is simpler and easier to implement in the field. The approach is based on the extension of Gasch's proportionality method, according to which maximum oscillation velocity $\nu_{\max}$ and maximum dynamic stress $\sigma_{\text{dyn,max}}$ are proportional by a factor p_{System} under harmonic excitation, as (2) points out [30].

$$\sigma_{\text{dyn,max}} = p_{\text{System}} \cdot |\nu_{\max}| \tag{2}$$

This factor p_{System} is referred to as the proportionality factor. It describes the dynamic behavior of the structure. A change in the proportionality factor can be interpreted as a change in the static structure and so be used as indicator for damage [9]. Values for maximum stress occur at the root of a WTRB, while the maximum oscillation velocity is found at the tip. Nevertheless, it has been proven, that the proportionality is also valid for velocities measured in the middle of the WTRB [9]. For the application of the proportionality method to WTRBs, a sensing concept was developed by ISD. The sensors used for stress monitoring are a four-sensor chain of FOS gauges. The FBG sensors explained in the previous chapter are a possible fit here. These four sensors are placed at the root area of the blade and are oriented in the longitudinal direction in combination with two novel sensor prototypes measuring deflections at parallel positions on the webs of the blade [30]. The latter sensors are specifically designed for use in WTRBs. The strain measurements at the root occurring from the FO deflection sensors are used for determining the maximum oscillation velocity. A glass fiber string is connected to the web attached to a single fixation hook, far inside the blade. This string is guided along the web to the root section, where the active sensor unit is also fixed to the web. The sensor positions mounted on the blade structure are shown in Figure 4. [9]

Deflection of the blade in the sensing location is possible in both the edgewise and flapwise directions, which results in a change of the string angle at the sensor, that is located near the root section. This angle can be related to the deflection amplitude, since the string tension is constant after the initial calibration process. The basic sensing system is fixed to the active unit. Directly behind the location of string fixation to the web, two orthogonally oriented load cells are connected to

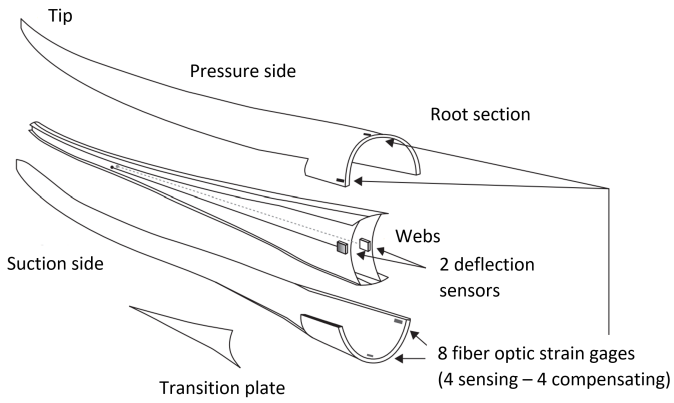

Fig. 4. Institute for Statics and Dynamics (ISDs) proportionality method for Structural Health Monitoring (SHM) of Wind Turbine Rotor Blades (WTRBs) [9]

the string. It appears that an extension of the proportionality method for the quantity of acceleration will also allow its use for the application of FO accelerometers, instead of the deflection sensor. [9]

III. Systems for Structural Health Monitoring (SHM) of Wind Turbine Rotor Blades (WTRBs)

The technologies presented in chapter II are used in systems, which are developed for SHM of WTRBs. These systems are commercially available and partially used in existing WTs. Table I shows a selection of systems based on the introduced technologies combined with further approaches. It is shown, that in some SHM systems, several sensor technologies are combined to achieve a higher accuracy in monitoring the condition of a WTRB. While each system has its own strengths and use cases, not all are suitable for continuous SHM. Thus, the inspection robot can achieve a highly detailed overview of the blade's condition, but can mainly be used purposefully for non-destructive testing. Non-destructive testing methods are more appropriate in the event that an SHM system has sent an alarm and a more detailed damage analysis has to be executed. For collecting and processing the SHM data, often a system called Supervisory Control and Data Acquisition (SCADA) is used. This combines not only the data from different systems that are intended for condition monitoring of WTRBs, bearings, gearbox etc. but also environmental data like wind speed or power production. The collected data can be processed and then accessed remotely, along with warnings and alarms triggered by SHM systems. [31]

IV. Conclusion

In this paper, sensory approaches for SHM systems are described. The implementation of SHM in WTRB are able to reduce down-times and operational expenses. Due to the fact that WTRBs make up a significant proportion of a WTs total cost and the remote locations of offshore wind parks, unforeseen failures of WTRBs are associated with high costs. Acoustic Emission (AE) sensors show a high detection rate of emerging damages, especially when used in a Structural Neural System (SNS). While AE sensing has some disadvantages, the approach is still suitable for smaller WTs and is sensitive to all kinds of damages.
Fiber optic (FO) strain sensors, particularly Fiber Bragg Gratings (FBGs) are among the most widely used sensors in WT structural monitoring. Even though, there are certain limitations to FO based sensors too, there are well-developed solutions for that approach. One described solution is the system made by the institute for statics and dynamics in Hanover, that monitors the maximum dynamic strain in the root of a WTRB and the maximum deflection or acceleration in the blade tip. These are proportional by a factor and a change in that factor can point out a change in the structural constitution of the WTRB. With this approach, fewer sensors are needed, while the detection of structural damages is still accurate. Better monitoring of the structural health can be achieved, when several approaches are combined. This is possible with systems like SCADA, which takes several data sources into account. Although SHM systems have some difficulties, their ability to lower the operating expenses of wind energy makes the usage of WTs more affordable and easier to use in high number.

TABLE I
Systems for Structural Health Monitoring (SHM) of Wind Turbine Rotor Blades (WTRBs) [9]

System	Method	Sensor
BLADEcontrol	Frequency-based method	Two accelerometers
MOOG RMS	Load control	Four FOS gauges per blade
SKF Blade Monitoring System	Blade deflection	Laser
LM Blade Monitoring	Breakage of embedded fiber	Embedded fiber
ISD SHM Proportionality Method	Proportionality method	Four strain gauges, One deflection sensor or FO accelerometers
Woelfel SHM.Blade	AE, Acousto-ultrasonics, Operational modal analysis	Piezoelectric sensors, Compatible strain gauges
SNS	AE	Piezoelectric sensors
Inspection robot	Thermography, Ultrasonic wave propagatoin	Infrared radiator, High resolution thermal camera, Ultrasonic system, High resolution camera

References

[1] "EEG amendment: More renewable energy for increased climate protection." (2022), [Online]. Available: https://www.bundesregierung.de/breg-de/themen/klimaschutz/amendment-of-the-renewables-act-2060448 (visited on 11/19/2022).

[2] BDEW. "Anteil Erneuerbarer Energieträger an der Bruttostromerzeugung in Deutschland in den Jahren 2020 und 2021 Statista." (2021), [Online]. Available: https://de.statista.com/statistik/daten/studie/171368/umfrage/struktur-der-bruttostromerzeugung-durch-erneuerbare-energien-in-deutschland/ (visited on 11/24/2022).

[3] BDEW. "Verteilung der Stromerzeugung aus Erneuerbaren Energien in Deutschland nach Ernergieträger im Jahr 2021 Statista." (2021), [Online]. Available: https://de.statista.com/statistik/daten/studie/173871/umfrage/stromerzeugung-aus-erneuerbaren-energien-in-deutschland/ (visited on 11/24/2022).

[4] W. Yang, Z. Peng, K. Wei, and W. Tian, "Structural health monitoring of composite wind turbine blades: Challenges, issues and potential solutions," *IET Renewable Power Generation*, vol. 11, no. 4, pp. 411–416, 2017.

[5] M. Moradi and S. Sivoththaman, "Mems multisensor intelligent damage detection for wind turbines," *IEEE Sensors Journal*, vol. 15, no. 3, pp. 1437–1444, 2015. DOI: 10.1109/JSEN.2014.2362411.

[6] I. Alsyouf, "The role of maintenance in improving companies' productivity and profitability," *International Journal of production economics*, vol. 105, no. 1, pp. 70–78, 2007.

[7] B. Hahn, M. Durstewitz, and K. Rohrig, "Reliability of wind turbines," in *Wind energy*, Springer, 2007, pp. 329–332.

[8] V. Giurgiutiu and A. Cuc, "Concepts for using piezoelectric wafer active sensors in rotorcraft health monitoring," *Structural Health Monitoring*, pp. 1173–1182, 2005.

[9] R. Rolfes, S. Tsiapoki, and M. Häckell, "Sensing solutions for assessing and monitoring wind turbines," in *Sensor technologies for civil infrastructures*, Elsevier, 2022, pp. 391–426.

[10] S. K. Chakrapani, D. J. Barnard, and V. Dayal, "Review of ultrasonic testing for nde of composite wind turbine blades," in *AIP Conference Proceedings*, AIP Publishing LLC, vol. 2102, 2019, p. 100 003.

[11] E. Jasinien, R. Raiutis, A. Voleiis, A. Vladiauskas, D. Mitchard, M. Amos, *et al.*, "Ndt of wind turbine blades using adapted ultrasonic and radiographic techniques," *Insight-Non-Destructive Testing and Condition Monitoring*, vol. 51, no. 9, pp. 477–483, 2009.

[12] R. Haj-Ali, B.-S. Wei, S. Johnson, and R. El-Hajjar, "Thermoelastic and infrared-thermography methods for surface strains in cracked orthotropic composite materials," *Engineering Fracture Mechanics*, vol. 75, no. 1, pp. 58–75, 2008.

[13] E. R. Jørgensen, K. K. Borum, M. McGugan, *et al.*, "Full scale testing of wind turbine blade to failure-flapwise loading," 2004.

[14] A. Anastassopoulos, D. Kouroussis, V. Nikolaidis, *et al.*, "Structural integrity evaluation of wind turbine blades using pattern recognition analysis on acoustic emission data.," *Journal of Acoustic Emission*, vol. 20, no. 1, 2003.

[15] P. Joosse, M. Blanch, A. Dutton, D. Kouroussis, T. Philippidis, and P. Vionis, "Acoustic emission monitoring of small wind turbine blades," *J. Sol. Energy Eng.*, vol. 124, no. 4, pp. 446–454, 2002.

[16] C. C. Ciang, J.-R. Lee, and H.-J. Bang, "Structural health monitoring for a wind turbine system: A review of damage detection methods," *Measurement science and technology*, vol. 19, no. 12, p. 122 001, 2008.

[17] G. R. Kirikera, V. Shinde, M. J. Schulz, A. Ghoshal, M. Sundaresan, and R. Allemang, "Damage localisation in composite and metallic structures using a structural neural system and simulated acoustic emissions," *Mechanical Systems and Signal Processing*, vol. 21, no. 1, pp. 280–297, 2007.

[18] M. Schulz and M. Sundaresan, "Smart sensor system for structural condition monitoring of wind turbines: 30 may 2002–30 april 2006," National Renewable Energy Lab.(NREL), Golden, CO (United States), Tech. Rep., 2006.

[19] F. M. Larsen, "New lightning qualification test procedure for large wind turbine blades," in *International Conference on Lightning and Static Electricity (ICLOSE), Blackpool, 2003*, 2003.

[20] T. Krause, S. Preihs, and J. Ostermann, "Acoustic emission damage detection for wind turbine rotor blades using airborne sound," in *Proceedings of the 10th International Workshop on Structural Health Monitoring (IWSHM2015)*, 2015.

[21] D. Ozevin, "Mems acoustic emission sensors," *Applied Sciences*, vol. 10, no. 24, p. 8966, 2020.

[22] R. Kashyap, *Fiber bragg gratings (optics and photonics series)*, 1999.

[23] R. Kashyap and J. Lopez-Higuera, "Fiber grating technology: Theory, photosensitivity, fabrication and characterization," *Handbook of Optical Fibre Sensing Technology*, pp. 349–374, 2002.

[24] J. M. López-Higuera, L. R. Cobo, A. Q. Incera, and A. Cobo, "Fiber optic sensors in structural health monitoring," *Journal of lightwave technology*, vol. 29, no. 4, pp. 587–608, 2011.

[25] J. M. López-Higuera, *Optical sensors*. Ed. Universidad de Cantabria, 1998.

[26] K. Schroeder, W. Ecke, J. Apitz, E. Lembke, and G. Lenschow, "A fibre bragg grating sensor system monitors operational load in a wind turbine rotor blade,"

Measurement science and technology, vol. 17, no. 5, p. 1167, 2006.

[27] M. López-Amo and J. M. López-Higuera, "Multiplexing techniques for fbg sensors," *Fiber Bragg Grating Sensors: Recent Advancements, Industrial Applications and Market Exploitation*, pp. 99–115, 2011.

[28] J. M. López-Higuera, L. R. Cobo, A. Q. Incera, and A. Cobo, "Fiber optic sensors in structural health monitoring," *Journal of lightwave technology*, vol. 29, no. 4, pp. 587–608, 2011.

[29] R. Ernst, P. Lauridsen, C. Klein, and B. Buffetti, "The effect of damage position on operational modal analysis of wind turbine blades for shm," in *Journal of Physics: Conference Series*, IOP Publishing, vol. 2265, 2022, p. 032 099.

[30] S. Zerbst, "Global approach for early damage detection on rotor blades of wind energy converters," ISD, 2011.

[31] W. Yang, R. Court, and J. Jiang, "Wind turbine condition monitoring by the approach of scada data analysis," *Renewable energy*, vol. 53, pp. 365–376, 2013.

Energy Harvesting: Grätzel Cell (Dye Sensitized Solar Cell DSSC)

Lucia Liedl
Faculty of Electrical Engineering and Information Technology
Ostbayerische Technische Hochschule Regensburg
Regensburg, Germany
lucia.liedl@st.oth-regensburg.de

***Abstract*—In order to favor carbon neutral processes for energy generation, photovoltaic cells became inevitable both in the industrial as in the domestic sector. Since the beginning of photoelectric experiments by Becquerel photovoltaic cells have been developed, which convert sunlight into electrical power and are usually made of silicon (Si). A special form of the photovoltaic cell is the Grätzel cell, which is named after its inventor Prof. Michael Grätzel. The Grätzel cell is a Dye-Sensitized Solar Cell (DSSC) and represents an alternative to the Si-based p-n junction photovoltaic devices by using nanocrystalline materials and conducting polymer films. The fabrication of these materials is affordable and easy and the conversion efficiencies are relatively high. This increases the ability of competing with conventional market-leading Si cells. Photons enter the cell and excite electrons of a dye, which diffuse and enter through an electrode an external circuit. The remaining holes move to the electrolyte and recombine in a redox reaction with the electrons re-entering the cell. Resulting in a smaller impact on the pollution of the environment, natural and organic dyes made of chlorophyll, flowers, fruits, or melanin were tested as dyes in the DSSCs. With a compound of different organic materials, the whole sun spectrum can be absorbed. In this review paper the structure, the operating principle and recent enhancements of the DSSCs are summarized. It concentrates on the improvements on the dye and the electrolyte by using organic and renewable materials.**

***Index Terms*—Organic Solar Cells, Solar Energy, Solar Photovoltaic, Photovoltaic**

I. Introduction

The topic of utmost importance for this generation is the climatic change. Many improvements have been made to save the planet. In 2019, the *European Union* decided the *European Green Deal* in order to achieve climate neutrality by 2050. The Greenhouse Gases should be reduced by 55% by 2030, which implicates a share of renewable energy of at least 38.5%. [1] Photovoltaic energy has been used a lot in the past to generate electrical energy. With $3 \cdot 10^{24}\,\mathrm{J}$ a year the sun provides about $10,000$ times as much energy as needed to cover the global population's consumes. Even a rapidly increasing need of energy could be covered by sun energy. In order to compete with fossil fuels, renewable energy has to reach higher efficiency rates and must be more profitable. [2] Therefore, photovoltaics has been improved immensely in the last decades. The photoelectric effect takes place when a photon interacts with a semiconductor and therefore generates an electron-hole pair. With the following potential difference between the two different materials in the cell a current is flowing. [2] Most cells use inorganic solid-state materials such as crystalline or amorphous silicon (Si) due to high conversion efficiencies of $15 - 20\%$. These cells need pure Si, their conversion efficiency decreases at high temperatures, and the manufacture is energy-intensive. [3] These problems were driving forces for scientists to avoid inorganic materials and investigate organic solar cells. This step was taken by Prof. *Michael Grätzel* in 1991 by inventing the Dye-Sensitized Solar Cell (DSSC), the so-called "Grätzel cell", which is a next generation photovoltaics able to compete with Si-based p-n junction cells. With these photovoltaics an even higher share of renewable energy can be harvested by using low-cost materials. [4]

II. Operating principle of the DSSC

A typical DSSC consists of at least three components: the photoanode, the counter electrode and the liquid electrolyte (see Fig. 1). The photoanode consists of a conducting glass substrate, which is glass with a thin layer of Transparent Conducting Oxide (TCO) plated on it, e. g. Tin Oxide (SnO_2). On this layer a mesoporous semiconductor, e. g. Titania (TiO_2), of about $3 - 12\,\mu$m is plated. The particles have a size of about $20 - 50\,$nm. The mesoporous semiconductor is sensitized with a monolayer of a dye. The dye is an electron transfer sensitizer and absorbs the light. The counter electrode also begins with a conductive glass substrate with a thin layer of an electrocatalytic material, e. g. platinum. The photoanode and counter electrode are put together with a liquid electrolyte in-between, which contains a redox mediator or a solid Hole Transporting Material (HTM). The electrolyte has a wide band gap to guarantee stability. Therefore, different materials have been tested in the DSSC. [4], [5]

The converting of solar in electrical energy in a DSSC is shown in Fig. 2. A photon (hv) interacts with the dye and excites an electron located on the valence band (VB). If the energy of the photon is higher than the band gap, the electron joins the conduction band (CB) (step 1). TiO_2 has a lower energy level of the CB than the dye. Because of an energy alignment the electron receives enough driving force to join the CB of TiO_2 (step 2). The electron diffuses through the mesoporous layer and endures many surface states (step 3) with the purpose of reaching the glass of the photoanode and joining an external circuit (step 4). Still in its oxidized state

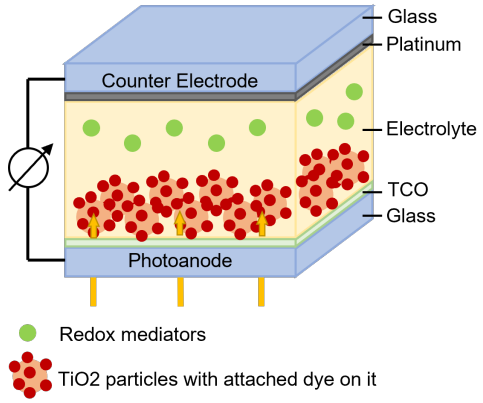

Fig. 1. The schematic structure of the Grätzel cell. [5], [9]

the dye waits to be regenerated by the redox mediator or the HTM of the electrolyte (step 5), which conversely will be regenerated at the counter electrode by the re-entering electrons (step 6). [5], [6], [7] Because the oxidized dye also oxidizes the electrolyte and reduction takes place as well, the occurring oxidation and reduction in the electrolyte are described as redox reaction [8].

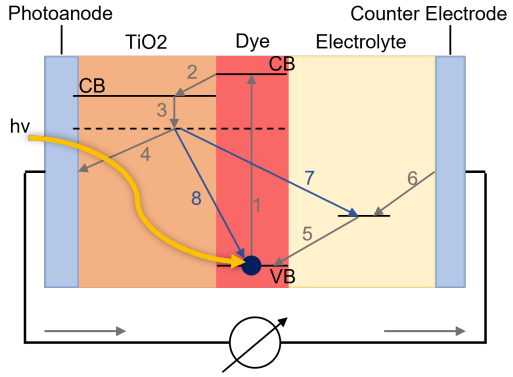

Fig. 2. The converting of solar in electric energy of the DSSC in detailed steps. [2], [3], [9]

There are also reactions working against the energy harvesting in the DSSC diminishing the generated photocurrent. Electrons from the CB of TiO_2 or the oxide substrate may combine with the redox mediator or the HTM (step 7) or they could recombine with the oxidized dye (step 8). By using an additional compact blocking TiO_2 layer the side effects can be prevented partly. [5], [7]
The operating principle can also be described with following equations. An absorbed photon (hv) excites the dye sensitizer (S) (eqn (1)). The excited sensitizer (S^*) injects an electron (e^-) into the CB of the semiconductor (eqn (2)), which moves via the semiconductor to the photoanode entering an external circuit. The electron reaches the counter electrode and reduces the redox mediator (I_3^-) (eqn (3)). The reduced redox mediator (I^-) regenerates the oxidized sensitizer (S^+) (eqn (4)). Counter reactions are the recombination of an injected electron with the oxidized redox mediator (I_3^-) (eqn (5)) or with the oxidized dye sensitizer (S^+) (eqn (6)). [9]

$$S + hv \rightarrow S^* \tag{1}$$

$$S^* \rightarrow S^+ + e^- \tag{2}$$

$$I_3^- + 2 \cdot e^-_{cathode} \rightarrow 3 \cdot I^- \tag{3}$$

$$S^+ + \frac{3}{2} \cdot I^- \rightarrow S + \frac{1}{2} \cdot I_3^- \tag{4}$$

$$I_3^- + 2 \cdot e^-_{TiO_2} \rightarrow 3 \cdot I^- \tag{5}$$

$$S^+ + e^- \rightarrow S \tag{6}$$

The Incident monochromatic Photon-to-current Conversion Efficiency (IPCE) is an important value to compare solar cells and is defined by the formula

$$\mathrm{IPCE}(\lambda) = \frac{1240}{\lambda} \cdot \frac{J_{SC}}{I_{in}(\lambda)} \tag{7}$$

with the incident monochromatic (consisting of only one wavelength) light wavelength λ in nm, the short-circuit photocurrent density J_{SC} in mAcm^{-2} and the incident monochromatic light intensity I_{in} in mWcm^{-2}. The constant 1240 is measured with the speed of light, Planck's quantum of action and the charge of a single electron. [10] The overall efficiency η_{global} is dependent on J_{SC}, the open-circuit photo-voltage V_{OC}, the fill factor of the cell ff and I_{in}. It is defined by the following formula: [2], [9], [11], [12]

$$\eta_{global} = \frac{J_{SC} \cdot V_{OC} \cdot ff}{I_{in}} \tag{8}$$

III. Historical Background

In 1839 *Alexandre Edmond Becquerel* illuminated metal halide salt solutions between two noble platinum electrodes. This photoelectric experiment showed a current flow in the aqueous solution between the electrodes. [5] First used to develop photography [2], photoelectrochemical conversion of solar to electrical or chemical energy became a scientific topic of paramount interest during the international energy crises in the 1970s. As it turned out, semiconductors are often effective in the absorption of the ultraviolet (UV) and near-ultraviolet share of the solar spectrum. [5] Due to similarities between photoelectrochemical energy harvesting and photography, photographic films were insensitive to mid-spectrum and red light. Due to the band gap of 2.7 to 3.2 electron volt (eV) of the silver halide grains, it can not absorb wavelengths longer than 460 nm. [2] By adding a dye to the silver halide the sensitized solution was able to absorb longer wavelengths in photography. Transferred to photoelectrochemistry with persistent dye sensitization and regeneration, the dye had to be able to absorb photons with wavelengths in the visible part of the solar spectrum. [2] Most used semiconductors have been TiO_2, SnO_2, Zinc Oxide (ZnO) and Silicon Carbide (SiC). The electrodes consisted of single crystals or polycrystalline materials resulting in a low surface area with a low integration capacity of the dye. Low amounts of dye implicated a low

absorption rate of photons and only a small photocurrent. As a result, most scientists focused on monocrystalline semiconductors resulting in the market-leading position of Si cells. [5] As an exception, *Michael Grätzel* and his group at the *Swiss Federal Institute of Technology* in Lausanne have been working with dye-sensitized photoelectrochemical cells and replaced the single crystal electrodes with porous ones. The big surface area of porous electrodes implicated a higher dye volume. [5] The manufacture of mesoporous TiO_2 included heat treatments to improve inter-particle contacts, which is important for a better transport of photogenerated electrons [12]. Another achievement was the faster separation of excited dye electrons in order to diminish counter reactions. The excited dye electrons were injected into the semiconductor with the holes remaining on the dye. [5] *Grätzel* then concentrated on porous TiO_2 electrodes and achieved with a carboxylate group as dye an IPCE of 44% at a wavelength of $460\,\mathrm{nm}$ [13]. According to their outstanding photochemical stability and redox characteristics relevant for photocatalysis, carboxylate groups have been appreciated as catalysts and photosensitizers since the 1960s. In DSSCs a redox mediator or shuttle was added to guarantee an effective regeneration of the oxidized dye. The mediator transferred the charge between the electrodes and consisted of redox mediator couples such as iodine/iodide or bromine/bromide. [5]

