TU Delft

Delft University of Technology

Would Magnonic Circuits Outperform CMOS Counterparts?

Mahmoud, Abdulqader; Cucu-Laurenciu, Nicoleta; Vanderveken, Frederic; Ciubotaru, Florin; Adelmann, Christoph; Cotofana, Sorin; Hamdioui, Said

DOI
10.1145/3526241.3530368

Publication date
2022

Document Version
Accepted author manuscript

Published in
GLSVLSI 2022 - Proceedings of the Great Lakes Symposium on VLSI 2022

Citation (APA)
Mahmoud, A., Cucu-Laurenciu, N., Vanderveken, F., Ciubotaru, F., Adelmann, C., Cotofana, S., & Hamdioui, S. (2022). Would Magnonic Circuits Outperform CMOS Counterparts? In *GLSVLSI 2022 - Proceedings of the Great Lakes Symposium on VLSI 2022* (pp. 309-313). (Proceedings of the ACM Great Lakes Symposium on VLSI, GLSVLSI). ACM. https://doi.org/10.1145/3526241.3530368

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.

Would Magnonic Circuits Outperform CMOS Counterparts?

Abdulqader Mahmoud,[1, a)] Nicoleta Cucu-Laurenciu,[1] Frederic Vanderveken,[2] Florin Ciubotaru,[2] Christoph Adelmann,[2] Sorin Cotofana,[1] and Said Hamdioui[1]

[1)] *Delft University of Technology, 2628 CD Delft, The Netherlands*

[2)] *Imec, 3001 Leuven, Belgium*

In the early stages of a novel technology development, it is difficult to provide a comprehensive assessment of its potential capabilities and impact. Nevertheless, some preliminary estimates can be drawn and are certainly of great interest and in this paper we follow this line of reasoning within the framework of the Spin Wave (SW) based computing paradigm. In particular, we are interested in assessing the technological development horizon that needs to be reached in order to unleash the full SW paradigm potential such that SW circuits can outperform CMOS counterparts in terms of energy consumption. In view of the zero power SWs propagation through ferromagnetic waveguides, the overall SW circuit power consumption is determined by the one associated to SWs generation and sensing by means of transducers. While current antenna based transducers are clearly power hungry recent developments indicate that magneto-electric (ME) cells have a great potential for ultra-low power SW generation and sensing. Given that MEs have been only proposed at the conceptual level and no actual experimental demonstration has been reported we cannot evaluate the impact of their utilization on the SW circuit energy consumption. However, we can perform a reverse engineering alike analysis to determine ME delay and power consumption upper bounds that can place SW circuits in the leading position. To this end, we utilize a 32-bit Brent-Kung Adder (BKA) as discussion vehicle and compute the maximum ME delay and power consumption that could potentially enable a SW implementation able to outperform its 7 nm CMOS counterpart. We evaluate different BKA SW implementations that rely on conversion- or normalization-based gate cascading and consider continuous or pulsed SW generation scenarios. Our evaluations indicate that 31 nW is the maximum transducer power consumption for which a 32-bit Brent-Kung SW implementation can outperform its 7 nm CMOS counterpart in terms of energy consumption.

[a)] Electronic mail: a.n.n.mahmoud@tudelft.nl

I. INTRODUCTION

In the last decade, there was an enormous increase in the raw data generation and collection because of information technology ubiquity. In addition, the processing and analysis of these huge amounts of data requires high performance computing platforms[1], which, during the last decades, where enabled by the CMOS downscaling iduced performance improvement. However, Dennard scaling becomes more and more difficult due to leakage, reliability, and economical walls[2],which predicts the near end of Moore's law. As a result, researchers have investigated alternative technologies such as graphene devices[3], memristors[4], and spintronics[5]. One of the spintronic technologies is the Spin Wave (SW) technology, the so called magnonics, which is promising for computing as[6–9]: i) it has ultra-low energy consumption as the electrons are spinning and not moving, ii) it is highly scalable as the SW wavelength, which is the only scalability limitation, can reach down to few nanometers, iii) it has an acceptable delay, and iv) it has natural support for parallelism feature as SWs with different frequencies can simultaneously propagate through the same waveguide without affecting each other.