Many adaptions followed, like the replacement of fractal by colloidal electrodes with a higher replicability. Li^+-ions were added to the electrolyte, which changed the CB edge resulting in an increased rate of electron injections. The Li^+-ions were replaced by Bu_4N^+-ions to achieve a higher V_{OC}. [5], [6] The most popular dye is inorganic ruthenium (Ru), which is a black dye with an outstanding $IPCE^{max}$ of 80% at $480\,\mathrm{nm}$ in the near-infrared (NIR) part of the spectrum. Most organic and metal-free dyes were instable, so the most DSSCs contained Ru dyes despite the high costs. [5]

IV. Advancements in DSSCs

A. The Dye

The dye of DSSCs should possess well-aligned energy levels regarding the CB edge of TiO_2 and the redox mediator to achieve an efficient electron transfer at the surface. The light absorption should correspond to the incident light's spectrum and the system must be stable. Besides the Ru-based sensitizers, metal-free and organic options were considered with a donor-pi-bridge-acceptor (D-π-A) motive. [4], [5], [14] The electron-donating part (D) is near the electrolyte and the electron-accepting part (A) is near TiO_2. The electronically conjugated building block, called the pi-linker, connects both parts and assures optimal electronic coupling. The D-π-A motive directs a push-pull effect to the semiconductor and simplifies the photoinduced charge separation. The excited electrons are located on the acceptor and the holes on the donor. With the further improved D-A-π-A motive the electron transfer and stability were upgraded. [5], [9], [11], [14]

The most common dyes are Ru-polypyridyl complexes. The transition metal coordination compounds absorb light of the panchromatic visible spectrum by implicating a complicated work process, pollution of the environment, expensive processes, and a limited amount. Avoiding these disadvantages natural dyes seemed to be suitable substitutes. Natural floral dyes are flowers, vegetables, leaves, or fruits. [15]

While Ru dyes absorb the whole spectrum, natural dyes often fail in the red and NIR parts. The key to achieve panchromatic sensitization is the combination of different natural dyes cosensitizing, the so-called *dye cocktails*. [9] Natural dyes are more difficult to conserve, their life span is short, they are less stable, and their availability depends on the seasons. [15] Natural dyes can be picked from fauna as well, e. g. melanins, which absorb light in the UV and visible spectrum. Melanins are pigments found in animals, fungi, plants, and bacteria with changing colors. They are electronically conductive, have redox reversibility qualities, are biocompatible, biostable and are used in bioelectronics and medical devices. The most popular melanin is sepia melanin, which is derived from the ink sack of a cuttlefish in an expensive and non-sustainable procedure. [16] The black powder with hydroxyl and carboxylic functional groups binds with the hydroxyl group of TiO_2 strongly. [15]

Melanins can also be gained from the cell wall of microorganisms, e. g. fungi or bacteria, which protect from chemical and environmental stresses, such as high temperature and UV light. Another idea was the use of marine waste, the *Posidonia oceanica* balls, known as *egagropili*. Found at the coasts of the Mediterranean Sea, *egagropili* powder can be used after some chemical processes as dye in DSSCs. [16]

B. The Redox Mediator

The redox mediator is part of the electrolyte and supports the regeneration of the oxidized dye after electron injection. In order to yield a high IPCE the dye must regenerate fast, the charge transfer resistance must be low, the system must be stable, and the charge carriers must diffuse fast. To guarantee these properties the electrolyte has to be chosen appropriately. [4], [5], [17] The mostly used I_3^-/I^- redox couple with a redox potential of $0.35 - 0.4\,\mathrm{V}$ fulfills most of the required properties besides the forming of some unstable radicals. The I_3^- transports holes to the counter electrode and is reduced back to I^- leading to a $0.3\,\mathrm{eV}$ potential drop. [5]

An important factor to compare different redox mediators is the regeneration half-time. It measures $100\,\mathrm{ns}$ to $10\,\mu\mathrm{s}$ in a typical Ru dye with I_3^-/I^- [7]. The recombination half-time of electrons from the electrode of TiO_2 with I_3^- measures $1\,\mu\mathrm{s}$ and the recombination half-time between the electrons from the electrode of TiO_2 and the oxidized dye counts $100\,\mathrm{ns}$ to $1\,\mathrm{ms}$. The recombination half-time should be relatively high so that the electrons and holes do not recombine too early. Therefore, Li^+-cations were added to the TiO_2 surface, leading to an increased driving force for electron injection. [4], [5]

With an I_3^-/I^- redox couple a very impressive IPCE of 12.4% was reached, but disadvantages prevented the break-through in solar modules, e. g. leakage of the electrolyte, bleaching aged electrolytes, high vapor pressure, corrosivity to metals and

solvent evaporation. Ionic liquids with lower volatility prevent the leakage of the electrolyte but have a slow diffusion of the mediator. Better performing electrolytes should be stable, achieve high IPCEs, and reach V_{OC}s over 800 mV (I_3^-/I^- only yielded 779 mV). The redox couple Br_3^-/Br^- with a V_{OC} of 1.156 V was researched, but the IPCE was too low. On the other hand, pseudohalogen couples like SCN_3^-/SCN^- had a bad stability. Other redox couples showed similar results: High V_{OC}s with low IPCEs or competitive IPCEs with low V_{OC}s. Other problems came along with these substitutes, like early recombinations, mass transport constraints, or high charge transfer resistances. In comparison with the I_3^-/I^- redox couple, these substitutes were less effective and affected the device performance exceptionally. [5] I_3^-/I^- will presumably stay unrivaled in the next time [7].
Hybrid complexes containing both metal and organic materials led to better performances and the need for lower driving forces due to fewer potential losses. Copper or cobalt as auspicious metal components yielded combined with a Ru dye IPCEs of 6.5% in 2001. Higher IPCEs were reached by using a cobalt complex in combination with D-π-A or D-A-π-A organic sensitizers. [4], [5] Compared to cobalt, copper coordination complexes led to fewer losses, because there are no spin transitions upon oxidation and reduction. With these coordination complexes science moved away from solution to solid-state DSSCs. [5]

V. New Developments

A. Emergence of Solid-State DSSCs

With aqueous DSSCs IPCEs of 14.2% were reached with cobalt redox shuttles as metal-organic dyes. Liquid electrolytes are challenging to encapsulate and evaporation affects the stability. Solid-state DSSCs with its solid HTMs do not have these problems, e. g. the by *Grätzel* in 1998 investigated spiro-OMeTAD with IPCEs of 8%. Disadvantages of spiro-OMeTAD are the speed of recombination, the hole mobility, and the high price, as well as thinner TiO_2 films for the pore-filling by the HTM leading to a smaller amount of dye holding on the film. For a commercialization good pore filling by the HTM has to be ensured, the stability must be improved, and the metal oxide films must be thicker. To eradicate the disadvantages, electronically conducting polymers and metal complexes were tested and reached as well high IPCEs by indoor light. In order to absorb more photons in the NIR range, a highly conductive p-type Copper Iodide (CuI) solid-state electrolyte was tested in combination with a multi-dye system. Wavelengths near 900 nm could be absorbed resulting in high J_{SC}s and IPCEs, but also lower V_{OC}s due to recombinations at the CuI-surface. [5]

B. Domestic Devices for the Internet of Things

Without facing harsh conditions in terms of temperature or humidity, DSSCs endure longer lifetimes. In domestic devices DSSCs only witnessed a small impact on the power output after 200 hours of illumination, while solid state DSSCs withstood 1000 hours. The outdoor light spectrum with a big share in the NIR and the red region of the visible part matches the sensitivity of Si solar cells. Indoor light from fluorescent lamps shows a broader spectrum of visible light. DSSCs are sensitive to the panchromatic spectrum and generate higher IPCEs with indoor light than Si cells, which just partly use the whole spectrum. DSSCs can be used for domestic devices, especially in applications for the *Internet of Things* (IoT). [18] They generate enough power indoor and are persistent power sources for IoT devices. With an increasing number of IoT applications in the future more DSSCs will be utilized. [3]

C. Energy Relay Delays

The previously mentioned co-sensitization with a dye cocktail reached a nearly panchromatic spectrum, but it can not be guaranteed that every dye molecule attaches to the limited surface of TiO_2. Co-sensitization expects efficient charge transfers and requires the dye to absorb strongly on the surface. The development of highly luminescent Energy Relay Dyes (ERDs) dissolved in the electrolyte overcame these obstacles. The ERDs absorb higher energy photons and then transfer the energy with the Förster Resonance Energy Transfer (FRET) to the dye. By adding ERDs to the typical DSSC, photons of different energies can be absorbed and more energy can be generated (see Fig. 3). [4], [9], [19]

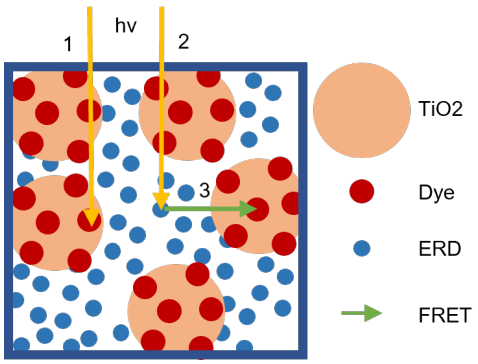

Fig. 3. The energy of the photons can excite a dye molecule (1) or an Energy Relay Dye (ERD) (2), which transfers the energy via Förster Resonance Energy Transfer (FRET) to a dye molecule (3). [9]

VI. Conclusion and Outlook

Since the first photoelectric experiments by *Becquerel* in 1839 solar cells have been investigated intensely. The idea of adding a dye to absorb the panchromatic spectrum was crucial for higher IPCEs and for the competitiveness with the Si-based p-n junction cells, the widespread market leaders. The aim of this paper is to give an overview of the DSSCs by mentioning the recent improvements, such as the modifications of the dye and the electrolyte. The dominant Ru dyes have been replaced step by step by organic dyes in order to save money and the environment. Nowadays, solar cells are a popular, efficient and inevitable approach of generating green energy.
In the future the amount of DSSCs will increase, especially as power sources for IoT devices. In order to achieve higher IPCEs the DSSCs will be investigated more intensely.

REFERENCES

[1] I. Kougias, N. Taylor, G. Kakoulaki, and A. Jäger-Waldau, "The role of photovoltaics for the European Green Deal and the recovery plan," in Renewable and Sustainable Energy Reviews, vol. 144, 2021.

[2] M. Grätzel, "Photoelectrochemical cells," in Nature, vol. 414, 2001, pp. 338–344.

[3] A. Aslam, U. Mehmood, M. H. Arshad, A. Ishfaq, J. Zaheer, A. U. H. Khan, and M. Sufyan, "Dye-sensitized solar cells (DSSCs) as a potential photovoltaic technology for the self-powered internet of things (IoTs) applications," in Solar Energy, vol. 207, 2020, pp. 874–892.

[4] M. Wang, C. Grätzel, S. M. Zakeeruddin, and M. Grätzel, "Recent developments in redox electrolytes for dye-sensitized solar cells," in Energy & Environmental Science, vol.5, 2012, pp. 9394–9405.

[5] M. Stojanović, N. Flores-Diaz, Y. Ren, N. Vlachopoulos, L. Pfeifer, Z. Shen, Y. Liu, S. M. Zakeeruddin, J. V. Milić, A. Hagfeldt, "The Rise of Dye-Sensitized Solar Cells: From Molecular Photovoltaics to Emerging Solid-State Photovoltaic Technologies," in Helvetica Chimica Acta, vol. 104, 2021.

[6] C. Teng, X. Yang, S. Li, M. Cheng, A. Hagfeldt, L. Wu, and L. Sun, "Tuning the HOMO Energy Levels of Organic Dyes for Dye-Sensitized Solar Cells Based on Br^-/Br_3^- Electrolytes," in Chemistry – A European Journal, vol. 16, 2010, pp. 13127–13138

[7] G. Boschloo, and A. Hagfeldt, "Characteristics of the Iodide/Triiodide Redox Mediator in Dye-Sensitized Solar Cells," in Accounts of chemical research, vol. 42, 2009, pp. 1819–1826.

[8] M. Grätzel, "Dye-Sensitized Solid-State Heterojunction Solar Cells," in Mrs Bulletin, vol. 30, 2005, pp. 23–27.

[9] J. Yum, E. Baranoff, S. Wenger, Md. K. Nazeeruddin, and M. Grätzel, "Panchromatic engineering for dye-sensitized solar cells," in Energy & Environmental Science, vol. 4, 2011, pp. 842–857.

[10] G. Xue, X. Yu, T. Yu, C. Bao, J. Zhang, J. Guan, H. Huang, Z. Tang, and Z. Zou, "Understanding of the chopping frequency effect on IPCE measurements for dye-sensitized solar cells: from the viewpoint of electron transport and extinction spectrum," in Journal of Physics D: Applied Physics, vol. 45, 2012.

[11] X. Li, X. Zhang, J. Hua, and H. Tian, "Molecular engineering of organic sensitizers with *o,p*-dialkoxyphenyl-based bulky donors for highly efficient dye-sensitized solar cells," in Molecular Systems Design & Engineering, vol. 2, 2017, pp. 98–122.

[12] S. R. Gajjela, K. Ananthanarayanan, C. Yap, M. Grätzel, and P. Balaya, "Synthesis of mesoporous titanium dioxide by soft template based approach: characterization and application in dye-sensitized solar cells," in Energy & Environmental Science, vol. 3, 2010, pp. 838–845.

[13] J. Desilvestro, M. Grätzel, L. Kavan, J. Moser, and J. Augustynski, "Highly Efficient Sensitization of Titanium Dioxide," in Journal of the American Chemical Society, vol. 107, 1985, pp. 2988–2990

[14] S. Ahmad, E. Guillén, L. Kavan, M. Grätzel, and M. K. Nazeeruddin, "Metal free sensitizer and catalyst for dye sensitized solar cells," in Energy & Environmental Science, vol. 6, 2013, pp. 3439–3466.

[15] A. Mbonyiryivuze, I. Omollo, B. D. Ngom, B. Mwakikunga, S. M. Dhlamini, E. Park, and M. Maaza, "Natural Dye Sensitizer for Grätzel Cells: Sepia Melanin," in Physics and Materials Chemistry, vol. 3, 2015, pp. 1–6.

[16] O. F. Restaino, M. Scognamiglio, S. F. Mirpoor, M. Cammarota, R. Ventriglia, C. V. L. Giosafatto, A. Fiorentino, R. Porta, and C. Schiraldi, "Enhanced *Streptomyces roseochromogenes* melanin production by using the marine renewable source *Posidonia oceanica* egagropili," in Applied Microbiology and Biotechnology, vol. 106, 2022, pp. 7265–7283.

[17] K. Sharma, V. Sharma, and S. S. Sharma, "Dye-Sensitized Solar Cells: Fundamentals and Current Status," in Nanoscale Research Letters, vol. 13, 2018, pp. 1–46.

[18] H. Michaels, M. Rinderle, R. Freitag, I. Benesperi, T. Edvinsson, R. Socher, A. Gagliardi, and M. Freitag, "Dye-sensitized solar cells under ambient light powering machine learning: towards autonomous smart sensors for the internet of things," in Chemical Science, vol. 11, 2020, pp. 2895–2906.

[19] E. Lee, C. Kim, and J. Jang, "High-Performance Förster Resonance Energy Transfer (FRET)-Based Dye-Sensitized Solar Cells: Rational Design of Quantum Dots for Wide Solar-Spectrum Utilization," in Chemistry – A European Journal, vol. 19, 2013, pp. 10280–10286.

GaN and SiC for power electronics: Comparison for various applications

Florian Thalhammer
Faculty of Electrical Engineering and Information Technology
Ostbayerische Technische Hochschule
Regensburg, Germany
florian1.thalhammer@st.oth-regensburg.de

***Abstract*—The requirements in modern power electronics are becoming increasingly challenging. The currently most commonly used power semiconductors made of silicon (Si) are reaching the limits of their thermal and electrical load capabilities. In order to meet the increased requirements, new semiconductors have been researched for some time already. These so-called wide band gap (WBG) semiconductors, such as silicon carbide (SiC) and gallium nitride (GaN), have the ability to operate at higher voltages, temperatures and frequencies, which makes them extremely suitable for future technologies. However, there is still some research to be done before the materials can be used on a widespread basis. While SiC has matured to the point where the technology can be used commercially, GaN still needs further study to exploit the full potential of the material. Although both technologies are wide bandgap, there are fundamental differences between GaN and SiC that make one more suitable than the other for certain topologies and applications. Semiconductors made of SiC, for example, are primarily used in high power applications, while semiconductors made of GaN are more commonly used in high frequency applications. But there are also overlaps. In order to provide a possibility to decide between one of the two semiconductor technologies, depending on the use case, this paper presents a systematic comparison between GaN and SiC power devices. For this purpose, both technologies are examined and compared in terms of their fundamental material properties, which allows conclusions to be drawn about their advantages and disadvantages in certain applications. The manufacturing process and the associated cost factor will also be discussed as an important issue for future viability. Finally, the results are evaluated and the technologies are assigned to the various applications in power electronics.**

***Index Terms*—power electronics, SiC, silicon carbide, GaN, gallium nitride, wide band gap, WBG, power devices**

I. Introduction

Current research and development in the field of power semiconductors is constantly trying to optimize the characteristics of electronic components. On the one hand, it is crucial to improve their efficiency. On the other hand, the constant miniaturization of electronic devices requires components to become increasingly smaller. For some time now, research has been conducted into ways of accomplishing this at the semiconductor level. One promising approach is to abandon silicon (Si) and use wide band gap (WBG) semiconductors, such as gallium nitride (GaN) and silicon carbide (SiC).
As the name implies, this type of semiconductor material has a larger band gap than silicon, which offers new technical opportunities. Power semiconductors made of GaN and SiC have better efficiency, can be used at higher switching frequencies and are overall more robust at higher temperatures than their silicon counterparts. [1]
Despite all the advantages, the rapidly changing currents and voltages cause harmonics resp. electromagnetic interference (EMI) as the switching frequency increases [2]. Therefore, filters must be used for interference suppression. [3]
The two technologies also need more work on research to reduce cost and manufacturing effort as well as to increase stability and reliability [4], [5].
Although GaN and SiC are both materials for wide band gap semiconductors and thus have some of the same advantages and disadvantages, their different material properties make them suitable for different applications. This work therefore presents a systematic comparison between GaN and SiC and is intended to help the readership make a quick decision for the desired application of power electronics. The paper is structured as follows: In Section II, some of the most important physical properties of the two semiconductor materials are described in detail. Furthermore, their effects on the behavior of the components made from them are explained. Then, in Section III, the two semiconductor technologies are compared in terms of their physical properties. This also includes a comparison of economic factors such as production effort and associated costs. Finally, the individual applications are assigned to GaN and SiC in Section IV. A final conclusion is given in Section V.

II. Electrical and Material Properties

For better clarity, the individual values have been summarized in Table I. The data is a collection of several literature references [2], [4]–[10]. Since different values for the physical properties occur in the literature references, the minimum and maximum values of these are therefore shown in the table. The individual properties will be discussed in more detail in the following subsections.

A. Bandgap E_G

In a solid, electrons exist at certain energy states and form energy bands. Figure 1 shows a simplified energy band diagram. The upper band is the conduction band. Underneath is the valence band. Between the two bands is the forbidden band, where no electrons exist. In order for a current to flow,

TABLE I
PHYSICAL PROPERTIES OF GAN AND SIC [2], [4]–[10]

Property	Unit	SiC	GaN
Bandgap E_G	eV	3.2–3.3	3.4–3.5
Electric breakdown field E_C	MV/cm	2.0–3.5	2.5–3.5
Electron mobility μ_n	$10^3 cm^2/Vs$	0.7–1.1	0.9–2.0
Thermal conductivity λ	W/cm·K	3.7–5.0	1.1–1.5
Saturated drift velocity v_{sat}	$10^7 cm/s$	2.0	2.0–2.8

electrons must move from the valence band to the conduction band, i.e. they must have a certain amount of energy. This energy respectively energy difference between conduction and valence band is the bandgap. It is specified in electron volts. In conductors, the forbidden band does not exist. Conduction band and valence band overlap. In non-conductors, the gap is so wide that electrons need a lot of energy to move from the valance band to the conduction band. In semiconductors the gap is smaller than for a non-conductor. [11]
The energy required by the electrons to reach the conduction band can be supplied in a controlled manner via an external voltage. With an increase in temperature, however, the thermal energy of the electrons in the valence band increases. As a result, the electrons can also reach the conduction band, which is an uncontrolled conduction. This intrinsic temperature at which the semiconductor becomes a conductor increases with the width of the bandgap. Thus, the wider the band gap, the more heat semiconductors can withstand without losing their electrical characteristics. [11]

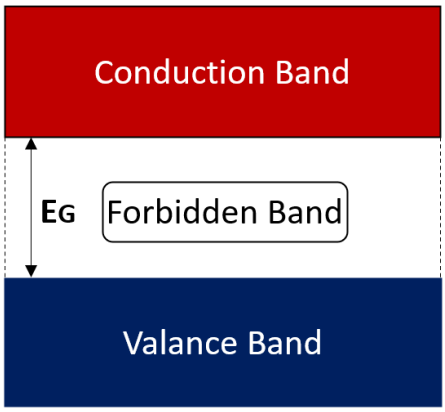

Fig. 1. Simplified energy band diagram (adapted from [11])

B. Electric Breakdown Field E_C and Electron Mobility μ_n

The electrical breakdown field directly determines the breakdown voltage, i.e. the voltage that may be applied in reverse direction. Furthermore, due to a higher electrical breakdown field, higher doping levels and thus smaller device layers can be achieved [11]. This reduces the width of the drift region, which affects the on-resistance of the component. Equation 1 shows the formula for the on-resistance of an unipolar device.

$$R_{on} = \frac{4 \cdot V_R}{\epsilon_0 \cdot \epsilon \cdot \mu_n \cdot {E_C}^3} \tag{1}$$

Whereby ϵ represents the dielectric constant, μ_n the electron mobility and V_R the breakdown voltage of the semiconductor material. Electron mobility and electric breakdown field strength are found in the denominator of the equation. Therefore, the higher these two parameters are, the smaller the on-resistance. The low on-resistance leads into a lower power dissipation of the devices, which is a key requirement for a better energy efficiency. [6], [12]

C. Thermal Conductivity λ

In order for a power semiconductor to effectively dissipate the resulting heat loss to the environment via the case and heat sink, it must have a high thermal conductivity. The thermal conductivity of the material is indirectly proportional to the junction-to-case thermal resistance $R_{th\text{-}jc}$ of the power semiconductor as seen in Equation 2.

$$R_{th\text{-}jc} = \frac{d}{\lambda \cdot A} \tag{2}$$

Thereby d is the layer thickness of the semiconductor material, and A the cross-sectional area of heat transfer. With a large cross-sectional area, the heat can be dissipated better. However, this goes along with a high space requirement. Components with high thermal conductivity therefore possess a low junction-to-case thermal resistance and can thus keep the space requirement low at the same time. This is critical for high-temperature operation as the device temperature increases more slowly. [6], [11]

D. Saturated Drift Velocity v_{sat}

The ability of a power component to be used for high switching frequencies is directly proportional to its saturated drift velocity which is the maximum speed that carriers can reach in the channel zone. The higher this velocity, the better the frequency response of the power component. A higher saturated drift velocity also enables a faster depletion of the charge in the depletion region of a diode. Therefore, the reverse current of wide band gap diodes is lower and the reverse recovery time is overall shorter than that of a silicon-based diode. [6], [11]

III. COMPARISON OF GAN AND SIC

A. Graphical Comparison

For the graphical evaluation (see Figure 2) of the electrical and material properties of the two semiconductor technologies, the maximum values of the data collection of Table I were used. Initial conclusions can already be drawn from this. Due to the much higher thermal conductivity and the approximately equal band gap respectively electric breakdown field of silicon carbide, power electronic components made of SiC are more

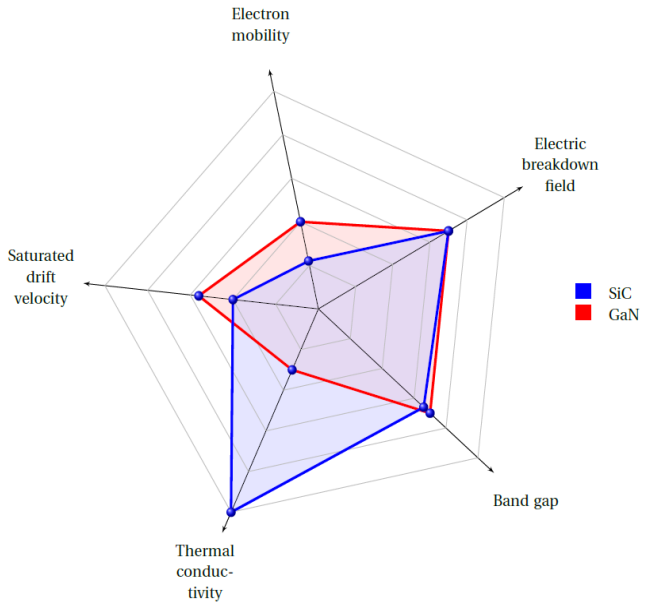

Fig. 2. Graphical comparison of physical properties of SiC and GaN [2], [4]–[10]

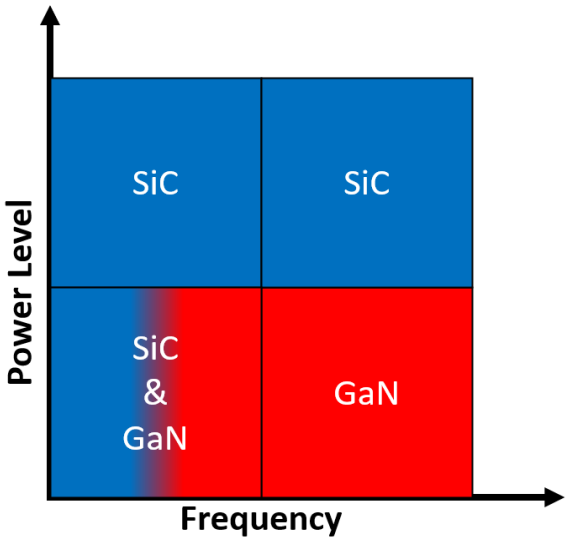

Fig. 3. Application trends for SiC and GaN technologies (adapted from [19])

likely to be used for high-power applications [13]–[17]. GaN, on the other hand, excels with a high saturation drift rate and a high electron mobility. Semiconductors made of GaN therefore feature better switching behavior overall [13]–[17]. According to [18], higher losses occur in GaN power semiconductors than in SiC at higher temperatures. The losses of GaN therefore depend very much on the junction temperature [18]. This is also supported by [19]. In this work, 91 power MOSFETs made of GaN, SiC and Si were used and the most important properties of the individual devices were investigated for their temperature and frequency dependence. From these results, an application trend was developed (see Figure 3). In it, the respective areas of application of GaN and SiC are classified according to frequency and power level.
Therefore, it can be stated that GaN is very suitable for high frequencies, but only if the switched power is correspondingly low. For high power SiC is used throughout. [13], [19], [20]
At relatively low power levels and frequencies, both technologies could theoretically be used. However, the advantages that amortize the high costs of devices made of GaN and SiC are not fully exploited here. Therefore, power semiconductors made of silicon would be used in this area at the current state of the art. [20]

B. Cost Comparison

To measure the future viability of the two technologies, the cost composition of a power semiconductor made of SiC and GaN must be considered. These costs are largely determined by the costs for substrate and fabrication. The more chips fit on a wafer, the lower the cost of a single component (see Equation 3). [13]

$$Device\ cost = \frac{substrate\ cost + fabrication\ cost}{chips\ per\ wafer} \tag{3}$$

Cost projections forecast that GaN devices will be cheaper to manufacture than Si devices in the coming years. According to these analyses, SiC will lag behind gallium nitride and silicon. [4], [13]
The individual factors that contribute to Equation 3 will be discussed in the following subsections.

1) Chips per Wafer: The number of chips per wafer is a function of various factors including wafer size and thermal resistivity of the used semiconductor technology. SiC has better thermal resistivity than GaN and therefore allows a higher number of chips per wafer. [13]

2) Substrate and Fabrication Cost: A significant disadvantage of power semiconductors made of SiC are the relatively high manufacturing costs. A large part of this is accounted for by the substrate costs, which are particularly high for SiC. While GaN devices can be grown on low-cost and already researched silicon substrates and manufacturers thus benefit from already existing manufacturing tools, SiC requires special substrate and manufacturing processes. This includes process steps for which temperatures of 2500 °C are required. Especially in times of a global energy crisis, this is a decisive disadvantage. [13]

IV. Applications

Power semiconductors made of GaN and SiC are applied in various fields of applications. In the product selections guide of a well-known semiconductor manufacturer [21], which serves as a reference for choosing the right field of application for their state-of-the-art products, the various application areas for GaN and SiC are divides into three categories: Automotive, Industrial and Consumer. Various application areas of power

electronics are assigned to these three supercategories, for which the corresponding components from their product range are recommended.

A. Automotive

The automotive sector includes applications in the field of e-mobility. Here, only power semiconductors made of SiC are used, as very high power must be provided. The individual applications are listed in Table II.

TABLE II
SiC and GaN in the automotive sector [21]

Application	SiC	GaN
On-board charger (OBC)	✓	✗
High-voltage to low-voltage DC-DC converter	✓	✗

B. Industrial

The industrial sector covers a very wide range of products. It includes applications in the field of energy, e-mobility, power supply for larger infrastructures as well as battery-powered applications. The required power therefore varies between medium and high, which also results in application areas for GaN. The individual applications are listed in Table III.

TABLE III
SiC and GaN in the industrial sector [21]

Application	SiC	GaN
Electric vehicle charging (EV charging)	✓	✗
LSEV (Industrial battery charger)	✓	✗
Battery formation	✓	✗
Server power supply	✓	✓
Telecom power supply	✓	✓
Industrial SMPS (Switch-mode power supplies)	✓	✓
Solar power systems	✓	✓
UPS (Uninterruptible Power Supply)	✓	✓

C. Consumer

The consumer segment mainly includes power supplies for entertainment electronics such as televisions, monitors and portable devices. Household appliances and LED lighting technology are also part of this area. The required power is therefore relatively low compared to the areas already presented. Power semiconductors made of GaN are suitable for this. The individual applications are listed in Table IV.

TABLE IV
SiC and GaN in the consumer sector [21]

Application	SiC	GaN
TV/PC/Gaming power supplies	✗	✓
Adapter and charger power supply	✗	✓
LED lighting	✗	✓
Audio power supply	✗	✓
Major home appliances	✗	✓

V. Conclusion

The aim of this paper is to present a categorical comparison between gallium nitride (GaN) and silicon carbide (SiC) for the various applications in power electronics in order to make a decision between these two innovative and seminal semiconductor technologies. Due to the favorable physical properties of gallium nitride, GaN devices possess outstanding switching behavior even at very high frequencies. However, due to the poor thermal conductivity of gallium nitride, they should only be used at low to medium power levels. Power semiconductors made of silicon carbide are characterized by excellent thermal conductivity and low switching losses at low to medium frequencies. Therefore, they are the best choice for high power applications at moderate frequencies. This paper has given a survey of the technologies and their suitable fields of application. However, the decision must always be made individually depending on the exact boundary conditions of the end product. The respective application engineers of the semiconductor manufacturer should always be consulted for assistance.

References

[1] I. C. Kizilyalli, Y. A. Xu, E. Carlson, J. Manser and D. W. Cunningham, "Current and future directions in power electronic devices and circuits based on wide band-gap semiconductors," 2017 IEEE 5th Workshop on Wide Bandgap Power Devices and Applications (WiPDA), 2017, pp. 417-417, doi: 10.1109/WiPDA.2017.8170583.

[2] X. Yuan, "Application of silicon carbide (SiC) power devices: Opportunities, challenges and potential solutions," IECON 2017 - 43rd Annual Conference of the IEEE Industrial Electronics Society, 2017, pp. 893-900, doi: 10.1109/IECON.2017.8216154.

[3] I. Aretxabaleta, I. M. de Alegría, J. I. Garate, A. Matallana and J. Andreu, "Wide-Bandgap Semiconductor HF-Oscillation Attenuation Method With Tuned Gate RLC Filter," in IEEE Transactions on Power Electronics, vol. 35, no. 8, pp. 8025-8033, Aug. 2020, doi: 10.1109/TPEL.2020.2964272.

[4] A. Ghazanfari, C. Perreault and K. Zaghib, "EV/HEV Industry Trends of Wide-bandgap Power Semiconductor Devices for Power Electronics Converters," 2019 IEEE 28th International Symposium on Industrial Electronics (ISIE), 2019, pp. 1917-1923, doi: 10.1109/ISIE.2019.8781528.

[5] X. Ding, Y. Zhou and J. Cheng, "A review of gallium nitride power device and its applications in motor drive," in CES Transactions on Electrical Machines and Systems, vol. 3, no. 1, pp. 54-64, March 2019, doi: 10.30941/CESTEMS.2019.00008.

[6] L. F. S. Alves et al., "SIC power devices in power electronics: An overview," 2017 Brazilian Power Electronics Conference (COBEP), 2017, pp. 1-8, doi: 10.1109/COBEP.2017.8257396.

[7] G. E. Town, ”Gallium nitride power electronic devices and circuits: A review,” 2015 IEEE 11th International Conference on Power Electronics and Drive Systems, 2015, pp. 1-3, doi: 10.1109/PEDS.2015.7203483.
[8] Flack, T.J., Pushpakaran, B.N. and Bayne, S.B. GaN Technology for Power Electronic Applications: A Review. J. Electron. Mater. 45, 2673–2682 (2016). https://doi.org/10.1007/s11664-016-4435-3
[9] R. Adappa, K. Suryanarayana, H. Swathi Hatwar and M. Ravikiran Rao, ”Review of SiC based Power Semiconductor Devices and their Applications,” 2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT), 2019, pp. 1197-1202, doi: 10.1109/ICICICT46008.2019.8993255.
[10] A. Choudhury, ”Present Status of SiC based Power Converters and Gate Drivers – A Review,” 2018 International Power Electronics Conference (IPEC-Niigata 2018 -ECCE Asia), 2018, pp. 3401-3405, doi: 10.23919/IPEC.2018.8507554.
[11] H. Jain, S. Rajawat and P. Agrawal, ”Comparision of wide band gap semiconductors for power electronics applications,” 2008 International Conference on Recent Advances in Microwave Theory and Applications, 2008, pp. 878-881, doi: 10.1109/AMTA.2008.4763184.
[12] Roccaforte, F., Fiorenza, P., Lo Nigro, R. et al. Physics and technology of gallium nitride materials for power electronics. Riv. Nuovo Cim. 41, 625–681 (2018). https://doi.org/10.1393/ncr/i2018-10154-x.
[13] Masoud Beheshti, ”Wide-bandgap semiconductors: Performance and benefits of GaN versus SiC,” [Online]. Available: https://www.ti.com/lit/an/slyt801/slyt801.pdf. [Accessed: 2022-11-24].
[14] P. Palmer, X. Zhang, E. Shelton, T. Zhang and J. Zhang, ”An experimental comparison of GaN, SiC and Si switching power devices,” IECON 2017 - 43rd Annual Conference of the IEEE Industrial Electronics Society, 2017, pp. 780-785, doi: 10.1109/IECON.2017.8216135.
[15] M. Treu et al., ”The role of silicon, silicon carbide and gallium nitride in power electronics,” 2012 International Electron Devices Meeting, 2012, pp. 7.1.1-7.1.4, doi: 10.1109/IEDM.2012.6478995.
[16] D. Ji, W. Li and S. Chowdhury, ”Potential of GaN vertical JFETs presented through a comprehensive discussion of dynamic performance compared to SiC JFETs,” 2016 IEEE 4th Workshop on Wide Bandgap Power Devices and Applications (WiPDA), 2016, pp. 114-117, doi: 10.1109/WiPDA.2016.7799920.
[17] G. Deboy, M. Treu, O. Haeberlen and D. Neumayr, ”Si, SiC and GaN power devices: An unbiased view on key performance indicators,” 2016 IEEE International Electron Devices Meeting (IEDM), 2016, pp. 20.2.1-20.2.4, doi: 10.1109/IEDM.2016.7838458.
[18] F. M. Shah, H. M. Xiao, R. Li, M. Awais, G. Zhou and G. T. Bitew, ”Comparative performance evaluation of temperature dependent characteristics and power converter using GaN, SiC and Si power devices,” 2018 IEEE 12th International Conference on Compatibility, Power Electronics and Power Engineering (CPE-POWERENG 2018), 2018, pp. 1-7, doi: 10.1109/CPE.2018.8372523.
[19] Prado, E.O.; Bolsi, P.C.; Sartori, H.C.; Pinheiro, J.R. ”An Overview about Si, Superjunction, SiC and GaN Power MOSFET Technologies in Power Electronics Applications”, Energies 2022, 15, 5244. https://doi.org/10.3390/en15145244.
[20] Infineon Technologies Austria AG, ”Infineon masters it all – for you,” [Online]. Available: https://www.infineon.com/dgdl/Infineon-Experience-the-difference-in-power-with-CoolMOS-CoolSiC_CoolGaN-ProductBrief-v06_00-EN.pdf?fileId=5546d462636cc8fb0163fe86543408b0. [Accessed: 2022-11-24].
[21] Infineon Technologies Austria AG, ”High-voltage switches 500V - 950V CoolMOS™ - CoolSiC™ - CoolGaN™,” [Online]. Available: https://www.infineon.com/dgdl/Infineon-Experience_the_difference_in_power_CoolMOS_CoolSiC_CoolGaN-ProductSelectionGuide-v02_00-EN.pdf?fileId=5546d46262b31d2e0162b59d4e45079a. [Accessed: 2022-11-24].

Description and Improvement of the MUSIC algorithm

Johannes Mauderer
Faculty of Electrical Engineering and Information Technology
Ostbayerische Technische Hochschule Regensburg
Regensburg, Germany
johannes.mauderer@st.oth-regensburg.de

***Abstract*—The estimation of a signal's direction finds practical application in different areas like radar, sonar and communication. In communication, the Direction of Arrival (DOA) estimation is applied in smart antennas. The information about the direction of the received signal can be used to adapt the antenna pattern dynamically. To get the DOA of the incoming signal, smart antennas are using antenna arrays and digital signal processing techniques. A known method in signal processing is the Multiple Signal Classification (MUSIC) algorithm. This algorithm can detect the directions of various incoming incoherent signals by searching for peaks in the spatial spectrum. Therefore, this paper describes the general function of the MUSIC algorithm. During the past years, there have been some improvements regarding the performance of the algorithm. It is shown that the amount of antenna elements in an array, the number of snapshots and the Signal to Noise Ratio (SNR) have an impact on the resolution of the spectrum and accordingly also on the accuracy of the direction estimation. Another technique reveals that by enhancing the algorithm the resolution can be improved as well. As the conventional MUSIC algorithm can only detect incoherent signals the algorithm has been adapted to calculate the DOA of coherent signals too. An alteration of the MUSIC algorithm is the root-MUSIC algorithm. This method calculates the direction of the signal out of a polynomial which is based on the same eigenvalues as the MUSIC algorithm. As the eigenvalues of these two methods are the same, this paper discusses the differences between the algorithms. Also, the performance of the root-MUSIC and the MUSIC algorithm is compared to each other.**

***Index Terms*—Multiple signal classification (MUSIC), Direction-of-arrival (DOA) estimation, signal processing, Root-MUSIC**

I. Introduction

In radar, sonar, communications and a lot more applications, it is important to know the direction from which a signal is received [1]. In radar for example it is necessary to know the location of various scattering centers [2].
To detect the direction of an incoming signal, a fixed antenna can be used. However, a fixed antenna has the disadvantage that its resolution of the Direction of Arrival (DOA) depends on its Main Lobe Beam Width (MLBW). To get a better resolution a sharper MLBW is needed. As the physical size of the antenna is inversely proportional to the MLBW, the size of the antenna has to be increased. This is not always practical and therefore, an antenna array is used instead of a fixed antenna. Antenna arrays offer a narrow MLBW which is able to enhance the resolution of the DOA [3], [4]. To process the received signals and to estimate the DOA, signal processing methods are used. Some of these methods are the Multiple Signal Classification (MUSIC) and root-MUSIC algorithm [4]. The combination of an antenna array and a signal processing method is found in smart antennas. These smart antennas use the information of the DOA to adapt the antenna pattern dynamically. One area of application for smart antennas is in communications [5], [6].
As both algorithms, MUSIC and root-MUSIC, can be used in smart antennas and both algorithms are based on calculating the eigenvalues out of the estimated covariance matrix [1], first, this paper will show the general function of the MUSIC algorithm in chapter II. Due to its invention in 1979 by R. Schmidt, there have been many improvements to the algorithm. Some of the enhancements are described and categorized in this paper. The description of some physical improvements can be found in chapter III-A. Several upgrades of the algorithm itself are explained in chapter III-B. The next chapter deals with the function of the root-MUSIC algorithm. It shows how the eigenvalues of MUSIC are used to get the polynomial for the DOA estimation. After that follows a comparison of the MUSIC and root-MUSIC algorithm in chapter V.

II. General function of the MUSIC algorithm

To use the MUSIC algorithm there has to be an antenna array, like it is shown in Figure 1 for example. M elements in the array are receiving N signal sources $s(t)$ from different angles Θ_i $(1 \leq i \leq N)$. For a good estimation of the DOA, there have to be more antenna elements than signal sources $(M > N)$. The elements are separated by a known distance (d)

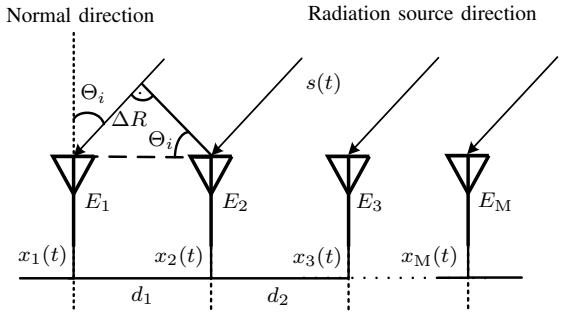

Figure 1. Structure of an one-dimensional antenna array. (based on [7])

from each other [3] [4]. An even distribution of the elements is not required. The distances can be described by a vector $[d_1, d_2, \cdots, d_\mathrm{M}]^\mathrm{T}$ [4]. The received signal can be expressed as in (1).

$$x(t) = \boldsymbol{A}\, s(t) + n(t) \tag{1}$$

$x(t)$ is the vector of the signals received by the antenna array:

$$x(t) = [x_1(t), x_2(t), \cdots, x_\mathrm{M}(t)]^\mathrm{T}. \tag{2}$$

$s(t)$ are the incident signals excited by N signal sources:

$$s(t) = [s_1(t), s_2(t), \cdots, s_\mathrm{N}(t)]^\mathrm{T}. \tag{3}$$

The noise which is received and generated by the different parts of the antenna is represented by

$$n(t) = [n_1(t), n_2(t), \cdots, n_\mathrm{M}(t)]^\mathrm{T}. \tag{4}$$

The noise consists of white gaussian noise with a variance of σ^2 and a mean value of zero. The array steering matrix $\boldsymbol{A}$ is defined as in (5). The matrix consists of N columns every column can be represented as in (6). Equation (7) shows the calculation of the phase difference ϵ between each antenna element. For a better understanding of the algorithm, the distances between the antenna elements are all the same in this case $(d_1 = d_2 = \cdots = d_\mathrm{M})$. The Term $d\sin(\Theta_i)$ in (7) describes the ΔR which can be seen in Figure 1. The wavelength of the signal is represented by λ [3]–[5], [8]–[10]. After getting the received signal vector $x(t)$ out of the antenna array the MUSIC algorithm calculates the covariance matrix

$$\boldsymbol{A} = [a(\epsilon_1), a(\epsilon_2), \cdots, a(\epsilon_\mathrm{N})]^\mathrm{T} \tag{5}$$

$$a(\epsilon_i) = \left[1, \mathrm{e}^{-j\epsilon}, \mathrm{e}^{-j2\epsilon}, \cdots, \mathrm{e}^{-j(M-1)\epsilon}\right]^\mathrm{T} \tag{6}$$

$$\epsilon = \frac{2\pi\, d}{\lambda}\sin(\Theta_i); \qquad 1 \leq k \leq M \tag{7}$$

with Equation (8). In the Equation E stands for the expectation value and $^\mathrm{H}$ indicates the hermitian conjugate transpose of the vector or matrix. $\boldsymbol{R}_\mathrm{s}$ is the covariance matrix of the signal $s(t)$, σ^2 embodies the noise variance and $\boldsymbol{I}_\mathrm{M}$ stand for the $M \times M$ identity matrix [3]–[5], [8]–[10]. Instead of the phrase covariance matrix some papers use correlation matrix but those expressions are used interchangeably in the referred papers [11].

$$\boldsymbol{R}_\mathrm{x} = \mathrm{E}\left[x\, x^\mathrm{H}\right] \tag{8}$$

$$\boldsymbol{R}_\mathrm{x} = \boldsymbol{A}\, \boldsymbol{R}_\mathrm{s}\, \boldsymbol{A}^\mathrm{H} + \sigma^2\, \boldsymbol{I}_\mathrm{M} \tag{9}$$

Next, the eigenvalues of $\boldsymbol{R}_\mathrm{x}$ have to be calculated by using the eigenvalue decomposition. When the decomposition is finished, there are M eigenvalues γ [3]–[5], [8]. To estimate the DOA these eigenvalues should be sorted so that γ_1 is the maximum and γ_M is the minimum. As the minimum eigenvalue of $\boldsymbol{A}\, \boldsymbol{R}_\mathrm{s}\, \boldsymbol{A}^\mathrm{H}$ equals zero there must be $M - N$ eigenvalues corresponding to σ^2 [3], [8]. Now it is possible to define two subspaces, the signal subspace with

$$\gamma_\mathrm{s} = [\gamma_1, \gamma_2, \cdots, \gamma_\mathrm{N}]^\mathrm{T} \tag{10}$$

and the noise subspace with

$$\gamma_\mathrm{n} = [\gamma_{\mathrm{N}+1}, \gamma_{\mathrm{N}+2}, \cdots, \gamma_\mathrm{M}]^\mathrm{T}. \tag{11}$$

Since the array steering vector from Equation (6) and (7) is orthogonal to the noise subspace the DOA can be calculated by searching for peaks in the spatial spectrum using (12) [3]–[5], [8]. The estimation of the DOA with the just shown MUSIC algorithm only works if the signals $s(t)$ are incoherent [5], [11].

$$P_\mathrm{MU}(\Theta) = \frac{1}{a(\Theta)^\mathrm{H}\, \gamma_\mathrm{n}\, \gamma_\mathrm{n}^\mathrm{H}\, a(\Theta)} \tag{12}$$

III. Improvements

As mentioned above, MUSIC has difficulties in estimating the DOA of two or more coherent signals [11]. Furthermore, the resolution of the spatial spectrum decreases when some signals which are incoherent lie very close to each other, thus the peaks can not be detected clearly [5]. Therefore, some improvements have been made to increase the performance of the MUSIC algorithm. The following methods are sorted by physical improvements which involve the antenna array, the Signal to Noise Ratio (SNR) and so on. The other category includes enhancements to the algorithm.

A. Physical Improvements

One approach to improve MUSIC is done by [6] and [12]. In those papers, it can be seen that the resolution and therefore the accuracy of the spatial spectrum depends on the SNR, the array elements and the number of snapshots. The snapshots are taken from the continuous receiving signal. The more snapshots are taken, the more array elements are involved and the higher the SNR is, the better the resolution of the spatial spectrum.
Another method is able to detect coherent and centralized underwater targets by rearranging the array elements. The distance between the elements is chosen randomly and the distribution is the normal distribution. [7]
Also, there is a way to improve the resolution in two dimensions. The method uses a Multiple Input Multiple Output (MIMO) antenna array. When the antenna array is receiving an Orthogonal Frequency-Division Multiplexing (OFDM) signal, this strategy aims to detect the distance and the angle of an incoming signal. The results depict that in a 2D spatial spectrum the 4×4 MIMO antenna performs better than a 2×2 antenna. [13]

B. Algorithm Improvements

An improvement of the resolution can not only be done in the physical area. It is also possible to enhance the resolution of the spectrum by adding some factors to the algorithm. These factors could be an array elements factor, a snapshot factor or a signal power factor and they enlarge the real parameter virtually. [5]
MUSIC combined with Blind Source Separation (BSS) is used in [14]. This method has the ability to detect a ten times weaker signal compared to the power of the other signals in the spectrum. Furthermore, two correlated sources are shown in the spatial spectrum due to the combination of MUSIC with

BSS. On the other side, the conventional MUSIC algorithm outperforms the modified version if the signals are close to each other.