Motivated by the SW interaction potential, researchers have investigated different SW logic gates and circuits[7–18]. In[10], a Mach-Zehnder interferometer has been utilized to design the first experimental SW NOT gate, whereas the Mach-Zehnder interferometer has been used to build different single output gates such as Majority, (N)AND, (N)OR, and X(N)OR gates in[10]. On the other hand, multi-output logic gates have been reported in[9,11,13]. Moreover, the unique parallelism feature of the SW technology has been utilized to demonstrate multi-frequency logic gates[8,12]. On the circuit side, different circuits have been illustrated at the conceptual level[14], simulation level[7,15–17], and physical millimeter scale level[18]. In[19–22], benchmarking of beyond CMOS technologies, e.g., SW, Graphene, SpinFET, has been performed by modelling these devices as a capacitor and parasitics, which is not quite appropriate for the SW case. Thus, to get better inside on the SW based computation paradigm potential, which is certainly a topic of great interest, a benchmarking based on more realistic device models and that considers the complications related to SW gate cascading and fanout achievement is required.

In view of the zero power SWs propagation through ferromagnetic waveguides, the overall magnionic circuit power consumption is determined by the one associated to SWs generation

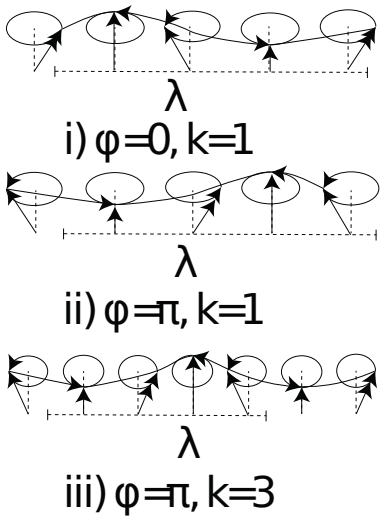

FIG. 1. Spin Wave Parameters.

and sensing by means of transducers. Recent developments indicate that magneto-electric (ME) cells[6] have a great potential for ultra-low power SW generation and sensing, but they have been only proposed at the conceptual level and no actual experimental demonstration has been reported.

Thus, since SW technology is still under development and the necessary insight for assessing its practical potential is not available, we reformulate in this paper the benchmarking process as a quest to identity the technological limits that need to be reached in order to fully unleash the SW potential such that SW circuits can outperform CMOS counterparts in terms of energy consumption. Thus, we assume a 32-bit Brent-Kung Adder (BKA) as discussion vehicle, and perform a reverse engineering process in order to determine the transducer power upper bound that allows its SW implementation to outperform the 7 nm CMOS counterpart. We evaluate different BKA SW implementations that rely on conversion- or normalization-based gate cascading, while considering continuous or pulsed SW generation scenarios. Our evaluations indicate that 31 nW is the maximum transducer power consumption for which a 32-bit Brent-Kung SW implementation can still outperform its 7 nm CMOS counterpart in terms of energy. Based on this assessment, we conclude the paper by discussing challenges and possible future directions toward building efficient magnonic circuits.

The paper organization is as follows. Section II explains the SW fundamentals and computing paradigm. Section III provides the possible implementation of the 32-bit Brent Kung Adder in SW technology. Section IV sums-up the paper.