Another strategy concentrates on detecting coherent signals as well. Therefore, the algorithm is extended by a new $M \times M$ matrix called Jordan matrix, see Equation (13). If a coherent signal is now received the Jordan matrix comes into play. Due to this extension, the covariance matrix $\boldsymbol{R}$ is calculated as in (14) where $\boldsymbol{R}_{\mathrm{Y}}$ can be described by (15). The covariance matrix of the received signal ($\boldsymbol{R}_{\mathrm{X}}$) is known from (8). The vector Y stands for the multiplication of the Jordan matrix with the complex conjugate of the received signal vector $x(t)$, which can be seen in Equation (16). [15]

$$\boldsymbol{J} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ \vdots & & \cdots & & \vdots \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix} \tag{13}$$

$$\boldsymbol{R} = \frac{\boldsymbol{R}_{\mathrm{X}} + \boldsymbol{R}_{\mathrm{Y}}}{2} \tag{14}$$

$$\boldsymbol{R}_{\mathrm{Y}} = \mathrm{E}\left[Y\, Y^{\mathrm{H}}\right] \tag{15}$$

$$Y = \boldsymbol{J}\, x(t)^{*} \tag{16}$$

Nowadays, Deep Learning (DL) is used to upgrade the performance of MUSIC. Simulation results show that for multiple signals and a SNR of $-20\,\mathrm{dB}$ the Root Mean Squared Error (RMSE) of the DL technique is roundabout two times lower than for the conventional MUSIC. For signals with a correlation coefficient from $0.6 - 1$, few received signals and a SNR of $30\,\mathrm{dB}$ the RMSE of DL MUSIC is greater than the one for MUSIC. Due to the combination with DL, the new approach has a lower computational complexity which can be seen in a faster computation time. MUSIC combined with DL needs $0.0020\,\mathrm{s}$ to calculate the spectrum. In contrast, the calculation of the spectrum with the conventional MUSIC algorithm takes $0.030\,\mathrm{s}$. [16]

As in chapter III-A, there is also a technique to improve the algorithm so it can be used as a two-dimensional DOA method. The spatial spectrum provides an accurate estimation of the distance and the angle of five closely spaced targets. The angular displacement lies between $-2°$ and $2°$ and the distance variates from $19.4\,\mathrm{m}$ to $20.6\,\mathrm{m}$. [17]

IV. The root-MUSIC algorithm

The root-MUSIC algorithm is an extension of the spatial MUSIC algorithm. Due to the fact that MUSIC has to sweep over the angles to detect the DOA, the implementation in applications can be very expensive. Root-MUSIC therefore provides a lower computational complexity than MUSIC [18]. Root-MUSIC is based on the same eigenvalues as MUSIC. Instead of searching for peaks in the spatial spectrum, root-MUSIC estimates the DOA out of a polynomial which uses the eigenvalues of the noise subspace [18]–[20]. In contrast to MUSIC, root-MUSIC can only be applied in systems where the array consists of a Uniform Linear Array (ULA) [21].

To get the polynomial, Equation (12) is expressed as in (17) with B according to the substitution in (18) [18]. As the antenna array is a ULA there is the same distance ($d_1 = d_2 = \cdots = d_{\mathrm{M}}$) between the elements. Therefore, the

$$P_{\mathrm{RMU}}(\Theta) = \frac{1}{a(\Theta)^{\mathrm{H}}\, B\, a(\Theta)} \tag{17}$$

$$B = \gamma_{\mathrm{n}}\, \gamma_{\mathrm{n}}^{\mathrm{H}} \tag{18}$$

steering vector in (6) and (7) can be expressed as in (19). By using the transformation in Equation (20) the steering vector can be expressed as in (21). The denominator of (17) depends now on

$$a(\Theta) = \left[1,\, \mathrm{e}^{-j\frac{2\pi\, d}{\lambda}\sin(\Theta)},\, \cdots,\, \mathrm{e}^{-j(M-1)\frac{2\pi\, d}{\lambda}\sin(\Theta)}\right]^{\mathrm{T}} \tag{19}$$

$$z = \mathrm{e}^{j\frac{2\pi\, d}{\lambda}\,\sin(\Theta)} \tag{20}$$

$$a(z) \quad = \left[1,\, z^{-1},\, \cdots,\, z^{-(M-1)}\right]^{\mathrm{T}} \tag{21}$$

the variable z, not Θ anymore, like Equation (22) depicts. From the spatial MUSIC it is known that the DOA matches with the root of the denominator of Equation (17). Hence, the roots of (22) also interfere with the DOA [12], [18]–[20].

$$D(z) = a(z)^{\mathrm{H}}\, B\, a(z) \tag{22}$$

After calculating the roots, there are $2\,(M-1)$ roots [18]. Since the number of received signals (N) has to be smaller than the number of antenna elements (M) there are more roots than received signals. Thus, to get the correct roots the N results which lie the closest to the unit circle are chosen. Every resulting root of the denominator can be expressed in the complex form of:

$$z = |z|\, \mathrm{e}^{j\,\arg(z)}. \tag{23}$$

In the formula, the magnitude of the root is represented by $|z|$. The argument of the complex root is shown by $\arg(z)$. Considering Equation (20) the angle of the received signal is calculated by comparing both exponents and solving the Equation for Θ. Accordingly, the formula for each arriving signal describes Equation (24) [12], [18], [19].

$$\Theta_i = -\sin^{-1}\left(\frac{\lambda}{2\pi\, d}\arg(z_i)\right); \qquad 1 \le i \le N. \tag{24}$$

V. Comparison of MUSIC and root-MUSIC

As seen in the previous chapter, the MUSIC and root-MUSIC algorithms are based on the same eigenvalues. Due to that, it is interesting to know which algorithm performs better in the case of the DOA estimation. The following comparisons consider only the spatial MUSIC without any improvement. Furthermore, for the measurement, ULAs and incoherent signals are used.

The first comparison, done by [12], shows that MUSIC, as well as root-MUSIC, estimate the impinging angle accurately. Additionally, it is mentioned that MUSIC needs fewer array elements in the same environment as root-MUSIC to perform the same. For a SNR $= 100\,\mathrm{dB}$ both algorithms can detect two signals separated by $1°$.

In another comparison, the authors find out that in areas with a

SNR of $0\,\mathrm{dB}$ up to $-10\,\mathrm{dB}$ and at least 200 snapshots MUSIC is able to detect the DOA more accurately than root-MUSIC [6]. As already Wameru et al. [12] has mentioned, MUSIC needs fewer array elements than root-MUSIC to detect the angles with the same accuracy. For only 100 snapshots, a SNR $= 10\,\mathrm{dB}$ and an array size equal to 16 the simulation of the root-MUSIC algorithm exposes a better accuracy and a roundabout 31 up to 1694 times lower mean squared error than for MUSIC [6].

As the simulation results in Figure 2 show, the MUSIC and root-MUSIC algorithm are performing quite equally for widely separated sources. The simulation for impinging signals which are separated by only two degrees is seen in Figure 3. MUSIC can not detect the five targets accurately. Root-MUSIC by contrast is able to detect all five signals. It has to be said that there is no information about the SNR and the number of snapshots, so it could be possible that the number of snapshots is too low for MUSIC to perform properly. The array elements in both simulations were the same. [22]

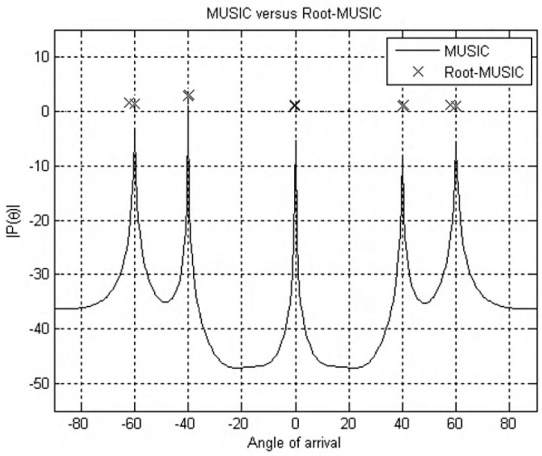

Figure 2. Comparison of the angle of arrival estimation between MUSIC and root-MUSIC. The used angles are -60, -40, 0, 40, 60. (Out of [22])

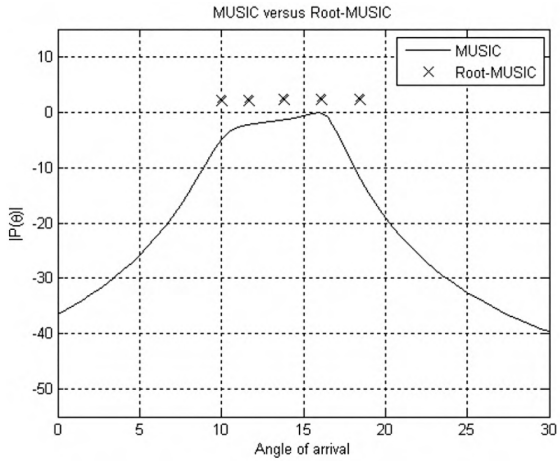

Figure 3. Comparison of the angle of arrival estimation between MUSIC and root-MUSIC. The used angles are 10, 12, 14, 16, 18. (Out of [22])

TABLE I
ACCURACY AMOUNT OF MUSIC (OUT OF [11])

Angles	DOA estimation average	Deviation from the mean	Variance of Angles
$\Theta_1 = 28°$	28.0005	-0.0005	0.0001
$\Theta_2 = 30°$	29.9985	0.0015	0.0001
$\Theta_1 = 29.2°$	29.2827	-0.0827	0.0050
$\Theta_2 = 30°$	29.9045	0.0955	0.0040
$\Theta_1 = 29.4°$	29.5067	-0.1067	0.0092
$\Theta_2 = 30°$	29.8867	0.1133	0.0069

TABLE II
ACCURACY AMOUNT OF ROOT-MUSIC (OUT OF [11])

Angles	DOA estimation average	Deviation from the mean	Variance of Angles
$\Theta_1 = 28°$	27.9998	0.0002	0.1525×10^{-3}
$\Theta_2 = 30°$	29.9997	0.0003	0.3479×10^{-3}
$\Theta_1 = 29.2°$	29.2827	-0.0827	0.005
$\Theta_2 = 30°$	29.9045	0.0955	0.004
$\Theta_1 = 29.4°$	29.4303	-0.0303	0.0047
$\Theta_2 = 30°$	29.9662	0.0338	0.0052

The last simulation gives an overview of the accuracy amount of the MUSIC and root-MUSIC algorithm, see Table I and II. For the simulation 1000 snapshots were taken, the SNR was $20\,\mathrm{dB}$ and the array consisted of 16 elements [11]. As the tables reveal root-MUSIC performs a little bit better than spatial MUSIC. However, the results in [11] correspond to the results in [6]. In [6] the authors did not measure the performance of both algorithms in an area of SNR $= 20\,\mathrm{dB}$ they only used a SNR of $10\,\mathrm{dB}$.

VI. CONCLUSION

In this paper, a brief explanation showed the functionality of the MUSIC and root-MUSIC algorithm. Furthermore, it turned out that the number of antenna elements, the number of snapshots and the height of the SNR have an impact on the resolution of the algorithm's spatial spectrum. The higher the number of antenna elements and snapshots and the higher the SNR, the better the resolution of the spectrum. Moreover, some methods carried out that it is possible to detect coherent signals too. Therefore, the algorithm has to be upgraded or the array elements have to be separated randomly. The comparison of both algorithms revealed that for a high number of snapshots the algorithms provide an accurate estimation of the DOAs of the impinging signals. Besides, for a SNR of $0\,\mathrm{dB}$ up to $-10\,\mathrm{dB}$ MUSIC performs better than root-MUSIC. When the SNR is high the accuracy of the DOA estimation is approximately the same. In addition, root-MUSIC needs more array elements in the same environment to perform the same as MUSIC.

REFERENCES

[1] L. C. Godara, *Smart Antennas*. CRC Press, 2004.

[2] J. W. Odendaal, E. Barnard, and C. W. Pistorius, "Two-dimensional super-resolution radar imaging using the music algorithm," *IEEE Transactions on Antennas and Propagation*, vol. 42, pp. 1386–1391, 1994.

[3] R. Joshi and A. Dhande, "Direction of arrival estimation using music algorithm," *International Journal of Research in Engineering and Technology*, vol. 3, pp. 633–636, 2014.

[4] P. Gupta and S. P. Kar, "Music and improved music algorithm to estimate direction of arrival," in *2015 International Conference on Communication and Signal Processing*. Institute of Electrical and Electronics Engineers Inc., 11 2015, pp. 757–761.

[5] Y. Liao and A. Abouzaid, "Resolution improvement for music and root music algorithms," *Journal of Information Hiding and Multimedia Signal Processing*, vol. 6, pp. 189–197, 2015.

[6] T. S. Dhope, D. Simunic, and M. Djurek, "Application of doa estimation algorithms in smart antenna systems," *Studies in Informatics and Control*, vol. 19, pp. 445–452, 2010.

[7] H. Zhang, Z. Gao, and H. Fu, "A high resolution random linear sonar array based music method for underwater doa estimation," in *Proceedings of the 32nd Chinese Control Conference*, 2013, pp. 4592–4595.

[8] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation*, vol. 34, pp. 276–280, 1986.

[9] R. D. DeGroat, E. M. Dowling, and D. A. Linebarger, "The constrained music problem," *IEEE Transactions on Signal Processing*, vol. 41, pp. 1445–1449, 1993.

[10] T. Gebauer and H. G. Göckler, "Channel-individual adaptive beamforming for mobile satellite communications," *IEEE Journal on Selected Areas in Communications*, vol. 13, pp. 439–448, 1995.

[11] M. Jalali, M. N. Moghaddasi, and A. Habibzadeh, "Comparing accuracy for ml, music, root-music and spatially smoothed algorithms for 2 users," in *2009 Mediterrannean Microwave Symposium (MMS)*, 2009, pp. 1–5.

[12] N. P. Wameru, B. O. Konditi, and P. K. Langat, "Performance analysis of music, root-music and esprit doa estimation algorithm," *International Journal of Electronics an Communication Engineering*, vol. 8, pp. 209–216, 2014.

[13] Y. L. Sit, C. Sturm, J. Baier, and T. Zwick, "Direction of arrival estimation using the music algorithm for a mimo ofdm radar," in *2012 IEEE National Radar Conference*, 2012, pp. 226–229.

[14] I. Jouny, "Improving music algorithm using bss," in *2007 IEEE Antennas and Propagation Society International Symposium*, 2007, pp. 5267–5270.

[15] M. M. Gunjal and A. A. B. Raj, "Improved direction of arrival estimation using modified music algorithm," in *2020 5th International Conference on Communication and Electronics Systems*, 2020, pp. 249–254.

[16] A. M. Elbir, "Deepmusic: Multiple signal classification via deep learning," *IEEE Sensors Letters*, vol. 4, pp. 1–4, 4 2020.

[17] F. Belfiori, W. van Rossum, and P. Hoogeboom, "Application of 2d music algorithm to range-azimuth fmcw radar data," in *2012 9th European Radar Conference*, 2012, pp. 242–245.

[18] A. Vesa, "Direction of arrival estimation using music and root-music algorithm," in *18th Telecommunications Forum, Pg*, vol. 18, 2010, pp. 582–585.

[19] H. K. Hwang, Z. Aliyazicioglu, M. Grice, and A. Yakovlev, "Direction of arrival estimation using a root-music algorithm," in *Proceedings of the International MultiConference of Engineers and Computer Scientists*, vol. 2, 2008, pp. 19–21.

[20] B. D. Rao and K. V. Hari, "Performance analysis of root-music," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, pp. 1939–1949, 1989.

[21] M. Rübsamen and A. B. Gershman, "Performance analysis of root-music-based direction-of-arrival estimation for arbitrary non-uniform array," in *2008 5th IEEE Sensor Array and Multichannel Signal Processing Workshop*. IEEE Computer Society, 2008, pp. 381–385.

[22] M. G. Porozantzidou, G. A. Mavridis, D. E. Anagnostou, and M. T. Chryssomallis, "Performance improvement of music algorithm by combined use of root and beam-space music versions," in *2009 IEEE Antennas and Propagation Society International Symposium*, 2009, pp. 1–4.

Wired vs. Wireless Communication: (Dis)Advantages, Applications, Trends

1st Maximilian Meier
Faculty of Electrical Engineering and Information Technology
Ostbayerische Technische Hochschule
Regensburg, Germany
maximilian.meier@st.oth-regensburg.de

***Abstract*—A communication system is an everyday item of use. Mostly it gets associated with Wi-Fi or LAN in a home network. But this is only a small part in the wide range of communication systems. Therefore, a communication system can be seen as everything that gives the possibility to exchange information between two locations. In general they can be separated into two categories: Wired and wireless systems. This paper does a general literature research on the advantages and disadvantages of wired and wireless communication and compares them in the points: data transfer rate, energy consumption, security, transmission range, stability, expandability, and installation costs. After this the paper takes a deeper look at different possible applications with communication systems. The focus therefore is on how the advantages and disadvantages discovered in the previous section affect these applications. Bringing all these things together shows, that the use of wired or wireless communication systems is generally determined by the systems operation environment and what the final system should be capable of. To close the paper a closer look is taken on future trends of wireless and wired communication and what is more reliable for future applications. This shows, that in the world of tomorrow everything must be connectable and that a lot of things like Smart Homes or the Internet of Things would not be even possible without a combination of wired and wireless communication systems.**

***Index Terms*—Wireless communication, Wireless LAN, Comparison, Applications, Future Trends, Wi-Fi, Zigbee, Ethernet, IoT, Smart Home**

I. Introduction

Communication is one of the biggest challenges in our modern world. However communication itself means to share information between two or more people and is basically done oral or via writing [1]. This principle can be transferred into the digital world: communication can be seen as an exchange of information between two or more devices. To give a device the ability to communicate with other devices they must be connected to each other. Therefore this can be realized over a wired or wireless connection [2]. By looking on communication systems there is a wide range full of different types of connections. On the one side there are many types of wireless communication systems which work over radio waves. They use protocols like Wi-Fi, Wi-Max, UWB, Bluetooth, GSM/GPRS or Zigbee to name just a few. The authors of [3] took a closer look to them and compared them. On the other side wired communication systems are similar various. There is for example the Ethernet protocol IEEE802.3 which evolved during the last decades to the standard communication protocol for computer networks worldwide [4]. But also there are variations like I2C or SPI described in [5], which are protocols, that are commonly used to connect microcontrollers.

Because this paper only does a general comparison between wired an wireless communication all these different kinds of connections will be summarized to the basic categories wired and wireless regarding to their type of connection. If there is a physical, conductive connection between the devices it is assumed as a wired communication system and if there isn't then it is categorized as a wireless communication system.

II. Advantages and Disadvantages

In this section a closer look will be made to the advantages and disadvantages of wireless and wired communication systems. As mentioned before there is not only one way to build a wired or wireless communication system. For this section a few of the widely used communication systems had been chosen to get an impression of what the communication systems are capable of. Table I and Table II sum up the results below and give some additional information.

A. Wireless Communication

Wireless communication is based on radio waves, where the data to be transferred is modulated on a carrier signal. The carrier signal therefore is dependent on the wireless communication system. A big advantage all wireless communication systems share is the nonnecessity of wires and the resulting possibility of the devices to join and leave the communication system on the fly [6]. But all kinds of radio waves have one common drawback. They interfere and therefore disturb other signals that are transmitted at the same time. This can result into a loss of data or in the worst case to a complete interruption of the wireless connection [7].

a) Bluetooth: Bluetooth was invented as a solution for short distance device to device communication. Therefore it was developed to be a cheap solution for replacing wires between two devices. For its functionality Bluetooth uses the 2.4 GHz band which is freely available in most countries. In addition to that Bluetooth is also designed to be cost effective and low in energy consumption. This allows the use of devices with low energy capacity available. So it can be used in nearly

any portable device to make it wirelessly accessible [6]. Bluetooth also offers security features to protect the transmitted data. Therefore 4 modes of security are defined which make different steps do ensure a secure data transmission between the devices. The authors of [8] give a good explanation for each of the security modes.

b) Wi-Fi: Unlike Bluetooth, Wi-Fi is designed to work as a short distance computer to computer communication system. It can therefore also be seen as a replacement for wired LAN networks but is often integrated in devices as extension to achieve more connectivity [6]. Wi-Fi is standardized in IEEE802.11 and gets since its invention continuously improved. The author of [9] gives a review on the different versions. In the newest version called Wi-Fi 6e the frequency bands $2.4\,\mathrm{GHz}$, $5\,\mathrm{GHz}$ and $6\,\mathrm{GHz}$ are used [10]. The data transfer therefore can be encrypted with the protocols WEP, WPA, WPA2 or WPA3 which had been designed to protect Wi-Fi communication. The authors of [11] give a good summary of the protocols.

c) Zigbee: Zigbee is a protocol which offers the ability to build up a stable cheap wireless communication system with low power devices with a low data rate [12]. A Zigbee network consists of 3 different types of devices: Coordinator, Router and End Device. A Coordinator can only be once in a Zigbee network and be seen as the origin of the communication system. The Coordinator manages the data transfer between different networks and acts as an access point through which a Zigbee network is accessible [12], [13]. The Router is responsible for the data exchange between the devices in the network and gives the Zigbee network the ability to expand [12]. A Zigbee End Device is a simple device that can only communicate with the Router or Coordinator. It can not route data from one device to another device. If an End Device is not used it can go in to a energy saving sleep mode and therefore extend its battery lifetime [14]. Therefore it is very hard to estimate the energy consumption of a Zigbee End Device because energy is only consumed during data transfer. The frequencies Zigbee uses vary from country to country. Worldwide the $2.45\,\mathrm{GHz}$ band is used. In America the $915\,\mathrm{MHz}$ Band is used and in Europe the $868\,\mathrm{MHz}$ Band is used [13]. To protect its data from unauthorized access Zigbee uses a 128-Bit AES encryption. The authors of [14] give a good explanation of how the encryption is applied.

B. Wired Communication

Wired communication systems afford in contrast to wireless communication systems a physical connection between the members of the network [15]. Without this connection a device is not able to enter the network. This makes every single wired communication network much safer than a wireless one, which everyone could join if they are able to receive the radio signal and knows how to bypass the security mechanisms. But on the other side this makes a wired communication system inflexible to integrate a new device and hard to scale for a flexible operating area [16].

a) I2C: For communication over the I2C bus two wires are needed. One is called the serial clock (SCL) and the other wire is called serial data (SDA). The SCL wire provides the clock for all members of the bus and the data gets transmitted via the SDA wire. The bus system itself consists of two device types, master and slave. Only a master can start a conversation to a slave and the slave only speaks when it was requested by the master [17]. Bus-systems can contain more than one master which can communicate among themselves. Every device in the bus must have its own unique 7 or 10-bit address which makes it possible to connect either 128 or 1024 devices [5]. The Data over the I2C bus is not encrypted and the transmitted signal is not a differential signal so that electromagnetic compatibility (EMC) can influence the signal and create false information.

b) CAN: The CAN-Bus is an asynchronous serial communication system, where the data to be send is transmitted over two wires as a differential signal. This makes the CAN-Bus resistant against electromagnetic influence from other devices. No clock signal is transmitted over an extra wire instead an internal oscillator is used. Different to I2C in a CAN-Bus system neither of the members has an unique address. The transmitted data is labeled with an message-ID to that the participants of the CAN-Bus can subscribe. So every device connected to the CAN-Bus receives every message and therefore has to determine if it belongs to it or not. This leads to an unlimited number of possible participants. Every member of the CAN-Bus can send data whenever it wants to. The arbitration of the participants is done via dominant and recessive bits where a logical 0 always wins over a logical 1. Therefore the participant that sends data, always has to read back the sent data to detect whether another data transmission is taking place at the same time and then abort its transmission if necessary [18]. The CAN protocol has some features to ensure the correctness of the transmitted data. But it has no type of encryption so everyone with access to the system can freely read and write data [19].

c) Ethernet: Ethernet is the most commonly used wired communication system. Based on the definition at the beginning of this paper only ethernet with copper wires is considered. For the data transfer at full speed 4 twisted pair copper wires are used. Differential signals are transmitted via the respective twisted pair wires, resulting in a very high immunity to external electromagnetic interference [28]. For addressing the devices in a network the internet protocol address version 4 (IPv4) and internet protocol address version 6 (IPv6) are used. While IPv4 is nowadays used in private networks to address all connected devices, IPv6 is used to address all devices connected to the internet. Because it is not possible to route data over ethernet from a device with a private IPv4 address to a device with a public IPv6 address a conversion from IPv4 to IPv6 has to be done. IPv4 consist of 32-bit and is therefore able to address theoretically 2^{32} devices. With IPv6 the address consists of 128-bit, which creates 2^{128} unique addresses [29]. The same addressing mechanism is also used by Wi-Fi which can be seen as the wireless equivalent

TABLE I
PROPERTIES OF THE SELECTED WIRELESS CONNECTIONS

Type	data transfer rate	energy consumption	transmission range	security	interference	expandability
Bluetooth	$\leq$2 MB/s [20], [21]	1 mA – 35 mA [6], [20]	1 m – 100 m [21]	✓	✓	✓
Wi-Fi	$\leq$10 GB/s [6]	—[a]	$\leq$100 m [21]	✓	✓	✓
Zigbee	$\leq$250 kB/s [12]	—[a]	10 m – 100 m [21]	✓	✓	✓

[a]Data was not available

TABLE II
PROPERTIES OF THE SELECTED WIRED CONNECTIONS

Type	data transfer rate	energy consumption	transmission range	security	interference	expandability
I2C	0.1 MB/s – 3.4 MB/s[a] [5]	—[b]	$\leq$10 m [22]	✗	✓	✓
CAN	$\leq$0.125 MB/s [18]	10 mA – 75 mA [23], [24]	$\leq$40 m[c] [18]	✗	✗	✓
Ethernet	10 GB/s [25]	0.5 W – 5 W [25]	$\leq$100 m [26]	✗	✓	✓

[a]5 MB/s absolute max but not bidirectional [27]
[b]Data was not available
[c]$\leq$1 km possible but bitrate decreases to 10 kbit/s

of ethernet. Therefore any device with an ethernet port can be connected to the communication system. Looking at the security features of ethernet, to join the network a physical access to it must exist. But on the other side if this access is existent everything is possible because an ethernet port is unlike Wi-Fi not encrypted. There are possibilities to secure the access to a wired LAN network like using a VLAN or IPsec as discussed in [30] or [31] but not without a certain effort.

III. APPLICATIONS OF WIRED AND WIRELESS COMMUNICATION

In the following section, the use of communication systems will be examined in various areas of application. Therefore a closer look is taken on how to realize these applications by using wired or wireless communication technologies.

A. Smart Grid

The energy turnaround is characterized by a move away from fossil fuels in energy generation. In contrast, more energy is now being generated from renewable sources. This change in energy generation is also causing a change in the current grid structure. The current centralized energy supply, with a few large energy suppliers, is becoming a decentralized energy supply with many small but widely distributed energy producers. [32]

In order to cope with the increased control effort, the current power grid must be modernized. This is what the term Smart Grid stands for. An energy network that can transport energy as loss-free and safely as possible from the producer to the consumer through intelligent regulation. [33]

To make this possible, all participating elements must be networked with each other. In this way, data transport is also made possible in addition to energy transport. Figure 1 is intended to illustrate this. This is a schematic representation of a Smart Grid. On the right in the center of Figure 1 is an

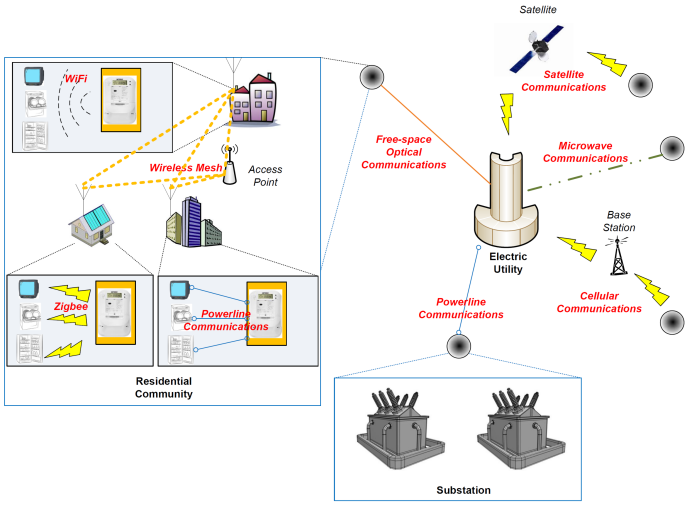

Fig. 1. Schematic structure of a Smart Grid [33]

electric utility. Starting from this, there are several connections to other grids, which are represented by black dots. Two of them are shown in more detail. One of them is a residential community and the other is a transformer substation. For communication between and within the different grids, different wired and wireless communication networks are used. For example, a Powerline connection is used to exchange data between the electric utility and the transformer substation. However, data is also exchanged with other subnetworks via Satellites, Mobile communications and Free-space optical communications. Within the residential community, data is exchanged via Wi-Fi, Zigbee or a Powerline connection.

B. Industry 4.0

The Industry 4.0 is widely seen as the next step in the industrial evolution [34]. While the previous industrial revolutions focused on introducing mass production and electrifying

industrial processes, the industry 4.0 is focused on digitalization, automation and connectivity [35]. Figure 2 shows this principle. This concept is related to the Internet of Things

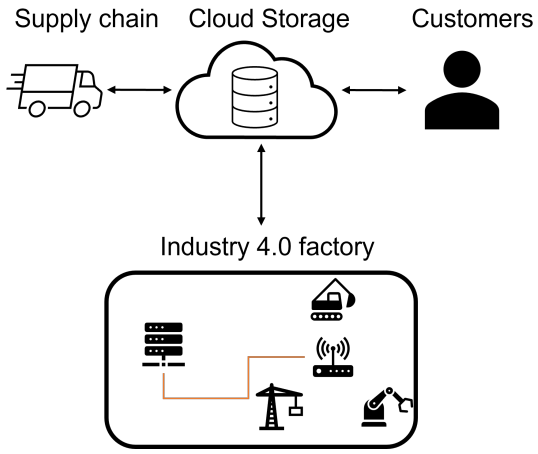

Fig. 2. Schematic structure of a smart factory [36]

(IoT) in that also every device is connectable with the internet. All the devices can be connected to each other via any wired or wireless communication system. It depends on what the device must be capable of. If the device is always on the same spot then a wired connection would be the best option due to its interference resistance. Depending on the amount of data that must be transferred either Ethernet or CAN could be a good choice. But if the device is constantly in motion or must be highly scalable, a wireless communication system would be a better option. Depending on factors like safety, transmission range, interference and data transfer rate a decision for example between Bluetooth or Wi-Fi could be made.

The concept of industry 4.0 also goes one step forward. The used machines can be cyber-physical systems which are able to work autonomously and improve their capabilities through machine learning. Figure 2 shows a network where all connected machines can take care about the product flow and organize consumables over the internet. Also customers could order products over the cloud and therefore over the internet which then will be fully automated delivered.

IV. Future trends

Figure 3 shows a forecast of the number of Machine-to-Machine communication connections from 2014 to 2022 in the mobile communication sector. This shows that an exponential growth of these compounds is to be expected. In this respect, it is to be expected that the wireless communications sector will continue to grow. However, this in turn also means that there is a general trend towards wireless communication technologies. More and more devices are being connected via radio technology. This trend can be explained mainly by the ease of integration and expandability of wireless communication systems. This increase in wireless communication also results in a possible change in the way buildings are constructed.This

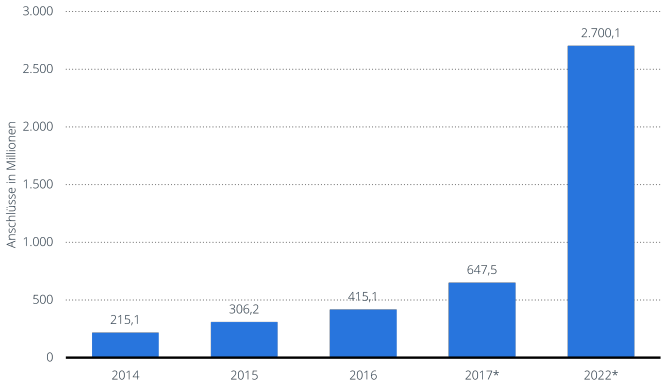

Fig. 3. Forecast of the evolution of Machine-to-Machine communication [37]

is the subject of the authors in [38]. They investigate how the attenuation of radio signals by walls, ceilings, etc. can be reduced by optimizing the building design. Despite all the advantages of wired networks, wireless networks have decisive disadvantages in terms of stability and interference. According to [39], current wireless communication systems are not capable of establishing a stable real-time connection. Real-time capability is one of the basic requirements in various applications especially for the industry 4.0 which requires a real-time dataflow [35]. This will a major topic in future to realize also wireless communication systems real-time capable in order to be able to replace the wired real-time applications.

V. Conclusion

All wireless and wired communication systems have specific benefits and drawbacks. While wireless communication systems are generally cheaper by installation and can be extended with low effort, wired communication systems are more stable and do not interfere much with their environment. Therefore wired communication systems demand a higher installation effort and every device that wants to join the network must have a physical connection to the system. This makes wired communication systems much more inflexible for expansion. But on the other side this inflexibility ends up in a higher level of security as any attacker must also have a physical connection to the system to be able to invade it. Unlike wireless networks, the attacker only needs to be within radio range to gain access. On the application side there is no absolute winner. Even when the trend clearly shows, that wireless communication systems will grow in number, in the point real-time capability wired communication systems will have an advantages that wireless communication do not have for now. Both communication systems are indispensable and cannot compensate for the disadvantages of the other. Therefore, for any good communication system, it is important to use the right communication systems according to their specifications in the right place and on the right device.

References

[1] N. E. Mastorakis, Ed., *Recent advances in financial planning and product development: Proceedings of the 5th international conference on finance, accounting and law (ICFA '14), proceedings of the 5th international conference on design and product development (ICDPD '14) ; Istanbul, Turkey, December 15-17, 2014* (Business and economics series). [Greece]: WSEAS Press, 2014, vol. 17, ISBN: 978-1-61804-261-3.

[2] V. C. Gungor, D. Sahin, T. Kocak, *et al.*, "Smart grid technologies: Communication technologies and standards," *IEEE Transactions on Industrial Informatics*, vol. 7, no. 4, pp. 529–539, 2011, ISSN: 1551-3203. DOI: 10.1109/TII.2011.2166794.

[3] S. Chakkor, E. A. Cheikh, M. Baghouri, and A. Hajraoui, "Comparative performance analysis of wireless communication protocols for intelligent sensors and their applications," 2014. DOI: 10.48550/arXiv.1409.6884.

[4] D. Law, D. Dove, J. D'Ambrosia, M. Hajduczenia, M. Laubach, and S. Carlson, "Evolution of ethernet standards in the ieee 802.3 working group," *IEEE Communications Magazine*, vol. 51, no. 8, pp. 88–96, 2013, ISSN: 0163-6804. DOI: 10.1109/MCOM.2013.6576344.

[5] F. Leens, "An introduction to i 2 c and spi protocols," *IEEE Instrumentation & Measurement Magazine*, vol. 12, no. 1, pp. 8–13, 2009, ISSN: 1094-6969. DOI: 10.1109/MIM.2009.4762946.

[6] E. Ferro and F. Potorti, "Bluetooth and wi-fi wireless protocols: A survey and a comparison," *IEEE Wireless Communications*, vol. 12, no. 1, pp. 12–26, 2005, ISSN: 1536-1284. DOI: 10.1109/MWC.2005.1404569.

[7] G. Shi, *Signal Interference in WiFi and ZigBee Networks* (Wireless Networks Ser). Cham: Springer International Publishing, 2016, ISBN: 978-3-319-47806-7. [Online]. Available: http://gbv.eblib.com/patron/FullRecord.aspx?p=4723667.

[8] R. Bouhenguel, I. Mahgoub, and M. Ilyas, "Bluetooth security in wearable computing applications," in *2008 International Symposium on High Capacity Optical Networks and Enabling Technologies*, IEEE, 2008, pp. 182–186, ISBN: 978-1-4244-2960-8. DOI: 10.1109/HONET.2008.4810232.

[9] S. Jaidev, *The wi-fi evolution*, Qorvo, Ed., 2020. [Online]. Available: https://www.qorvo.com/resources/d/qorvo-the-wi-fi-evolution-white-paper.

[10] P. Drahos, E. Kucera, O. Haffner, R. Pribis, and L. Beno, "Trends in industrial networks including apl, tsn, wifi-6e and 5g technologies," in *2022 Cybernetics & Informatics (K&I)*, IEEE, 2022, pp. 1–7, ISBN: 978-1-6654-8775-7. DOI: 10.1109/KI55792.2022.9925965.

[11] B. Indira Reddy and V. Srikanth, "Review on wireless security protocols (wep, wpa, wpa2 & wpa3)," *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, pp. 28–35, 2019. DOI: 10.32628/CSEIT1953127.

[12] N. V. R. Kumar, C. Bhuvana, and S. Anushya, "Comparison of zigbee and bluetooth wireless technologies-survey," in *2017 International Conference on Information Communication and Embedded Systems (ICICES)*, IEEE, 2017, pp. 1–4, ISBN: 978-1-5090-6135-8. DOI: 10.1109/ICICES.2017.8070716.

[13] T. Kumar and P. B. Mane, "Zigbee topology: A survey," in *2016 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT)*, IEEE, 2016, pp. 164–166, ISBN: 978-1-5090-5240-0. DOI: 10.1109/ICCICCT.2016.7987937.

[14] M. Gupta and S. Singh, "A survey on the zigbee protocol, it's security in internet of things (iot) and comparison of zigbee with bluetooth and wi-fi," in *Applications of Artificial Intelligence in Engineering*, ser. Algorithms for Intelligent Systems, X.-Z. Gao, R. Kumar, S. Srivastava, and B. P. Soni, Eds., Singapore: Springer Singapore, 2021, pp. 473–482, ISBN: 978-981-33-4603-1. DOI: 10.1007/978-981-33-4604-8{\textunderscore}38.

[15] Y. Zou, J. Zhu, X. Wang, and L. Hanzo, "A survey on wireless security: Technical challenges, recent advances, and future trends," *Proceedings of the IEEE*, vol. 104, no. 9, pp. 1727–1765, 2016, ISSN: 0018-9219. DOI: 10.1109/JPROC.2016.2558521.

[16] M. Kuzlu, M. Pipattanasomporn, and S. Rahman, "Review of communication technologies for smart homes/building applications," in *2015 IEEE Innovative Smart Grid Technologies - Asia (ISGT ASIA)*, IEEE, 2015, pp. 1–6, ISBN: 978-1-5090-1238-1. DOI: 10.1109/ISGT-Asia.2015.7437036.

[17] W. Gay, "I2c," in *Beginning STM32*, W. Gay, Ed., Berkeley, CA: Apress, 2018, pp. 195–221, ISBN: 978-1-4842-3623-9. DOI: 10.1007/978-1-4842-3624-6{\textunderscore}11.

[18] G. Schnell and B. Wiedemann, *Bussysteme in der Automatisierungs- und Prozesstechnik*. Wiesbaden: Springer Fachmedien Wiesbaden, 2019, ISBN: 978-3-658-23687-8. DOI: 10.1007/978-3-658-23688-5.

[19] M. Bozdal, M. Samie, and I. Jennions, "A survey on can bus protocol: Attacks, challenges, and potential solutions," in *2018 International Conference on Computing, Electronics & Communications Engineering (iCCECE)*, IEEE, 2018, pp. 201–205, ISBN: 978-1-5386-4904-6. DOI: 10.1109/iCCECOME.2018.8658720.

[20] Laird, *Datasheet-bl5340-series*, Laird, Ed., 2022. [Online]. Available: https://www.lairdconnect.com/documentation/datasheet-bl5340-series.

[21] P. P. Parikh, M. G. Kanabar, and T. S. Sidhu, "Opportunities and challenges of wireless communication technologies for smart grid applications," in *IEEE PES General Meeting*, IEEE, 2010, pp. 1–7, ISBN: 978-1-4244-6549-1. DOI: 10.1109/PES.2010.5589988.

[22] T. Addabbo, A. Fort, M. Mugnaini, S. Parrino, A. Pozzebon, and V. Vignoli, "Using the i2c bus to set up long range wired sensor and actuator networks in smart buildings," in *2019 4th International Conference on Computing, Communications and Security (ICCCS)*, IEEE, 2019, pp. 1–8, ISBN: 978-1-7281-0875-9. DOI: 10.1109/CCCS.2019.8888085.

[23] Microchip Technology Inc, *High-speed can transceiver*, Microchip Technology Inc, Ed., 2010. [Online]. Available: https://ww1.microchip.com/downloads/aemDocuments/documents/APID/ProductDocuments/DataSheets/20001667G.pdf.

[24] Microchip Technology Inc, *Stand-alone can controller with spi interface*, Microchip Technology Inc., Ed., 2012. [Online]. Available: https://ww1.microchip.com/downloads/aemDocuments/documents/APID/ProductDocuments/DataSheets/MCP2515-Family-Data-Sheet-DS20001801K.pdf.

[25] K. Christensen, P. Reviriego, B. Nordman, M. Bennett, M. Mostowfi, and J. Maestro, "Ieee 802.3az: The road to energy efficient ethernet," *IEEE Communications Magazine*, vol. 48, no. 11, pp. 50–56, 2010, ISSN: 0163-6804. DOI: 10.1109/MCOM.2010.5621967.

[26] A. B. Semenov, "Limit length of long ethernet telecommunication lines for industrial information systems with centralized structure," *Journal of Physics: Conference Series*, vol. 1582, no. 1, p. 012 075, 2020, ISSN: 1742-6588. DOI: 10.1088/1742-6596/1582/1/012075.

[27] Joseph Wu, *Basics of i2c: The i2c protocol*, Texas Instruments, Ed., 2021. [Online]. Available: https://training.ti.com/sites/default/files/docs/basics-of-i2c-the-i2c-protocol-presentation.pdf.

[28] W. Riggert, *Rechnernetze: Grundlagen - Ethernet - Internet* (Lehrbücher zur Informatik), 5., aktualisierte Auflage. München: Hanser, 2014, ISBN: 978-3-446-44096-8.

[29] M. Sailan, R. Hassan, and A. Patel, "A comparative review of ipv4 and ipv6 for research test bed," in *2009 International Conference on Electrical Engineering and Informatics*, IEEE, 2009, pp. 427–433, ISBN: 978-1-4244-4913-2. DOI: 10.1109/ICEEI.2009.5254698.

[30] J. Sommer, S. Gunreben, F. Feller, *et al.*, "Ethernet – a survey on its fields of application," *IEEE Communications Surveys & Tutorials*, vol. 12, no. 2, pp. 263–284, 2010, ISSN: 1553-877X. DOI: 10.1109/SURV.2010.021110.00086.

[31] T. Kiravuo, M. Sarela, and J. Manner, "A survey of ethernet lan security," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 3, pp. 1477–1491, 2013, ISSN: 1553-877X. DOI: 10.1109/SURV.2012.121112.00190.

[32] L. Gailing and A. Röhring, "Was ist dezentral an der energiewende? infrastrukturen erneuerbarer energien als herausforderungen und chancen für ländliche räume," *Raumforschung und Raumordnung — Spatial Research and Planning*, vol. 73, no. 1, pp. 31–43, 2015, ISSN: 0034-0111. DOI: 10.1007/s13147-014-0322-7.

[33] X. Fang, S. Misra, G. Xue, and D. Yang, "Smart grid — the new and improved power grid: A survey," *IEEE Communications Surveys & Tutorials*, vol. 14, no. 4, pp. 944–980, 2012, ISSN: 1553-877X. DOI: 10.1109/SURV.2011.101911.00087.

[34] A. C. Pereira and F. Romero, "A review of the meanings and the implications of the industry 4.0 concept," *Procedia Manufacturing*, vol. 13, pp. 1206–1214, 2017, ISSN: 23519789. DOI: 10.1016/j.promfg.2017.09.032.

[35] A. Rojko, "Industry 4.0 concept: Background and overview," *International Journal of Interactive Mobile Technologies (iJIM)*, vol. 11, no. 5, p. 77, 2017. DOI: 10.3991/ijim.v11i5.7072.

[36] A. Taivalsaari and T. Mikkonen, "On the development of iot systems," in *2018 Third International Conference on Fog and Mobile Edge Computing (FMEC)*, IEEE, 2018, pp. 13–19, ISBN: 978-1-5386-5896-3. DOI: 10.1109/FMEC.2018.8364039.

[37] F. Tenzer, *Prognose zur anzahl der machine-to-machine-anschlüsse (m2m) im mobilfunk weltweit von 2014 bis 2022*, Statista, Ed., 2022. [Online]. Available: https://de.statista.com/statistik/daten/studie/413052/umfrage/prognose-zur-anzahl-der-m2m-anschluesse-im-mobilfunk-weltweit/ (visited on 11/29/2022).

[38] S. Suherman, "Wifi-friendly building to enable wifi signal indoor," *Bulletin of Electrical Engineering and Informatics*, vol. 7, no. 2, pp. 264–271, 2018, ISSN: 2089-3191. DOI: 10.11591/eei.v7i2.871.

[39] D. Bankov, E. Khorov, A. Lyakhov, and M. Sandal, "Enabling real-time applications in wi-fi networks," *International Journal of Distributed Sensor Networks*, vol. 15, no. 5, p. 155 014 771 984 531, 2019, ISSN: 1550-1477. DOI: 10.1177/1550147719845312.

Technology Review of Battery Energy Storage Systems for Home Applications

Florian Geitner
Faculty of Electrical Engineering and Information Technology
Ostbayerische Technische Hochschule Regensburg
Regensburg, Germany
https://orcid.org/0000-0001-6177-9431

***Abstract*—To counteract climate change and simultaneously become more self-sufficient, more and more private households are generating electricity from renewable energy sources (RESs). Those mostly include wind and solar power. The problem with RESs is that they can rarely provide an immediate response to demand. If feeding excess produced energy into the public grid is not profitable, it makes sense for homeowners to install a battery energy storage system (BESS). BESSs in conjunction with solar power generation at home is often referred to as a solar home system (SHS). Within this paper, electrochemical storage technologies used in stationary battery systems will be compared based on the literature. The energy storage technologies discussed include those that currently have the highest market share as well as promising new technologies. The battery types with the highest market share in Germany are lithium-ion batteries (LIBs) and lead acid (LA) batteries. Promising new technologies include redox flow batteries (RFBs) and sodium ion batteries (SIBs) ("saltwater batteries"). Key quantitative properties that are used to compare the battery types in this paper include cyclability, energy density, storage efficiency, self-discharge rate, and cost. Qualitative parameters taken into account include the maturity of the respective technology, environmental aspects, and operational constraints (especially related to safety concerns, e.g. fire hazards, leaks, etc.).**

***Index Terms*—energy storage, batteries, energy management, renewable energy sources**

I. Introduction

In 2021, solar power was the second largest gross renewable energy source in Germany behind onshore wind power [1], [2]. With 51.2 TWh of generated electrical energy, nearly 9% of Germany's gross electricity consumption in 2021 was covered by solar power [1], [2]. Since 2014, German solar power generation has increased by more than 160%, but it has to grow even more in the coming years [1]–[3]. According to [3], solar and wind power are expected to provide more than 50% of overall power generation in the G20 countries by 2030.

Private individuals are also contributing to the energy transition. More than 1.4 million private homes in Germany already have a solar power plant installed [4]. This corresponds to about 3.6% of all private homes in Germany [4]. Due to planned new legislation, it will probably soon be mandatory for all new German buildings to install a solar system on their roofs [5]. If the new law is enacted, this number is expected to grow rapidly. But the power of solar and wind power is usually not available on demand and can also be very variable in quantity, which puts stress on the grid if unregulated [6], [7]. This behavior is illustrated in Fig. 1.