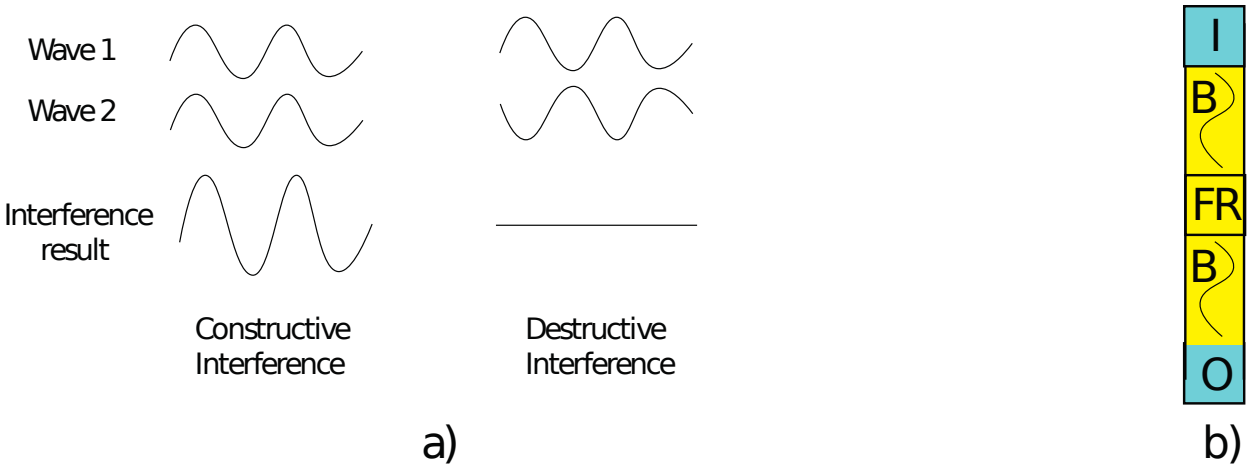

FIG. 2. a) Constructive and Destructive Interference, b) Spin Wave Device.

II. SPIN WAVE FUNDAMENTAL AND COMPUTING PARADIGM

For out of equilibrium magnetization states, the magnetization dynamics is governed by the Landau-Lifshitz-Gilbert (LLG) equation[6]

$$\frac{d\vec{M}}{dt} = -|\gamma|\mu_0 \left(\vec{M} \times \vec{H}_{eff}\right) + \frac{\alpha}{M_s}\left(\vec{M} \times \frac{d\vec{M}}{dt}\right), \tag{1}$$

where γ is the gyromagnetic ratio, μ_0 the vacuum permeability, M the magnetization, M_s the saturation magnetization, α the damping factor, and H_{eff} the effective field, which consists of the external field, exchange field, demagnetizing field, and magneto-crystalline field.

For small magnetic disturbances, Equation (1) can be linearized and results in wave-like solutions, known as Spin Waves (SWs), which can be seen as the collective excitation of the electron spins in the magnetic material[6]. A SW is described by its frequency f, amplitude A, phase ϕ, wavelength λ, and wavenumber $k = \frac{2\pi}{\lambda}$[6], which are graphically presented in Figure 1. The relation between the SW wavenumber and frequency is known as the dispersion relation and it is crucial for the SW circuit design[6].

SW amplitude, phase, and frequency can be utilized to encode information[6,8], which processing is governed by the wave interference principle. An example of SW interference is presented in Figure 2 a) where SWs are interfering constructively in the first case as they have the same phase $\Delta\phi = 0$, and destructively in the second case as they are out-of-phase $\Delta\phi = \pi$. Moreover, assuming SW phase information encoding, i.e., phase 0 and π represents logic 0 and 1, respectively, SWs interaction supports Majority function evaluation.

For instance, if 3 SWs having the same amplitude, frequency, and wavelength interfere in the same waveguide, the resultant SW has 0 phase, if at least 2 SWs have 0 phase, whereas the resultant SW has π phase, if at least 2 SWs have a π phase. Note that such an implementation in CMOS requires 18 transistors, whereas it can be directly implemented in SW technology with a single waveguide. One can easily deduce that more complex interference patterns can occur for SWs with different amplitude, frequency, wavenumber, and wavelength, which can be of great interest for developing future SW based computing paradigms.