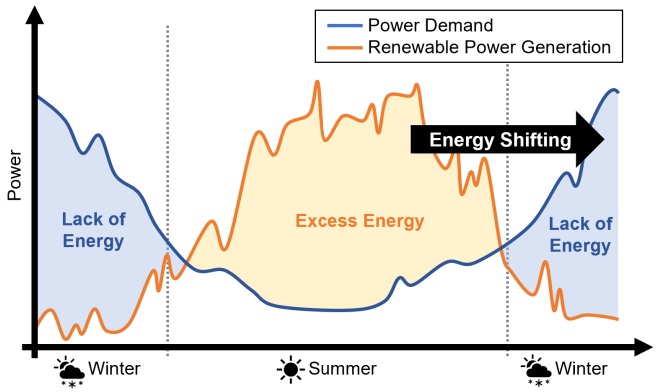

Fig. 1. Exaggerated representation of the seasonal correlation between solar power generation and power demand in Europe. Y-axis shows the power versus the time of year. [6]

A promising solution are energy storage systems (ESSs). ESSs are energy shifters, they store excess electrical energy when available and release the stored energy when the electrical power demand supersedes electrical power generation. This stabilizes the grid and ensures its reliability. ESSs are therefore indispensable in an energy infrastructure that is driven by renewable energy sources (RESs). [6]–[11]

A subset of energy storage systems are electrochemical storage systems or battery energy storage systems (BESSs). These are attractive for private households due to their often modular structure and relatively simple design. Furthermore, they can be installed anywhere without geological and geographical restrictions [12].

However, comparisons of different electrochemical energy storage systems in the end-user market are rather uncommon. This paper is intended to assist consumers in finding the appropriate technology for their solar home system (SHS). For this purpose, four different storage technologies are compared based on a literature review.

The four storage technologies being compared are:

a) Lithium Ion: Lithium-ion batteries (LIBs) have the biggest market penetration with a share of more than 90% of the German BESS market [13]–[15].

b) Lead Acid: Lead acid (LA) batteries are the second most used BESS technology in the German market (around 5% market share) [13].

c) Redox Flow: Redox flow batteries (RFBs) are considered the most realistic future of BESSs. This is because power and storage capacity can be scaled independently of each other. This makes the battery interesting from an economic and safety point of view. [12]

d) Sodium Ion ("Salt Water Batteries"): Sodium-ion batteries (SIBs) are supposed to be a future alternative to LIBs. The advantage of the SIB is the raw materials, which are less toxic and more abundant on earth and therefore cheaper. [16]

In addition, the basic operation of BESSs and their effects on a SHS will be described. SHSs in this context are solar power generation at home combined with a BESS.

II. Functionality and Impact of Battery Energy Storage Systems

A. Functionality of Traditional Electrochemical Cells

Traditional battery cells all share a similar structure, which is shown in Fig. 2. Electrochemical secondary cells can be recharged after being discharged. A battery consists of one or more cells. The two electrodes are made of different metals and are contained in an electrolyte. The electrolyte can either be liquid or solid. A separator prevents physical contact and thus a short circuit between the electrodes. This is especially important in densely packed cells. Both the separator and the electrolyte must allow ion flow and prevent electron flow. [9], [10]

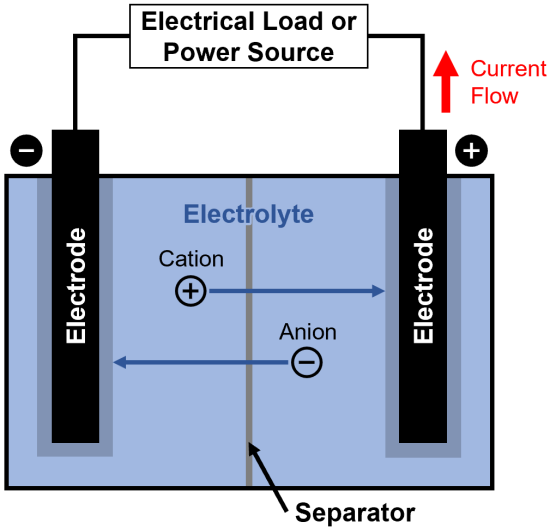

Fig. 2. Functional diagram of a traditional battery cell. The conventional current direction is used for the shown current flow arrow (not to be confused with the electron flow direction). The current flows in this direction assuming that the battery is charged and an electrical load is connected between the terminals. [9]

The voltage difference between the two terminals corresponds to the difference between the two metals in the electrochemical voltage series. The more noble a metal is, the higher it is in the electrochemical voltage series. The electrode made of nobler metal is the positive electrode. The electrode made of less noble metal is the negative electrode. [9]

When a load is connected between the electrical terminals (i.e. the battery is discharged), the negative electrode repels electrons (chemical oxidation) due to the electrical potential difference. The electrons flow towards the positive electrode and are absorbed (chemical reduction). As a result, a charge imbalance occurs. This is compensated by the electrolyte solution. Positive ions (cations) flow in the direction of the positive electrode to compensate for the excess of negative charges. Likewise, negative ions (anions) flow in the direction of the negative electrode. This process is reversed, when an electrical source is connected between the terminals (i.e. the battery is charged). [9], [10]

The electrodes determine the energy capacity and are responsible for energy conversion. The energy and power of a battery cell are therefore not independently adjustable. [9]

The operation principle of LIBs, LA batteries, and SIBs can be described with the model of the traditional electrochemical cell.

B. Functionality of Redox Flow Batteries

The basic structure of a RFB is shown in Fig. 3. The energy is stored in two electrolytes circulating in separate circuits. The electrolyte reservoir is contained in two tanks. When the electrolyte is pumped through the cell stack, a membrane prevents mixing but allows ion exchange. One electrolyte is reduced (electron gain), and the other is oxidized (electron loss). The charge imbalance of the electrolytes is rebalanced by ion exchange, and the remaining free electrons are collected by the electrode. [10]

RFBs are tertiary electrochemical elements, so they could supply power continuously (without recharging), provided the depleted electrolyte of both tanks is replaced instead of recharged. [9]

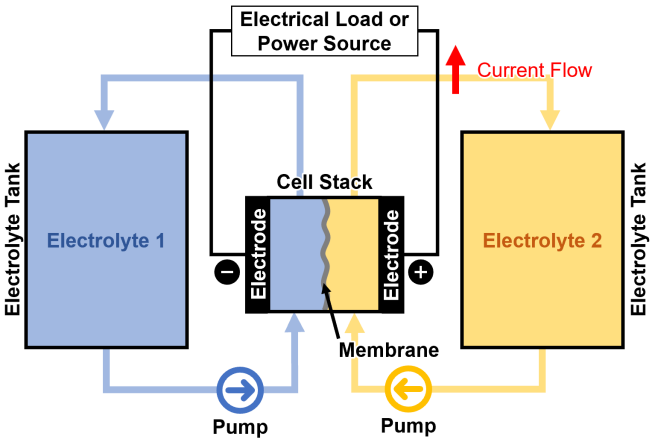

Fig. 3. Functional diagram of a RFB. The current flows in this direction assuming that the electrolytes are charged and an electrical load is connected between the two electrodes. [10]

The amount of energy that can be stored is directly proportional to the size of the electrolyte tanks. The generated power

TABLE I
COMPARISON MATRIX OF THE FOUR EVALUATED ELECTROCHEMICAL ENERGY STORAGE CELLS

	Lithium Ion Battery LFP [3], [14], [17]–[19]	**Lead Acid Battery** Flooded LA Battery [3], [17], [19], [20]	**Redox Flow Battery** Organic RFB (projected) [12], [21]	**Sodium Ion Battery** "Salt Water Batteries" [18], [22], [23]
Cyclability [#Cycles]	1,000–10,000	200–2,500	potentially unlimited	500–4,000
Volumetric Energy Density $[\frac{\mathrm{Wh}}{\mathrm{dm}^3}]$	200–600	50–100	210	180
Gravimetric Energy Density $[\frac{\mathrm{Wh}}{\mathrm{kg}}]$	90–120	30–50	150	100–130
Nominal Cell Voltage [V]	3.60	2.40	1.50	1.0–4.2
Maximum Discharge Rate [C]	5	> 30	Cell Stack specific	30
Storage Efficiency [%]	> 90	80	75–90	> 90
Self-Discharge Rate $[\frac{\%}{\mathrm{M}}]$	< 1	5	0	≥ 10[a]
Depth of Discharge [%]	> 90	50–60	100	100
Deep Cycle Resistant	✓	✗	✓	✓
Maturity	in use since 1999	in use since late 1800s	planned start of serial production in 2023[b]	in development[c]
Cost $[\frac{€}{\mathrm{kWh}}]$	750–1,150	250–500	*n.a.*	$< 1{,}150$[d]
Market Share[e] [%]	92.3	5.4	0.5	0.008
Environmental Aspects	• usage of rare earths	• toxic	• usage of abundant materials • not explosive • not flammable	• usage of abundant materials • not flammable • not toxic
Advantages	• mature technology • cathode is less flammable than that of other LIBs • partly recyclable	• mature technology • recyclable	• power and storage capacity can be configured separately • recyclable	• projected to be cheaper than LIBs • recyclable
Disadvantages	• protection circuit required to cope with thermal instability issues • expected price increases due to projected depletion of lithium before 2100 [24]	• frequent maintenance required	• inorganic counterparts still more advanced in many aspects	• not widely available yet • lower volumetric energy density than LIBs

n.a. if no data was available.
[a]LIBs have a "low" self-discharge rate while SIBs have a "low to medium" one, according to [18].
[b]As announced by the company CMBlu ([21]) in [25].
[c]Some SIBs are (or were) already commercially available but in the prototypic stage. See [26], [27].
[d]It is uncertain if SIBs will be cheaper than LIBs due to lower material cost but also lower volumetric energy density [18].
[e]German market data from 2017, acquired from [13]. Market share of the respective parent category (i.e. for LFP the market share of LIBs is used).

can be changed via the size and number of the cell stacks. The RFB is the only electrochemical storage element where power and energy are independently configurable. [9], [10]

C. Impact on Households Equipped with Battery Energy Storage Systems

The patterns of electrical power demand and renewable electrical power generation often do not match. This is why homeowners often do not use their own generated energy and are forced to instead purchase it from the public power grid. However, this is usually not ideal, as the feed-in tariff is lower than the consumption price (at least in Germany [19]). In this case, it makes sense to install a BESS to increase self-consumption and therefore reduce expenditure on electricity. [28]

A British study found, that a SHS with a properly designed BESS capacity can save up to 88% of the annual electricity bill. In addition, a distributed home BESS infrastructure that can be coordinated by electric suppliers could provide lower pricing for all consumers. This is because at peak demand times, the excess stored renewable energy can be drawn from SHSs. [7]

However, the maximum savings potential depends on many factors, including how well electric suppliers expand central BESSs. Higher values here lead to lower (but still profitable) savings potential through SHS. [7]

III. Comparison of the Battery Energy Storage Systems

A. Characteristic Parameters used for the Comparison

Batteries are characterized by numerous parameters. In this paper, the technologies are compared in Table I based on the following characteristics. The definitions were taken from [9], [10], [28] and slightly adapted in some cases.

a) Cyclability: Describes how many times a battery can be fully charged and discharged until it only has 80% of its original energy storage capacity left. The unit of measurement is $\#\mathrm{Cycles}$. The cyclability of a battery is not a variable to be considered alone and is often related to the maximum allowed depth of discharge (DOD). In most cases, higher DOD values are related to lower cycle capability.

b) Depth of Discharge: Describes how much discharge of the battery is tolerated. For example, if a battery is specified for a DOD value of 80%, the battery can be discharged to 20% residual charge capacity. For SHS high DOD values (near 100%) are critical.

c) Deep Cycle Tolerance: Describes, whether electrochemical storage can handle a DOD of 100% with little to no damage to the overall battery performance.

d) Volumetric Energy Density: Describes the maximum amount of energy a battery can theoretically store within a given space. The unit of measurement is watt-hours per cubic decimeter ($\frac{\mathrm{Wh}}{\mathrm{dm}^3}$).

e) Gravimetric Energy Density: Sometimes also called the specific energy. It describes the maximum amount of energy a battery can theoretically store within a given mass. The unit of measurement is watt-hours per kilogram ($\frac{\mathrm{Wh}}{\mathrm{kg}}$).

f) Storage Efficiency: Ratio of the total energy that can be taken from the battery divided by the total energy stored in the battery. It is measured in percent ($\%$). The self-discharge rate is not taken into account here.

g) Self-Discharge Rate: Percentage of stored energy that a battery loses during idle operation. Measured in percent lost per month ($\frac{\%}{\mathrm{M}}$).

h) Maximum Discharge Rate: Sometimes also called C-rate. Describes the maximum output current of a battery relative to its capacity. If a battery for example has a capacity of $10\,\mathrm{Ah}$, a maximum discharge rate of $1\,\mathrm{C}$ means that the peak continuous output current is $10\,\mathrm{A}$.

i) Nominal Cell Voltage: Typical voltage between the terminals of a battery cell. Measured in volts (V).

j) Environmental Aspects: Environmental aspects described in this paper include whether rare earth or abundantly available materials are used. It also considers whether used materials are polluting and/or toxic.

k) Cost: The cost of BESS technology per kilowatt hour ($\frac{€}{\mathrm{kWh}}$). If the respective source uses a currency other than Euro, the value was converted at the current exchange rate (November 2022).

l) Maturity: How long the technology has been in use (if it is even already commercially available yet). Allows qualitative conclusions about the maturity of the technology and its reliability.

B. Battery Types covered in this Work

1) Lithium Iron Phosphate Technology: The term "LIB" is merely an umbrella term that reflects the basic principle of the technology. The LIB type with the highest market share in the stationary market is the lithium iron phosphate (LFP) battery [13]. The term LFP characterizes the positive electrode of the battery, which is made of $LiFePO_4$. The negative electrode of most commercially available LFP batteries is made out of graphite. [29]

2) Flooded Lead Acid Battery: The flooded LA battery has hardly been changed since its invention. To this day, in many commercially available LA batteries, the positive electrode consists of PbO_2 and the negative electrode consists of pure Pb. The electrolyte is sulfuric acid. [10]

3) Organic Redox Flow Battery: Organic (carbon-based) redox couple electrolytes are used for energy storage. They can be inexpensive and are abundant, unlike inorganic redox couples. The RFBs that are currently already commercially successful are based on vanadium. However, vanadium is an inorganic material and will not be discussed further. [12]

4) Sodium Ion Battery: "SIB" is an umbrella term. In contrast to LIBs, sodium ions are responsible for the ion exchange, not lithium ions. SIBs are supposed to replace LIBs in some sections. [24]

IV. Conclusion

Lithium-ion batteries (LIBs) are unlikely to become the future of stationary energy storage. Safety concerns are too great for that, potential consumers may not appreciate putting a potential fire hazard in their own homes. In addition, mineable deposits of lithium are limited and would run out before 2100 (according to current estimates). However, they will likely continue to dominate the mobile market due to their high volumetric energy density.

Lead acid (LA) storage is also likely to be of little interest for solar home system (SHS) due to its toxicity, lower depth of discharge (DOD) value, and subpar cyclability and storage efficiency.

Organic redox flow battery (RFB) and sodium-ion battery (SIB) are likely to become the dominant battery energy storage system (BESS) technologies. Acceptable energy densities (both gravimetric and volumetric), cyclability values comparable to those of LIBs, low self-discharge rates, and high DOD values make them ideal for SHSs. But their greatest advantage lies in the materials needed for them. Carbon (needed for organic RFBs) as well as sodium (needed for SIBs) can be found in abundance everywhere on earth. Thus, the need for battery energy storage can be met in a relatively environmentally friendly manner without having to continue to exploit the environment.

References

[1] Statista, *Bruttostromerzeugung aus Erneuerbaren Energien in Deutschland nach Energieträger im Jahresvergleich 2011 und 2021*. [Online]. Available: de . statista . com / statistik / daten / studie / 37612 / umfrage / stromerzeugung - durch - erneuerbare - energie - in - 2008/ (visited on 11/24/2022).

[2] BDEW, *Stromerzeugung und -verbrauch in Deutschland*, Nov. 2022. [Online]. Available: https://www.bdew.de/energie/stromerzeugung-und-verbrauch-deutschland/ (visited on 11/24/2022).

[3] *Electricity Storage and Renewables: Costs and Markets to 2030*. International Renewable Energy Agency (IRENA), Oct. 2017, ISBN: 978-92-9260-038-9. [Online]. Available: https://www.irena.org/publications/2017/oct/electricity-storage-and-renewables-costs-and-markets (visited on 11/24/2022).

[4] Statistisches Bundesamt, *2,2 Millionen Photovoltaik-Anlagen in Deutschland installiert*. [Online]. Available: https://www.destatis.de/DE/Presse/Pressemitteilungen/%202022/06/PD22_N037_43.html (visited on 11/23/2022).

[5] Fraktion BÜNDNIS 90/DIE GRÜNEN, *Entwurf eines Gesetzes zur Beschleunigung des Ausbaus von Solaranlagen zur Stromerzeugung auf Gebäuden (Solaranlagenausbaubeschleunigungsgesetz – SolarBeschlG)*, Aug. 2021.

[6] "Energy Storage Targets 2030 and 2050," European Association for Storage of Energy (EASE), Tech. Rep., Jun. 2022. [Online]. Available: https://ease-storage.eu/publication/energy-storage-targets-2030-and-2050/ (visited on 11/19/2022).

[7] B. Zakeri, G. C. Gissey, P. E. Dodds, and D. Subkhankulova, "Centralized vs. distributed energy storage – Benefits for residential users," *Energy*, vol. 236, p. 121 443, Dec. 2021, ISSN: 0360-5442. DOI: 10.1016/j.energy.2021.121443.

[8] H. E. O. Farias and Luciane Neves Canha, "Battery Energy Storage Systems (BESS) Overview of Key Market Technologies," in *2018 IEEE PES Transmission & Distribution Conference and Exhibition - Latin America (T&D-LA)*, Lima: IEEE, Sep. 2018, ISBN: 978-1-5386-5844-4. DOI: 10.1109/TDC-LA.2018.8511796.

[9] M. Doppelbauer, "Energiespeicher," in *Grundlagen der Elektromobilität: Technik, Praxis, Energie und Umwelt*, M. Doppelbauer, Ed., Wiesbaden: Springer Fachmedien, 2020, ISBN: 978-3-658-29730-5. DOI: 10.1007/978-3-658-29730-5_6.

[10] M. Sterner and I. Stadler, *Energiespeicher - Bedarf, Technologien, Integration*. Berlin, Heidelberg: Springer, 2014, ISBN: 978-3-642-37379-4. DOI: 10.1007/978-3-642-37380-0.

[11] REN21, Ed., *Renewables 2022 Global Status Report*. Paris, 2022, ISBN: 978-3-948393-04-5.

[12] P. Leung, A. A. Shah, L. Sanz, *et al.*, "Recent developments in organic redox flow batteries: A critical review," *Journal of Power Sources*, vol. 360, pp. 243–283, Aug. 2017, ISSN: 0378-7753. DOI: 10.1016/j.jpowsour.2017.05.057.

[13] Sandia National Laboratories, *DOE Global Energy Storage Database*. [Online]. Available: https://sandia.gov/ess-ssl/gesdb/public/statistics.html (visited on 11/19/2022).

[14] J. Figgener, P. Stenzel, Kairies, Kai-Phillipp, and Linßen, Jochen, "The development of stationary battery storage systems in Germany - A market review," *Journal of Energy Storage*, vol. 29, DOI: 10.1016/j.est.2019.101153.

[15] J. Figgener, P. Stenzel, K.-P. Kairies, *et al.*, "The development of stationary battery storage systems in Germany – status 2020," *Journal of Energy Storage*, vol. 33, p. 101 982, Jan. 2021, ISSN: 2352-152X. DOI: 10.1016/j.est.2020.101982.

[16] C. Vaalma, D. Buchholz, M. Weil, and S. Passerini, "A cost and resource analysis of sodium-ion batteries," *Nature Reviews Materials*, vol. 3, no. 4, p. 18 013, Apr. 2018, ISSN: 2058-8437. DOI: 10.1038/natrevmats.2018.13.

[17] B. Diouf and R. Pode, "Potential of lithium-ion batteries in renewable energy," *Renewable Energy*, vol. 76, pp. 375–380, Apr. 2015, ISSN: 0960-1481. DOI: 10.1016/j.renene.2014.11.058.

[18] F. Duffner, N. Kronemeyer, J. Tübke, J. Leker, M. Winter, and R. Schmuch, "Post-lithium-ion battery cell production and its compatibility with lithium-ion cell production infrastructure," *Nature Energy*, vol. 6, no. 2, pp. 123–134, Feb. 2021, ISSN: 2058-7546. DOI: 10.1038/s41560-020-00748-8.

[19] H. C. Hesse, R. Martins, P. Musilek, M. Naumann, C. N. Truong, and A. Jossen, "Economic Optimization of Component Sizing for Residential Battery Storage Systems," *Energies*, vol. 10, no. 7, p. 835, Jul. 2017, ISSN: 1996-1073. DOI: 10.3390/en10070835.

[20] K. Divya and J. Østergaard, "Battery energy storage technology for power systems—An overview," *Electric Power Systems Research*, vol. 79, no. 4, pp. 511–520, Apr. 2009, ISSN: 03787796. DOI: 10.1016/j.epsr.2008.09.017.

[21] *Organische Solid-Flow Energiespeicher — CMBlu Energy AG*. [Online]. Available: https://www.cmblu.com/de/home/ (visited on 11/27/2022).

[22] A. M. Skundin, T. L. Kulova, and A. B. Yaroslavtsev, "Sodium-Ion Batteries (a Review)," *Russian Journal of Electrochemistry*, vol. 54, no. 2, pp. 113–152, Feb. 2018, ISSN: 1608-3342. DOI: 10.1134/S1023193518020076.

[23] X. Pu, H. Wang, D. Zhao, *et al.*, "Recent Progress in Rechargeable Sodium-Ion Batteries: Toward High-Power Applications," *Small*, vol. 15, no. 32, p. 1 805 427, 2019, ISSN: 1613-6829. DOI: 10.1002/smll.201805427.

[24] J.-Y. Hwang, S.-T. Myung, and Y.-K. Sun, "Sodium-ion batteries: Present and future," *Chemical Society Reviews*, vol. 46, no. 12, pp. 3529–3614, 2017, ISSN: 0306-0012, 1460-4744. DOI: 10.1039/C6CS00776G.

[25] *quer mit Christoph Süß: Fränkische Wunderbatterie bald in den USA? — ARD Mediathek*. [Online]. Available: https://www.ardmediathek.de/video/quer-mit-christoph-suess/fraenkische-wunderbatterie-bald-in-den-usa/br-fernsehen/Y3JpZDovL2JyLmRlL3ZpZGVvL2I5NGFlZmQ2LTE1NTMtNDAzMS04YTU2LTllMTBlN2RmNmEwZA (visited on 11/28/2022).

[26] *Aquion Energy - Energy storage, clean and simple.* [Online]. Available: https://www.aquionenergy.com/ (visited on 11/28/2022).

[27] *Salzwasser-Akku*. [Online]. Available: https://www.usp-austria.eu/produkte/salzwasser-akku/ (visited on 11/28/2022).

[28] H. Ibrahim, A. Ilinca, and J. Perron, "Energy storage systems—Characteristics and comparisons," *Renewable and Sustainable Energy Reviews*, vol. 12, no. 5, pp. 1221–1250, Jun. 2008, ISSN: 13640321. DOI: 10.1016/j.rser.2007.01.023.

[29] G. Zubi, R. Dufo-López, M. Carvalho, and G. Pasaoglu, "The lithium-ion battery: State of the art and future perspectives," *Renewable and Sustainable Energy Reviews*, vol. 89, pp. 292–308, Jun. 2018, ISSN: 1364-0321. DOI: 10.1016/j.rser.2018.03.002.