Figure 2 b) presents a general structure of the SW device, which consists of four main regions[6,23]: i) Excitation region (I), ii) Waveguide (B), iii) Functional Region (FR), and iv) Detection region (O). At the excitation region, the SW is excited by means of voltage- or current- driven techniques such as microstrip antennas[6], MagnetoElectric (ME) cells[24], or spin orbit torques[6]. After the excitation, the SWs propagates through the magnetic waveguide and reach the functional region, where it can be manipulated, i.e., amplified, normalized, or interfere with other SWs. Finally, at the detection region, the SW is detected by similar or different methods than in the excitation region[6,23].

III. SW TRANSDUCER POWER UPPER BOUND

As stated into the introduction, our goal is to determine the technological limits that need to be reached in order to unleash the SW computing paradigm full potential such that magnonic circuits can outperform CMOS counterparts in terms of energy consumption. In view of the zero power SWs propagation through ferromagnetic waveguides, the overall magnionic circuit power consumption is determined by the one associated to SWs generation and sensing by means of transducers. Thus we focus our analysis on determining transducer power consumption acceptable upper bounds that need to be achieve in future transducer implementations, e.g., ME cells. For this study, we choose a 32-bit Brent-Kung prefix adder (BKA)[25], as discussion vehicle and compute the maximum transducer power values that potentially enable a BKA SW implementation able to outperform its 7 nm CMOS counterpart. We note that the Brent-Kung adder is a Parallel Prefix Adder (PPA) form of the Carry-Look Ahead adder (CLA) that exhibits structure regularity, low wiring congestion, and reasonable area-performance ratio, which make it quite attractive for practical implementations[25]. To asses the representativity of our choice, we also determined transducer power upper bound

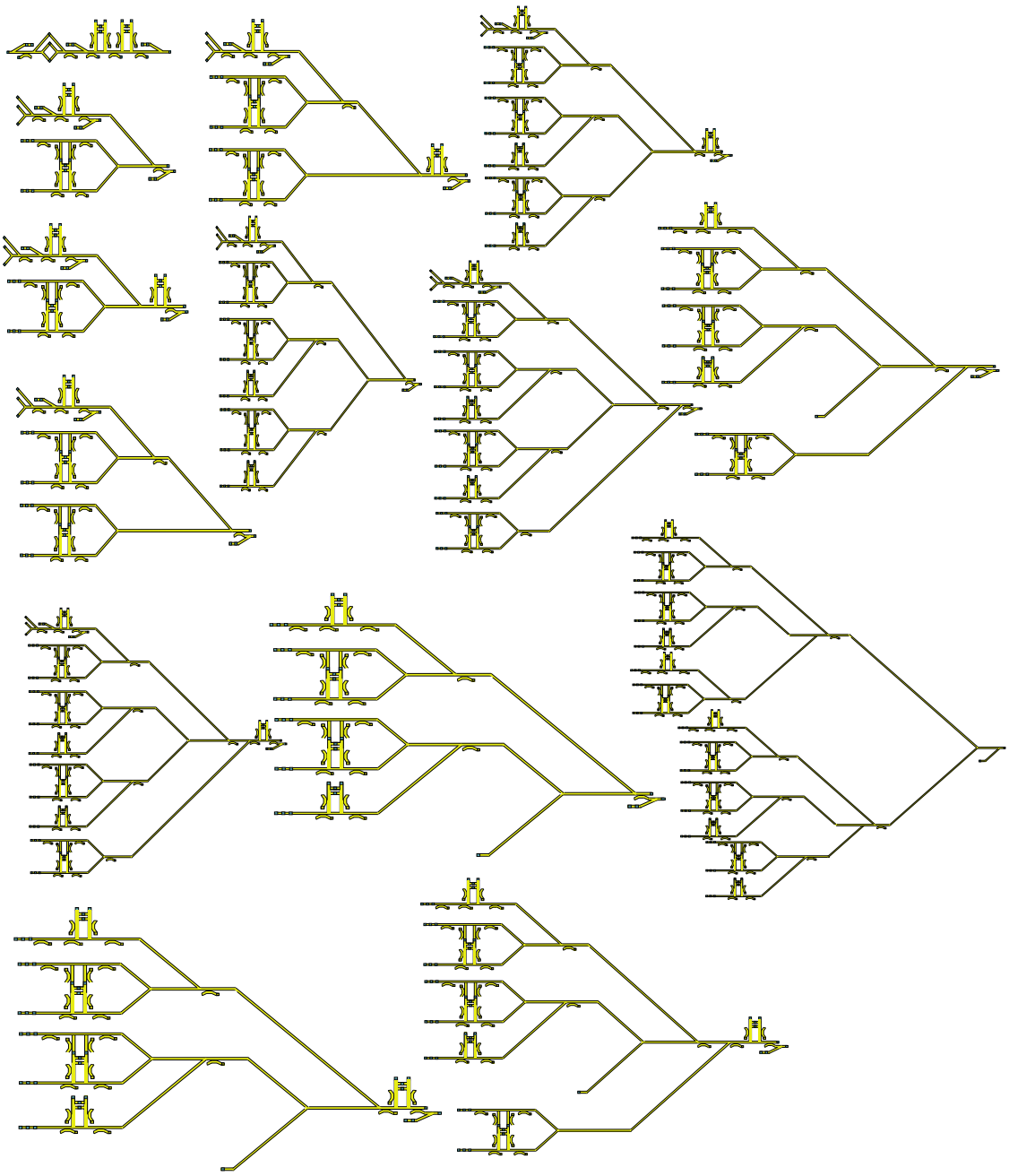

FIG. 3. 32-bit SW Brent Kung Prefix Adder Using Normalizers Only.

values for: 32-bit Wallace Tree Multiplier, 32-bit Dadda Tree Multiplier, 32-bit Brent Kung Adder, 64-bit Dadda Tree Multiplier, 4-operand 64-bit Han-Carlson adder, 4-operand 64-bit Carry Skip Adder, 32-bit Multiply Accumulate, 32-bit Divider, 17-bit Galois-Field Multiplier, and 32-bit Cyclic redundancy check. Our results indicate that Brent Kung Adder requires the lowest transducer upper bound (worst case), therefore, our choice as discussion vehicle is relevant for the purpose of this analysis.