Concepts for the integration of hydrogen into the energy supply system

Jakob Bindl
Faculty of Electrical Engineering and Information Technology
OTH Regensburg
Regensburg, Germany
jakob.bindl@st.oth-regensburg.de

***Abstract*—With the rising share of renewable power sources, most notably wind and solar, more and more countries are looking for ways to store surplus energy to stabilize the power grid in times of insufficient power production. In most regions sufficient energy storage systems that can account for hourly or even seasonal deviations in power production, must be implemented to achieve fully renewable power production. Due to its high energy density, hydrogen is often viewed as a potential storage technology. The objective of this article is to describe and analyze different application scenarios of hydrogen energy storage systems that are currently being discussed. Furthermore, the benefit these technologies could have on the transition to renewable energy sources is highlighted. The most notable of these concepts include the integration of hydrogen in microgrids(MG), wind-to-hydrogen projects and the repurposing of the existing gas infrastructure. The results of this analysis show that hydrogen is a technology that can benefit the move to renewable power generation. It differentiates itself from other storage technologies such as batteries and pumped-storage hydroelectricity with its unique properties and is mainly fitted for long-term storage. There are however still some technical challenges that need to be solved to utilize hydrogen storage systems on a broader scale. These are most notably the energy demand for storing and transporting hydrogen and the replacement of components in the natural gas infrastructure. It also needs to be further researched which components in the existing infrastructure can handle pure hydrogen.**

***Index Terms*—Hydrogen, Microgrids, Renewable energy sources, Energy storage, Fuel cells, Electrolyzer**

I. Introduction

The main technical challenge when expanding the production of renewable power sources is the growing demand for the operating reserve [1]. Shown in Fig.1 is the power production of Germany in the timespan from the 8th of October to the 11th of October. The deviations of solar power production(yellow) and wind power production(blue) can be seen. The existing fossil power plants vary their power production according to the current grid load to satisfy the overall power demand. Not only daily but also seasonal changes in power production have to be accounted for when expanding renewable power production. With the phase-out of fossil fuel power plants, the demand for storage technologies arises. Where batteries and pumped-storage hydroelectricity are adequate solutions for the short-term storage of energy, hydrogen seems to be a storage technology that can account for the seasonal deviations in power production [2]. This literature research explains different concepts for the integration of hydrogen into the energy supply system. It is split into four chapters that focus on different application scenarios for hydrogen. The first two aspects in Section II. concern the integration of hydrogen into a direct current (DC) Microgrid (MG). With the rising share of DC technologies such as battery-electric-vehicles, photovoltaics, and home-storage batteries the concept of MGs that are based on DC become more popular in the scientific field. Higher efficiency and more independence from the power grid (PG) are the main advantages of this technology. As hydrogen electrolyzers (HE) and fuel cells are also DC-based appliances, further concepts of MGs that include hydrogen will be analyzed [3]. The other subsection focuses on the application of hydrogen in off-grid wind installations. A significant part of the construction cost of wind parks is related to the connection to the PG, this is especially true for offshore installations. In this part, the concept of direct hydrogen production is presented, as it offers the possibility to reduce the construction cost and time for the electric connection [4]. Section III. focuses on the storage and transportation of hydrogen on a large scale. As hydrogen is the smallest of all elements and therefore relatively volatile, the requirements for storage facilities are high. Established ways of storing and transporting hydrogen will be examined and compared to different concepts. The possibility of repurposing the existing natural gas infrastructure will be presented and analyzed.

II. DC Microgrids

MGs are local electric grids, that can operate in an island-mode or a grid-connected mode. The benefit of this concept is the local production and consumption of electrical energy, which means it is independent from the central PG. This decentral approach reduces the strain on the PG and gives further security in the case of power outages [6]. In the concept of a DC MG the power sources and consumers are connected directly or via a DC/DC converter to a DC bus. As wind power turbines, photovoltaics, batteries, battery-electric-vehicles, HEs and fuel cells are operating on DC, these machines can be connected over the DC bus. In this way, the conversion losses from alternating current (AC) to DC and vice versa can be avoided. Furthermore, the number of rectifiers and inverters can be reduced, which reduces the overall cost [7]. An example of a DC MG in a residential application is shown

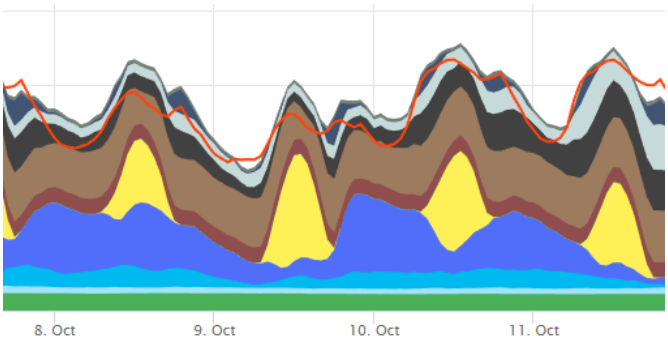

Fig. 1. Electricity generation and production in Germany [5]

Green	Biomass	Light Brown	Lignite
Blue	Wind power	Black	Hard coal
Yellow	Solar power	Grey	Natural gas
Dark Brown	Nuclear	Red	Grid load

in Fig. 2. In this scenario, the grid is separated into three different sectors. The 230V AC side is directly connected to the power grid and supplies household appliances such as a washing machine. The second sector operates on 380V DC and consists of a photovoltaics system, a battery storage, and a battery-electric vehicle. The third sector operates on low-voltage DC and supplies small DC-based appliances, such as a smartphone. In this concept, it is possible to directly supply the home battery storage or the battery-electric-vehicle with power from the photovoltaics system. The central AC/DC converter is only used to exchange energy in case of insufficient power production in one sector. In this way, the conversion losses can be reduced. In this chapter different concepts for the integration of hydrogen in DC MGs will be shown.

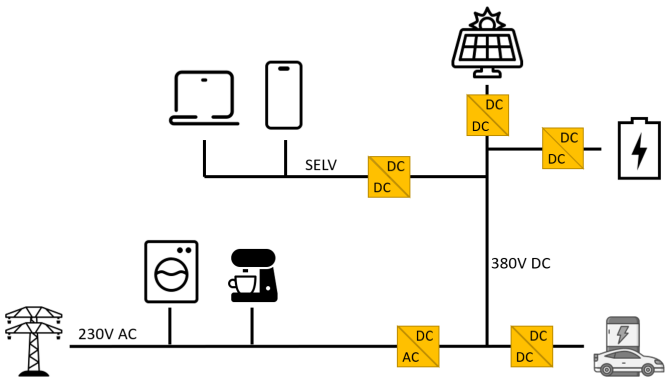

Fig. 2. DC MG for a residential application

A. Hydrogen MGs

One use case for hydrogen in a MG is for residential applications. The power generation of a photovoltaic system in the winter months is usually decreased, which means private consumers have to rely on the power grid to supply their homes with energy. With the integration of an electrolyzer and a fuel cell into the residential microgrid, shown in Fig. 2, the excess power production during the summer can be utilized to generate hydrogen. The hydrogen can then be converted back to electric energy in the winter to supply the house. The efficiency of the hydrogen system can further be increased by feeding the waste heat into the hot water supply of the house. There are already different commercially available solutions, although these are currently designed for an AC topology.[8] A further concept is the interconnection of such DC MGs, as shown in Fig. 3. Based on the local prerequisites there are different kinds of consumers and producers. In this example, there is an Industrial complex, a residential area, a storage grid, and a power production grid. Usually, these sectors are connected by using the central PG. In this concept, however, the sectors are directly interconnected with one another and form a DC MG cluster, which reduces the reliance on the central PG. As the production and consumption of electrical energy are decentralized, the strain on the PG can be reduced and the resilience of the power supply can be increased. By using a hydrogen storage system in this concept, seasonal deviations in power production can be compensated. This concept is further explained by the author [8], who also presents an algorithm to scale the individual components of this cluster.

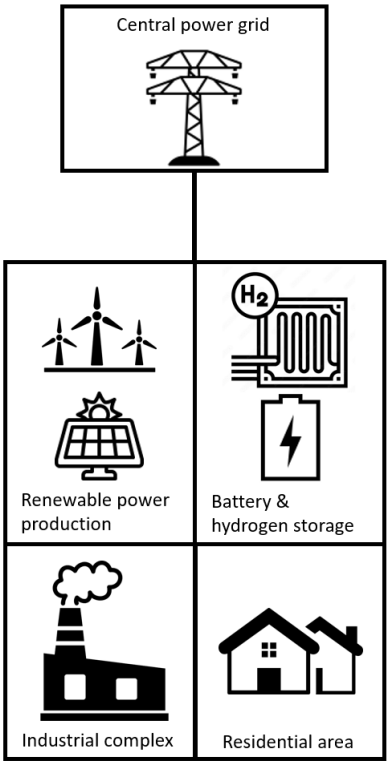

Fig. 3. DC Microgrids for decentral energy distribution based on [9]

B. Hydrogen wind farms

A further application scenario for DC MGs can be found in wind farms. Countries such as the UK, Germany, and

Denmark, are further increasing the number of offshore wind farms to increase the production of renewable energy[10]. One challenge, in this case, is the construction of the transmission lines, as the wind farms often have a distance of more than 100 kilometers to the shore. The required undersea transmission lines and the connection to the PG result in meaningful costs. The same goes for onshore wind turbines, as these are usually installed in rural areas. Due to local limitations in the PG wind turbines often have to be turned down, as the capacity of the transmission lines is limited[11]. There are however certain hydrogen-based concepts that can solve these problems. In Fig. 4 a diagram for an off-grid wind turbine is shown. In an off-grid system, the power is directly fed into the HE to generate hydrogen. This concept is especially useful for offshore wind farms, as no undersea cables have to be constructed. The hydrogen can then be transported with a ship or a hydrogen pipeline to be later used in the industry or for long-term energy storage [12]. This concept also increases the efficiency of the wind turbines, as the DC can be directly fed into the DC-based electrolyzer. In PG-connected wind turbines, the AC from the generator is first rectified and then gets inverted back AC to match the frequency of the PG. The losses of the invertion back to AC can be avoided in this concept. The authors in [13] simulated these concepts and showed that by combining wind power with HEs the capacity factor of wind turbines can be increased, auxiliary equipment can be saved and the cost of renewable electric energy can be reduced. In this way, the shown concepts can benefit the move towards fully renewable power production.[14]

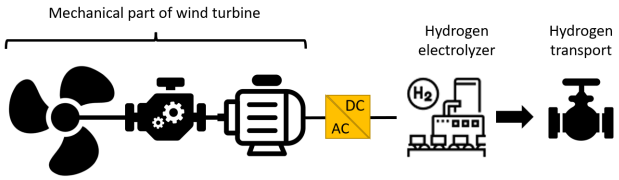

Fig. 4. DC MG concept for off-grid wind farms based on [13]

III. Storage and Transportation

One technical challenge in establishing a hydrogen-based energy supply system is the construction of a transportation system. The hydrogen needs to be transported from the HEs to the consumers. The diagram in Fig. 5 shows one potential layout. This layout is similar to a natural gas transportation system and is also equipped with a compressor station, a pipeline, and storage tanks. Due to different material properties, the transportation of hydrogen differs in certain aspects. The possibilities for storing hydrogen and transporting it with pipelines will be presented. It will also be shown how the existing natural gas infrastructure can be repurposed to handle hydrogen.

A. Storage technology

There are currently two established ways of storing hydrogen. One option is by compressing the hydrogen and storing

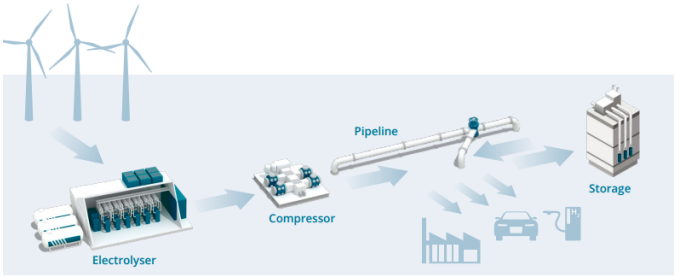

Fig. 5. Transportation system for hydrogen[15]

it in vessels with pressures of up to 800 bars. The currently dominant way for storing high volumes of hydrogen is by liquefying it. For this the hydrogen needs to be cooled below -240°C with a maximum pressure of 1,300kPa to reach a state below the critical point. Liquid hydrogen has a higher volumetric energy density than compressed hydrogen and transportation with vehicles is, therefore, more efficient. The energy demand for liquefying hydrogen is estimated to be around 18 – 30 % of the total stored energy. Additionally, as the hydrogen in the tank regasifies due to the ambient temperature some of the hydrogen gas has to be blown off with a pressure relief valve. This is estimated to account for daily losses of up to 1 % of the stored energy.[16]
Hydrogen can be stored in its liquid or gaseous state. Due to the stated boil-off losses hydrogen is usually stored in its gaseous state, when talking about long-term storage. As the role of hydrogen as an energy storage technology mainly concerns long-term storage, the storage of gaseous hydrogen will be further explained. The two most common types of natural gas storage are pore storage and cavern storage. This way of storing natural gas has proven to be the most efficient, therefore the possibilities of repurposing these storage facilities for hydrogen will be shown. In most cases, pore storages are depleted natural gas reservoirs, where the gas was stored in porous rocks. Natural gas can be pressed back into the reservoir. This type of storage can receive large volumes of gas; however, the storage process is time-consuming and therefore does not fit the demand for hydrogen storage. cavern storages are the result of mining salt. Brine is formed by pumping freshwater into underground salt reservoirs. When the brine is removed to extract the salt, an underground cavern remains. This cavern is then often repurposed for storing energy. 2/3 of the German storage capacities are cavern storage [17]. The injection and withdrawal are faster compared to pore storage, as the volume gets controlled by regulating the amount of brine in the cavern. This is beneficial for the storage of hydrogen, as it will be mainly used to compensate for the short-term fluctuations in renewable power production. The withdrawal time for natural gas is usually no limiting factor, as these storages mostly serve the long-time supply of energy [18]. The author in [15] states several technical challenges that still need to be researched, such as the permeability of the salt cavern, the long-term stability, and the qualification of

the materials of components. Another problem is the mixing with methane due to remaining stocks or eventual microbial activities in the cavern. The first German pilot project for the underground storage of hydrogen started in 2021 [19].

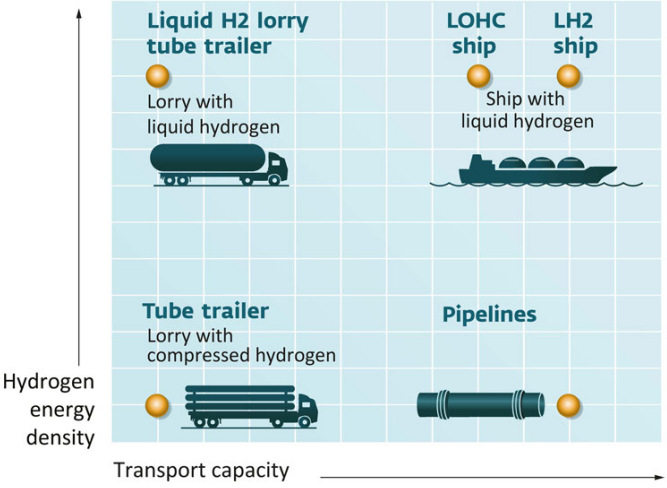

Fig. 6. Use-cases for transportation technologies [20]

B. Transportation

As the hydrogen is not directly consumed where it is produced, it needs to be transported. As with storing hydrogen, there are different ways of transporting it. Fig. 6. shows the different use cases for compressed and liquefied hydrogen. Liquid hydrogen provides a higher energy density than compressed hydrogen and therefore reduces the transportation cost, although the liquefication process requires an additional amount of energy. The adequate way of transportation depends on the available infrastructure and the demand of the consumer. For low energy demands the transport of compressed or liquefied hydrogen with a truck is the adequate solution, as a dedicated pipeline is not available or too expensive to construct. The transport of liquefied hydrogen with ships is beneficial in areas where hydrogen needs to be transported over long distances. One notable example is the Canada-Germany Hydrogen Alliance. Beginning in 2025 shipments of green hydrogen from Canada are planned to arrive in Germany to reduce the climate impact of the german industry [21]. Transportation with ships is the best solution for such distances and capacities and is a scalable approach. It is also possible to trade with different regions of the world.

The last possibility is transportation with a pipeline network. Dedicated hydrogen pipelines are already in use for energy-intensive companies. One example is the German chemical industry which uses around 400 kilometers of hydrogen pipelines [22]. With the phase-out of natural gas, it is only rational to repurpose the already existing pipeline network to handle hydrogen. The technical challenges will be presented in this chapter.

One important aspect is the transport capacity of the pipeline network. As the author [15] states, the pipeline capacity when using hydrogen is slightly lower than for the usage of natural gas. In general, the capacity of a pipeline depends on several values, such as volume, density, or flow velocity of the transport gas. The upper calorific value of hydrogen is about three times lower than the calorific value of natural gas. As the density of hydrogen is lower than the density of natural gas, the resulting volumetric flow of hydrogen is higher at the same pressure level. To achieve an identical energy transport capacity, the volumetric flow of the pipelines has to be increased. New compressors have to be installed to achieve this higher volumetric flow capacity. According to the author [15], the compressors, that are currently sold are already able to handle hydrogen from a materialwise perspective.

One technical challenge when handling hydrogen is the danger of hydrogen embrittlement. As hydrogen atoms are small, they can permeate solid metals and thereby reduce the ductility of the material. This raises the probability of cracks in the surface. The author in [15] estimates this risk to be low, as the hydrogen in the pipeline is in a molecular state. Additionally, hydrogen embrittlement results in the growth of crack-like structures in this scenario. As these crack-like structures are rare in pipelines, the overall risk can be evaluated as low. A report published by various European transmission system operators comes to the same conclusion and adds the possibility of internally coating the pipelines to further reduce the risk of hydrogen embrittlement. As part of this report tests were conducted to research the suitability of valves and seals in pipelines. The authors state that partial or full replacements of these components will be needed depending on regional variations. The authors estimate that the cost of repurposing existing natural gas pipelines to handle hydrogen is between 10 to 25 % of the construction cost of a dedicated hydrogen pipeline [23]. It can be seen that the repurposing of natural gas pipelines is technically and economically viable on a large scale.

IV. Conclusion

This literature research shows that there are many different concepts for the integration of hydrogen into the energy supply system. The shown concepts are mainly based on new developments due to the growing share of renewable energies. The storage and transportation of hydrogen on a large scale can be compared to that of natural gas, although there are key differences and technical challenges that still need to be further researched. Whilst existing natural gas infrastructure can be repurposed to handle hydrogen, further testing needs to be conducted and investments are required. The success of these concepts is strongly dependent on the regional prerequisites and the framework of specific renewable energy projects.

References

[1] L. Bongers, N. Mararakanye and B. Bekker, "Operating Reserve Dimensioning Methodologies for Renewable Energy Aligned Power Systems," *56th International Universities Power Engineering Conference (UPEC)* 2021, pp. 1-6, doi: 10.1109/UPEC50034.2021.9548228.

[2] S. O. Amrouche, D. Rekioua and T. Rekioua, "Overview of energy storage in renewable energy systems," in *3rd International Renewable and Sustainable Energy Conference (IRSEC)* 2015, pp. 1-6, doi: 10.1109/IRSEC.2015.7454988.
[3] K. Shimomachi, Y. Mishima, R. Hara and H. Kita, "Fuel cell and Electrolyzer System for Supply and Demand Balancing in DC," in *IEEE Third International Conference on DC Microgrids (ICDCM)* 2019, pp. 1-4, doi: 10.1109/ICDCM45535.2019.9232742.
[4] H. Ahmad, S. Coppens and B. Uzunoglu, "Connection of an Offshore Wind Park to HVDC Converter Platform without Using Offshore AC Collector Platforms," in *IEEE Green Technologies Conference (GreenTech)* 2013, pp. 400-406, doi: 10.1109/GreenTech.2013.68.
[5] Dr. Thomas Müller, "Electricity generation and consumption in germany" smard.de, https://www.smard.de/en (accessed November 28, 2022)
[6] L. Meng, M. Savaghebi, F. Andrade, J. C. Vasquez, J. M. Guerrero and M. Graells, "MG central controller development and hierarchical control implementation in the intelligent MG lab of Aalborg University," in *IEEE Applied Power Electronics Conference and Exposition (APEC)* 2015, pp. 2585-2592, doi: 10.1109/APEC.2015.7104716.
[7] L. E. Zubieta and P. W. Lehn, "A high efficiency unidirectional DC/DC converter for integrating distributed resources into DC MGs," in *IEEE First International Conference on DC MGs (ICDCM)* 2015, pp. 280-284, doi: 10.1109/ICDCM.2015.7152054.
[8] HPS Home Power Solutions AG, "Produkt" homepowersolutions.de https://www.homepowersolutions.de/en/product/(accessed December 30, 2022)
[9] B. Li, "Build 100% renewable energy based power station and MG clusters through hydrogen-centric storage systems," in *4th International Conference on HVDC (HVDC)* 2020, pp. 1253-1257, doi: 10.1109/HVDC50696.2020.9292836.
[10] European commission, "Offshore renewable energy", https://energy.ec.europa.eu/topics/renewable-energy/offshore-renewable-energy_en (accessed November 28, 2022)
[11] Huenteler, Joern, Tang, Tian, Chan, Gabriel, Anadon, Laura "Why Is China's wind power generation not living up to Its potential?", in *Environmental Research Letters* 13. doi: 10.1088/1748-9326/aaadeb.
[12] J. Li and S. Obara, "Study on Hydrogen energy supply system with natural gas pipeline in China," *IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC)* 2019, pp. 1-4, doi: 10.1109/APPEEC45492.2019.8994477.
[13] Z. Yan and W. Gu, "Research on integrated system of non-grid-connected wind power and water-electrolytic hydrogen production," in *World Non-Grid-Connected Wind Power and Energy Conference* 2010, pp. 1-4, doi: 10.1109/WNWEC.2010.5673278.
[14] Q. Wang and L. Pan, "Study on floating offshore wind power hydrogen production and hydrogen production ship," in *IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA)* 2022, pp. 899-905, doi: 10.1109/ICAICA54878.2022.9844498.
[15] Peter Adam, Frank Heunemann, Christoph von dem Bussche, Stefan Engelshove, Thomas Thiemann, "Hydrogen infrastructure - the pillar of energy transition", 2020
[16] Aziz, M., "Liquid Hydrogen: A Review on Liquefaction, Storage, Transportation, and Safety" in Energies 2021, 14, 5917. https://doi.org/10.3390/en14185917
[17] U. Lubenau et al., "H_2 short study: Hydrogen quality in an overall German hydrogen network", 2022
[18] Y. Zhu, "From Hydrogen Production to Storage: A Process for Sustainable Development of Underground Hydrogen Storage in Caverns," in *3rd International Academic Exchange Conference on Science and Technology Innovation (IAECST)* 2021, pp. 1658-1663, doi: 10.1109/IAECST54258.2021.9695812.
[19] VNG AG, Energiepark Bad Lauchstädt, https://energiepark-bad-lauchstaedt.de/ (accessed November 28, 2022)
[20] N.V. Nederlandse Gasunie, "Hydrogen infrastructure" theworldofhydrogen.com, https://www.theworldofhydrogen.com/gasunie/infrastructure/ (accessed December 30, 2022)
[21] Daniela Ewers, "Speeding up the roll-out of green hydrogen: Canada and Germany set up Hydrogen Alliance" bmwk.de, https://www.bmwk.de/Redaktion/EN/Pressemitteilungen/2022/08/20220823-speeding-up-the-roll-out-of-green-hydrogen-canada-and-germany.html (accessed December 30, 2022)
[22] TÜV Nord, "Building a national hydrogen grid with hydrogen pipelines", https://www.tuev-nord.de/en/company/energy/hydrogen/hydrogen-pipelines-grids (accessed November 28, 2022)
[23] Anthony Wang, Kees van der Leun, Daan Peters, Maud Buseman, "How a dedicated hydrogen Infrastructure can be created" in *European Hydrogen Backbone* 2020

Low power microcontrollers: Existing power saving concepts in state-of-the-art microcontrollers and how to apply them

Reinhold Belz
Fakultät Elektro- und Informationstechnik
Ostbayerische Technische Hochschule Regensburg
Regensburg, Deutschland
reinhold.belz@st.oth-regensburg.de

***Abstract*—By introducing power-saving modes, microcontrollers can be operated with batteries or renewable energy sources. This makes it possible to develop applications within the context of the Internet of Things that can be operated self-sufficiently without the need for a power grid. However, having limited energy storage requires the application to be operated in the most energy-efficient way possible. For this reason, there are various approaches to enable power-efficient operation. This paper addresses the energy efficient operation of state-of-the-art microcontrollers. In this regard, the main aspects of energy consumption when operating a microcontroller using power saving modes are considered. In view of these aspects, low power hardware components and logic circuits can be used to implement power saving concepts, which are presented here. Various concepts are applied through software, but have different impact on the overall possible energy savings and on the availability of the device. Depending on the requirements, a compromise must be found between these properties, which will be elaborated in this article. The total energy demand can be reduced by configuring duty cycles. They distinguish between an active phase, in which the actual function is executed, and an energy-saving phase, in which as little energy as possible is consumed. Energy can be saved by keeping the active phase as short as possible and further optimizing the software through understanding the energy cost of different operations. In addition, proper configuration of input and output logic is also an important factor when it comes to energy savings.**

***Index Terms*—microcontroller, energy efficient, low power, energy consumption, Internet of Things, embedded systems, power demand**

I. Introduction

Low power microcontrollers are designed to minimize power consumption while still providing the necessary processing capabilities for a wide range of applications. These microcontrollers (MCU) are often used in battery-powered devices and other systems where power consumption is a critical concern. The Internet of Things sector in particular benefits from the use of MCUs with low power consumption. Examples of use cases include waste management with the aid of sensors in smart cities, air quality monitoring with gas measurement sensors, or sensor-based activity recognition in smart homes [1]–[4]. Wireless sensor networks or portable devices also use these MCUs to ensure long operation without a constant power supply. There are a variety of power-saving concepts and techniques that have been researched over the past 50 years. In the following, concepts are presented that are used in state-of-the-art microcontrollers, such as clock gating, dynamic voltages and others. While all these concepts aim to minimize power consumption, they have different impacts on the characteristics of the operation of a MCU to meet. For this reason, the application of the concept of different energy saving modes has become widely adopted key features of low power MCUs. They allow a broad application of the concepts to meet different requirements for different use cases. In addition, there are several other techniques that can be combined with the power saving modes to further optimize energy consumption.