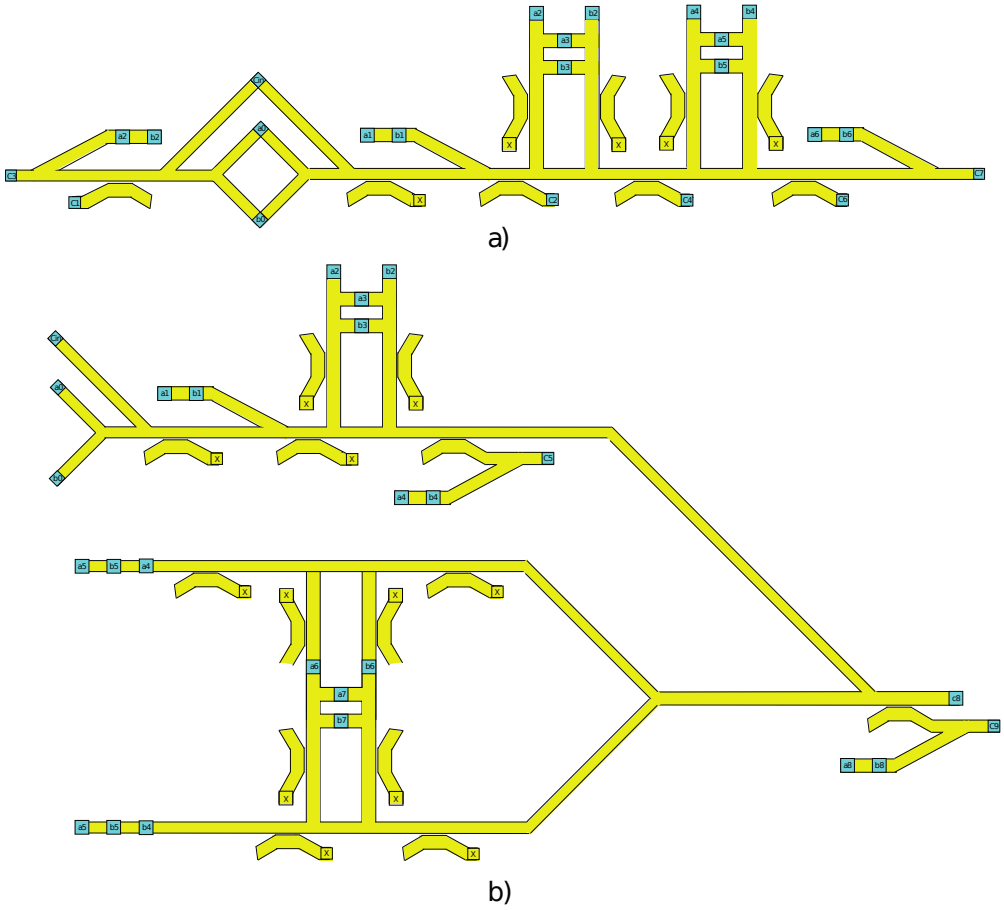

FIG. 4. a) Carry-out1 to Carry-out7 Calculation using Normalizers only, b) Carry-out8 to Carry-out9 Calculation using Normalizers only.

A. Possible Implementations

We evaluate different 32-bit BKA SW implementations based on Majority gates and compare them with the 7 nm CMOS implementation. Note that all our SW circuits rely on the majority tailored implementation method introduced in[26] for the 8-bit BKA case. As previously mentioned SW gate cascading is not straightforward[7] and to this end we evaluated 32-bit BKA implementations built with: (i) Ideal gate cascading (S1), (ii) Normalizers after each logic gate (S2), (iii) Normalizers and signal conversion back and forth between the electrical and spin wave domain (S3), and (iv) All-in-SW approach (S4). Note that in the implementations, we utilized a combination of fanout enabled ladder shaped Majority gate, programmable logic gates[9,11,13], triangle shape Majority gate[27], in-line Majority gate[28], and normalizers (directional couplers)[7].

Regardless of the gate cascading method the 32-bit BKA requires 98 transducers as it has

65 inputs and 33 outputs. This is the case for S1, which assumes direct SW gate cascading, i.e., no normalizers or signal conversion between electrical and SW domain are required to build the adder, and provides the best possible but practically unachievable adder performance. S2, which provides practically achievable performance data, makes use of directional couplers to normalize SW gate outputs. Figure 3 presents its structure, which contains 1040 transducers as a result of gate replication induced by: 1) unavailability of SW gates with larger than 4 fanout, 2) unavailability of SW splitters and amplifiers, and 3) layout limitations (waveguides crossovers are not allowed). For example, Figure 4 presents the SW circuit for calculating the carry-outs C1, C2, C3, ..., C9. As it can be observed from the Figure, C1 to C7 are calculated using the SW circuit in Figure 4 a), which requires 17 excitation transducers. On the other hand, C8 to C9 are detected using the SW circuit in Figure 4 b), which requires 23 excitation transducers, where 9 transducers are replicated because of fanout limitations. S3 diminishes the number of required replication and Figure 4 presents the SW circuit for calculating C1, C2, C3, ..., C9 by utilising normalizers and domain conversion (SW to/from electrical). This implementation requires a total of 43 transducers including excitation, intermediate, and detection transducers, whereas S2 implementation requires 49, thus we save 6 transducers for the calculation of the first 9 carry-outs.

However, as back and forth domain conversion cost is not yet available, the actual advantage of S3 cannot be accurately assessed. S4 implementation makes use of normalizers, splitters, amplifiers, and enables line crossover and its structure depicted in Figure 6 makes use of 98 transducers (65 excitation and 33 detection transducers), 72 splitters (directional couplers), and 72 amplifiers.

B. Transducer Power Upper Bound

To determine the transducer power consumption upper bound we first need to estimate the power and delay of our reference, i.e., CMOS 32-bit BKA. For this we utilize a commercial 7 nm FinFET technology, with regular threshold voltage standard cells, and typical process corner (V_{DD}=0.7 V, T=25°C). The adder was evaluated by means of Cadence simulation, which reported a power consumption of 2.58 μW and a delay of 1.033 ns that translates to an energy consumption of approximately 2.67 fJ for the 7 nm CMOS 32-bit BKA. In order to outperform the 7 nm CMOS BKA, the SW implementations should exhibit a maximum

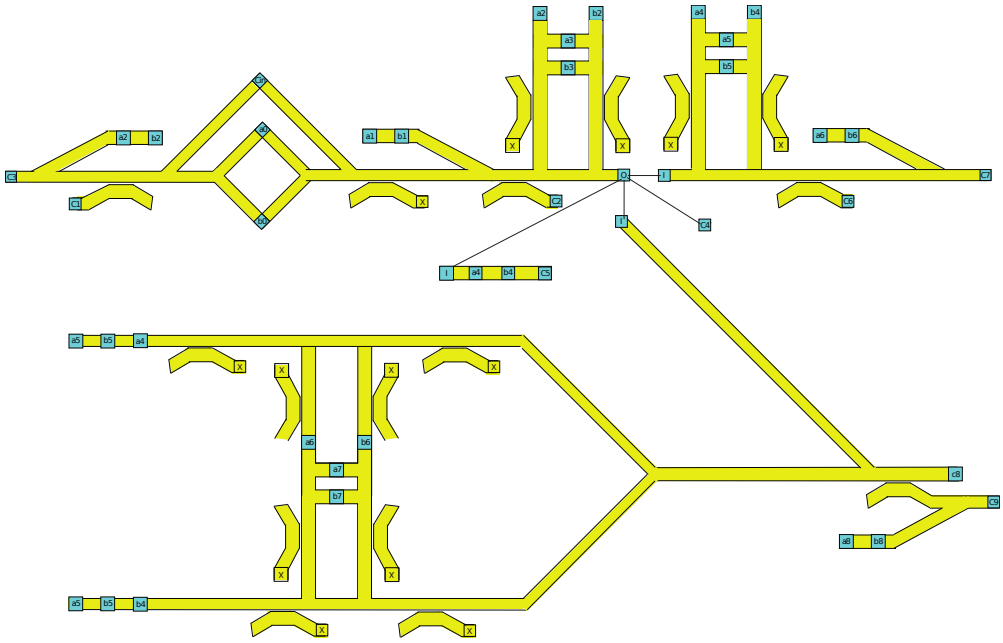

FIG. 5. Carry-out1 to Carry-out9 Calculation using Normalizers and Converters.

energy consumption $E_{SW} < E_{CMOS}$, and based on this, we can determine the performance constraint that the transducer needs to fulfil.