This paper presents the concepts for energy optimization and examines their application. Section II provides information on the power management of a low-power MCU before presenting the most common power saving concepts in Chapter III. Chapter IV describes how the application of the discussed techniques affects the MCU's computational performance, availability, and energy cost. A summary of the application of a power-saving MCU is presented in Chapter V.

II. Energy consumption aspects of a low power microcontroller

The components of a MCU are mainly digital CMOS circuits, which can be divided into a dynamic power demand while the circuit is active and a static power demand during an idle state of the circuit. The dynamic power demand consists mainly of the power required to switch transistor states and can be described by the following (1) [5]–[7]:

$$P_{\text{switch}} = C_{\text{eff}} \cdot V_{\text{dd}}^2 \cdot f \tag{1}$$

In contrast, when a component of the MCU is deactivated, the static current demand is dominated by the leakage current generated by the voltage applied to the transistors in the inactive, non-conducting state. This quiescent leakage current is caused by multiple sources, e.g. tunnel leakage current [8], [9]. In addition to fine-tuning transistor parameters, leakage

current can be reduced by lowering or removing the supply voltage applied to the transistor.

Using power from batteries or renewable energy sources, the MCU has to manage its supply voltage of the different digital circuits through a low dropout regulator (LDO) [10], [11]. In simplified terms, a LDO uses transistors to provide a controlled and thus stable output voltage. There are various sizes of the components for different voltage levels [12]. A low power MCU has multiple LDOs with different device parameters that allow switching between different voltage settings. A change in voltage comes with a power demand consisting of switching transistors and loading and discharge capacities. Not only a voltage alteration causes energy drain, but also the initialization of components that are either switched on or change from an inactive state to the active state.

The overall energy consumption can be described as:

$$E(t) = \int_{t_0}^{t_1} P(t)\,dt \tag{2}$$

This equation shows that the energy consumption depends both on the momentary power consumption and on the time period in which it takes effect. Therefore, the following concepts take advantage of the different energy characteristics of CMOS circuits by either spending less time with a component in active operation or reducing leakage currents, resulting in lower overall power consumption. To realize the following power saving concepts, a low power MCU has components designed to operate with different power supply modes. A generic structure can be seen in Figure 1. The structure varies from MCU to MCU, but all have a mechanism such as a real-time clock (RTC) that is responsible for switching between different operation modes.

III. Power saving concepts

It is possible to save energy by either reducing the dynamic energy demand by switching off components of a MCU or by reducing the energy demand during dynamic operation of a CMOS circuit. In the following, various energy saving techniques are presented which a developer can apply via software, either by using special power-saving hardware designs or by avoiding unnecessary energy losses due to inefficient operation.

A. Clock gating

In low power applications the central processing unit (CPU) is not permanently active but sporadically executes an action. Considering (1), energy can be saved by running the CPU at a lower clock frequency for the period it is idle or by selectively removing the clock signal for that period. The latter can be described as clock gating, where the clock signal is being prevented from reaching the circuit of a component during the idle state via a logic block and control signals [5], [14]. This strategy can be applied to all components of a MCU and effectively halt their operation to preserve energy.

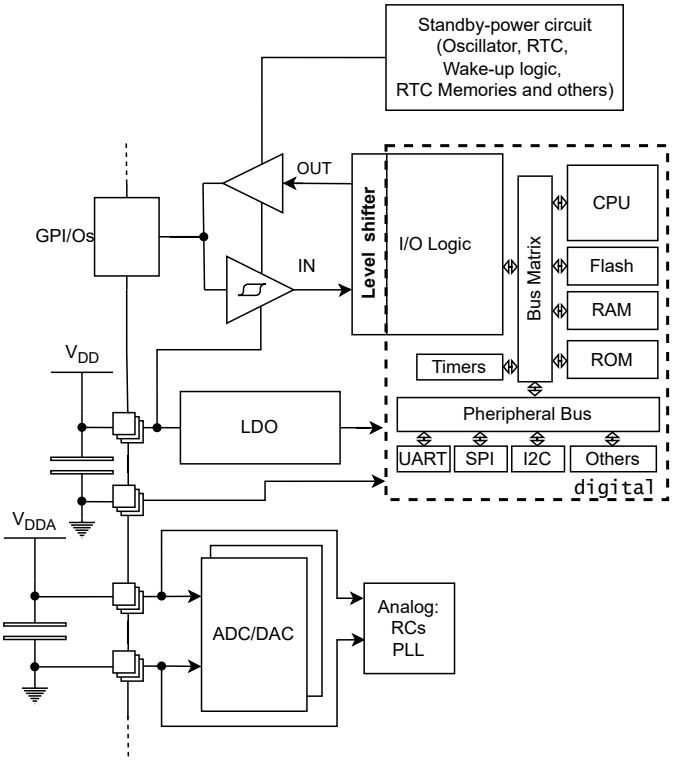

Fig. 1. Basic Structure of a MCU [13]

B. Application of a fast clock rate

Components of a MCU, which can be operated with a certain frequency, like e.g. the CPU or analog-to-digital converter (ADC), require more power during operation on the one hand, but less time to perform an operation on the other hand, as can be seen by formula (1) and (2). Chung et al. [15] stated that a minimum of this ratio can be achieved in a certain frequency range. However, in the context of a low-power MCU, the energy requirements of the operating mode change have an additional effect. The active phase of a device requires much more power than an inactive phase. Therefore selecting the fastest possible clock frequency shortens this phase, resulting in a lower overall energy cost for operation [16]–[18].

C. Reducing supplied voltage

When a CMOS circuit is active, the supply voltage V_{dd} must be high enough to allow maximum load current for proper operation of the CMOS logic gates. However, when a component is inactive, the supply voltage causes a static current demand consisting of leakage current, which can be reduced by lowering the operating voltage by switching between the available LDOs [14], [19], [20].

D. Optimization of software

One power saving technique is to use a code analysis tool, which gives the programmer of a low power MCU hints about potential energy savings. But even without such tooling a software can reduce the overall energy cost by applying knowledge of several properties of a MCU.

The execution of software is a sequential flow of instructions, which are loaded from a memory into the CPU and if necessary new information is written into the memory. The MCU requires energy for each memory access and instruction call. For instance this can account for about 70 percent of energy consumption during active operation, according to [21], [22]. Therefore, it makes sense from an energy point of view to perform only necessary write and read operations and to use deliberate register allocation in order to reduce the cost of memory accesses [23], [24]. Another aspect is the different length of the pipeline cycles of an instruction. For example, a multiplication may require two pipeline cycles, while a division requires twelve cycles [18]. The precise rewriting of certain operations can thus also contribute to energy-efficient operation.

E. Low power GPIO configuration

When it comes to operating peripherals with a MCU, a developer can configure general purpose input and output (GPIO) pins. For each GPIO pin there are two transistors which control digital output signals. In addition, for each pin there is either a pull-down or a pull-up resistor to pull the pin's line up or down. The configuration of a GPIO takes effect on the energy consumption especially during a longterm inactive period. When I/O logic is inactive, there can be energy loss due to current flowing through a resistor, producing unwanted thermal energy. Also if a unused pin stays undefined, the floating voltage on that pin may drive the transistor as a resistive component, resulting in increased current draw. A low-power application should set all existing GPIO pins to a state where power consumption is lowest [25].

F. Power saving modes

With the help of a RTC the MCU is able to make use of duty cycles. They enable the determination of time windows in which different operating modes can be utilized. This makes it possible to temporarily switch to an active mode for the sporadic execution of a task and to operate in a low-power mode during the time when the MCU is not supposed to execute any functions. In the energy-saving modes, individual components or all components can either be operated with a lower clock frequency. Their operation can be stopped by means of clock gating or switched off by removing the supply voltage. The possible combinations are versatile and can be adapted to each individual application by the developer. After a program has completed its task, the MCU switches to the configured energy-saving mode. In this mode, only a few components are still active, including the RTC [26], [27]. When the RTC has exceeded a certain amount of time, it triggers an interrupt. This starts a wakeup routine, which changes to an active operating mode and the chain of events can start the next cycle. The interrupt that causes the wake-up can also be triggered by other sources, e.g. to prompt a radio module to wake up the MCU when an incoming signal is registered. In addition, a low-power RTC memory can be integrated to store register values for the inactive time of a CPU or other components to reduce wake-up time. By using an ultra low power co-processor, peripherals can be operated that also allow an interrupt to be triggered.

IV. Application

The power saving concepts are implemented by different low power profiles that are predefined by the MCU manufacturer or can be configured by the programmer. The application of these power saving techniques requires a trade-off between energy saving and the constraints imposed by the concept used. Switching between profiles involves a wake-up process that consumes power and requires time to activate the supply voltage via the LDO regulator. In addition, the clock signals must be initialized and the memory registers must be accessed to load instructions and process data. The resulting wake-up time and the energy required for wake-up are different for each change to active mode, since different levels of activity must be performed. Therefore, we will discuss the most common profile in state-of-the-art MCUs and the associated system constraints below. The listed application profiles are arranged in an increasing rate of energy saving, but also in an increasing wake-up time.

1) Active state: All components of a MCU are switched on and are supplied with their corresponding clock signals. The MCU starts in this mode and is fully functional as conventional MCUs.

2) Sub-active state: This state is similar to the active state, except that the supply voltage of the CPU and the clock frequency are lowered to save energy, as described in section III-C. This low-voltage operation is suitable for applications where a parallel process must be awaited, and operation at a higher signal frequency would not reduce the total time in the active state. To switch back to the fully active state, the clock signal must be reactivated and the voltage regulator needs to step up the supply voltage. Even if the wake-up time is short, there is a constant dynamic energy demand that is always higher than the static energy demand of inactive components.

3) Sleep state: In sleep mode, the CPU is disabled via clock gating, but other system components including the system clock continue to operate. This allows monitoring an event to reactivate the CPU, while the peripherals are already initialized and therefore do not require additional power and time to wake up. The energy now required is reduced by the dynamic demand of an active CPU.

4) Deep sleep state: This state is also called stop state. Here the system clock, CPU, flash memory and the peripherals are switched off. The voltage regulator enables low voltage supply. The RTC is still running and can be used to trigger a periodic return to the active state. It is also possible to trigger a wake-up with an external component by any GPIO pin [29]. If the RTC is not needed it can also be deactivated. The required power in this state is reduced even further by saving energy needed from the system clock, the flash memory and the peripherals. The wake-up time is longer than the wake-up time from sleep mode since the system clock must also be restarted and this process requires additional time to stabilize [27].

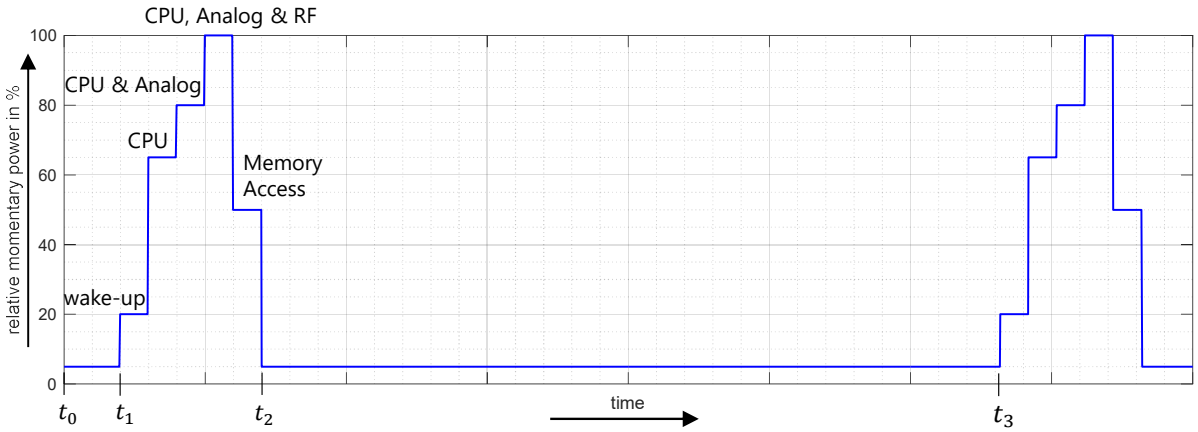

Fig. 2. Exmaple application of power profiles. Following [28].

5) Standby: In standby mode, the system clocks and the internal voltage regulator are switched off. The digital components of the MCU are completely prevented from operating. All volatile memory is lost, but the values of the registers can be stored in a power-saving memory. The RTC is still operating in this state and provides a wake-up mechanism. There is also the possibility to wake up the MCU with special wake-up pins. This state has the longest wake-up time, but also the lowest energy consumption before the MCU is completely shut down.

With the use of these profiles, an application that sporadically performs an action can be operated in an energy-saving manner. This can be, for example, an analog sensor that detects a value and then sends it via a radio module. In Picture 2 this example is shown as momentary power consumption over time. For generalization, the momentary power is given as a relative value, where 100% corresponds to the maximum power draw and the energy-saving mode requires about 5%. Here the MCU enters the active mode and after a wake-up period it executes its measurement and reports the results via a radio module to another system. Before switching to deep sleep mode it writes their low power memory. In this scenario, the MCU enters active mode at time t_1, performs its measurement after a wake-up phase and reports the results to another system via a radio module. Before entering deep sleep mode at time t_2, it writes information for the next cycle to the power saving memory. At time t_3, the cycle starts anew. The time spent in deep sleep between t_2 and t_3 is much longer than in an active state between t_1 and t_2.

The use of the different profiles makes it possible to optimize the energy consumption of an application according to its requirements by reducing the energy costs for active waiting or inactivity. Considering the equation (2) in relation to the example, one can see the great impact the different power modes have, since the area under the long inactive section can be greatly reduced by the low power profiles. This effect can even be enhanced by the efficient use of software, which can reduce the energy required in each active phase of a MCU by using special hardware or avoiding energy-consuming operations by using alternative methods, such as replacing a division with a multiplication [18]. But also the inactive part can be influenced by configurations. According to [25] the wrong configuration of GPIO pins can require up to 1500 times more power during the sleep phase. The optimization of software can either be achieved by the knowledge of the system behind the MCU or by the use of static code analysis tools such as Ultra Low Power Advisor [30].

V. Conclusion

The power saving concepts can be examined on the basis of the energy-relevant properties of the power management of a MCU. Based on state-of-the-art MCUs, we have summarized the most common concepts and provided information on their application. A low-power MCU can be operated in duty cycles to divide the power demand into a high but short-term active phase and a long-term inactive phase, where unused components are switched off or the operation of the CPU is adapted to the specific application. This is possible through the use of wake-up signals generated by an event from a RTC or peripheral device, and multiple low-power profiles. The transition from one state to the other requires a wake-up time and wake-up energy. We have provided some insight into the cause of these characteristics. Using this information, an application can be designed for energy efficient operation and even further optimized by applying energy important configurations, spending as little time as possible in active mode, using a fast clock rate for ADC or CPU, and optimizing software.

References

[1] M. Cerchecci, F. Luti, A. Mecocci, S. Parrino, G. Peruzzi, and A. Pozzebon, “A low power iot sensor node architecture for waste management within smart cities context,” *Sensors*, vol. 18, no. 4, Apr. 2018.

[2] K. Pardini, J. J. P. C. Rodrigues, S. A. Kozlov, N. Kumar, and V. Furtado, "Iot-based solid waste management solutions: A survey," *Journal of Sensor and Actuator Networks*, vol. 8, no. 1, Jan. 2019.

[3] S. Nedelcu, K. Thodkar, and C. Hierold, "A customizable, low-power, wireless, embedded sensing platform for resistive nanoscale sensors," *Microsystems & Nanoengineering*, vol. 8, no. 1, Jan. 2022.

[4] D. Park, J. M. Youn, and J. Cho, "A Low-Power Microcontroller with Accuracy-Controlled Event-Driven Signal Processing Unit for Rare-Event Activity-Sensing IoT Devices," *Journal of Sensors*, vol. 2015, p. 10, Aug. 2015.

[5] D. Flynn, R. Aitken, A. Gibbons, and K. Shi, *Low power methodology manual*, ser. Integrated Circuits and Systems. New York: Springer, Dec. 2010.

[6] Y. Dhillon, A. Diril, A. Chatterjee, and H.-H. S. Lee, "Algorithm for achieving minimum energy consumption in cmos circuits using multiple supply and threshold voltages at the module level," in *ICCAD-2003. International Conference on Computer Aided Design*, Nov. 2003, pp. 693–700.

[7] T. Burd and R. Brodersen, "Energy efficient cmos microprocessor design," in *Proceedings of the Twenty-Eighth Annual Hawaii International Conference on System Sciences*, vol. 1, Jan. 1995, pp. 288–297.

[8] N. Kim, T. Austin, D. Baauw, T. Mudge, K. Flautner, J. Hu, M. Irwin, M. Kandemir, and V. Narayanan, "Leakage current: Moore's law meets static power," *Computer*, vol. 36, no. 12, pp. 68–75, 2003.

[9] K. Roy, S. Mukhopadhyay, and H. Mahmoodi-Meimand, "Leakage current mechanisms and leakage reduction techniques in deep-submicrometer cmos circuits," *Proceedings of the IEEE*, vol. 91, no. 2, pp. 305–327, 2003.

[10] K. Keikhosravy and S. Mirabbasi, "A 0.13-μm CMOS Low-Power Capacitor-Less LDO Regulator Using Bulk-Modulation Technique," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 61, no. 11, pp. 3105–3114, Nov. 2014.

[11] A. Maity and A. Patra, "A single-stage low-dropout regulator with a wide dynamic range for generic applications," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 24, no. 6, pp. 2117–2127, 2016.

[12] D. Khan, M. Basim, Q. u. Ain, S. A. A. Shah, K. Shehzad, D. Verma, and K.-Y. Lee, "Design of a Power Regulated Circuit with Multiple LDOs for SoC Applications," *Electronics*, vol. 11, no. 17, p. 2774, Jan. 2022.

[13] *Ultra-low-power 32-bit MCU ARM®-based Cortex®-M3, 128KB Flash, 16KB SRAM, 4KB EEPROM, LCD, USB, ADC, DAC*, STMicroelectronics, 4 2016, rev. 12. [Online]. Available: https://www.st.com/resource/en/datasheet/stm32l151r6.pdf

[14] M. Chakraverty, H. PS, and V. Ruparelia, "Low power design practices for power optimization at the logic and architecture levels for VLSI system design," in *2016 International Conference on Energy Efficient Technologies for Sustainability (ICEETS)*, Apr. 2016, pp. 727–733.

[15] A. Dzhagaryan and A. Milenković, "Impact of thread and frequency scaling on performance and energy in modern multicores: a measurement-based study," in *Proceedings of the 2014 ACM Southeast Regional Conference*, ser. ACM SE '14. New York, NY, USA: Association for Computing Machinery, Mar. 2014, pp. 1–6.

[16] H. Wu, C. Chen, and K. Weng, "An Energy-Efficient Strategy for Microcontrollers," *Applied Sciences*, vol. 11, no. 6, p. 2581, Jan. 2021.

[17] I. Tsekoura, G. Rebel, P. Glösekötter, and M. Berekovic, "An evaluation of energy efficient microcontrollers," in *2014 9th International Symposium on Reconfigurable and Communication-Centric Systems-on-Chip (ReCoSoC)*, May 2014, pp. 1–5.

[18] M. M. Al-Kofahi, M. Y. Al-Shorman, and O. M. Al-Kofahi, "Toward energy efficient microcontrollers and Internet-of-Things systems," *Computers & Electrical Engineering*, vol. 79, p. 106457, Oct. 2019.

[19] K. Mazumdar, S. Bartling, S. Khanna, and M. Stan, "A digitally-controlled power-aware low-dropout regulator to reduce standby current drain in ultra-low-power MCU," in *Sixteenth International Symposium on Quality Electronic Design*, Mar. 2015, pp. 98–102.

[20] M. Lueders, B. Eversmann, J. Gerber, K. Huber, R. Kuhn, M. Zwerg, D. Schmitt-Landsiedel, and R. Brederlow, "Architectural and Circuit Design Techniques for Power Management of Ultra-Low-Power MCU Systems," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 22, no. 11, pp. 2287–2296, Nov. 2014.

[21] W. J. Dally, J. Balfour, D. Black-Shaffer, J. Chen, R. C. Harting, V. Parikh, J. Park, and D. Sheffield, "Efficient Embedded Computing," *Computer*, vol. 41, no. 7, pp. 27–32, Jul. 2008.

[22] V. Konstantakos, A. Chatzigeorgiou, S. Nikolaidis, and T. Laopoulos, "Energy Consumption Estimation in Embedded Systems," *IEEE Transactions on Instrumentation and Measurement*, vol. 57, no. 4, pp. 797–804, Apr. 2008.

[23] V. Tiwari, S. Malik, A. Wolfe, and M.-C. Lee, "Instruction level power analysis and optimization of software," in *Proceedings of 9th International Conference on VLSI Design*, Jan. 1996, pp. 326–328.

[24] S. Liu, K. Pattabiraman, T. Moscibroda, and B. G. Zorn, "Flikker: Saving dram refresh-power through critical data partitioning," *SIGPLAN Not.*, vol. 46, no. 3, p. 213–224, Mar 2011.

[25] X. Ji, X. Zhou, M. Xu, W. Xu, and Y. Dong, "OPCIO: Optimizing Power Consumption for Embedded Devices via GPIO Configuration," *ACM Transactions on Sensor Networks*, vol. 16, no. 2, pp. 16:1–16:28, Jan. 2020.

[26] "Enabling Timekeeping Function and Prolonging Battery Life in Low Power Systems." [Online]. Available: https://www.digikey.com/en/articles/enabling-timekeeping-function-and-prolonging-battery-life-in-low-power-systems

[27] T. Ishizu, K. Furutani, Y. Yakubo, A. Isobe, M. Fujita, T. Atsumi, Y. Ando, T. Murakawa, K. Kato, M. Fujita, and S. Yamazaki, "An energy-efficient normally off microcontroller with 880-nw standby power, 1 clock system backup, and 4.69- μ s wakeup featuring 60-nm caac-igzo fets," *IEEE Solid-State Circuits Letters*, vol. 2, no. 12, pp. 293–296, 2019.

[28] W. Lutsch and L. Reese, "Energy Optimization Tools for Ultra-Low-Power Microcontrollers." [Online]. Available: https://eu.mouser.com/applications/low-power-ewc-design/

[29] *Ultra-low-power 32-bit MCU Arm®-based Cortex®-M0+, 64KB Flash, 8KB SRAM, 2KB EEPROM, LCD, USB, ADC, DAC, AES*, STMicroelectronics, 6 2019, rev. 8. [Online]. Available: https://www.st.com/resource/en/datasheet/stm32l063r8.pdf

[30] V. Iossifov, "Programming Techniques for Energy-Efficient Software," in *Advanced Computing in Industrial Mathematics*, ser. Studies in Computational Intelligence, I. Georgiev, H. Kostadinov, and E. Lilkova, Eds. Cham: Springer International Publishing, 2021, pp. 154–169.