To evaluate the delay of a spin wave implementation, we have to identify its critical path, evaluate its physical length and determine the number of transducers it contains. Considering the ladder shape Majority (and their programmable logic gate version) gates[9,11,13] and assuming that the maximum propagation length per Majority gate is 336 nm, we can evaluate the length of the input SWs trajectory towards outputs in each implementation. First, we note that the critical path encompasses 16 Majority gates for S1, S2, S3, and S4. Based on this we derive the following critical path lengths: (i) 5.4 μm for S1, (ii) 50 μm for S2, (iii) 43 μm for S3, and (iv) 85 μm for S4. Although S3 has the shortest critical path length because it includes the least amount of directional couplers, it does not have the shortest delay because of the domain conversion circuitry. S4 has the longest critical path length because it make use of amplifiers and splitters to avoid transducer replications.

To derive the actual critical path delay, the SW propagation speed is required, which equals the SW group velocity that can be obtained from the dispersion relation material specific slope. Based on the critical path length and SW group velocity, we calculated the delay of the 4 implementations based on CoFeB waveguide as this material provides the highest SW group velocity. In addition, the following assumptions were made for the delay of the separate elements: 0.42 ns transducer delay[22], a 20 ns normalizer delay[7], a 0.03 ns

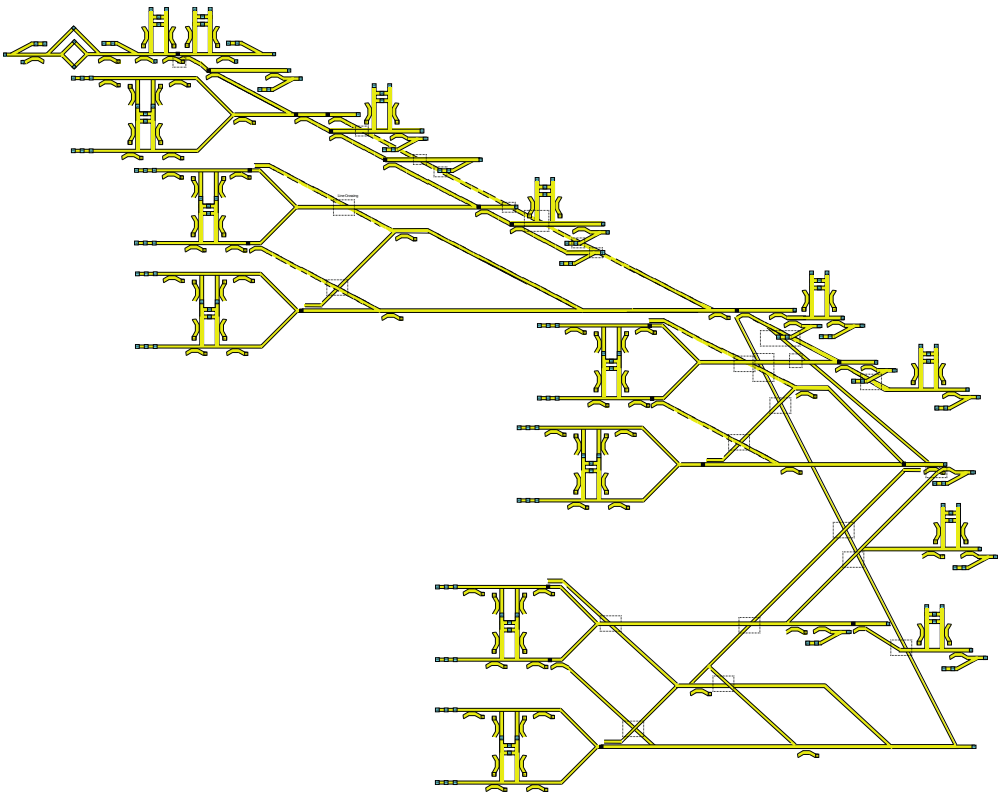

FIG. 6. 32-bit SW Brent Kung Prefix Adder Using Hybrid Approach.

peripheral circuit for the converters[22]. Based on this, we derive the following overall delays: (i) 1.92 ns for S1, (ii) 14.3 ns for S2, (iii) 20 ns for S3, and (iv) 24.3 ns for S4.

To proceed with the investigation on the SW adder energy consumption, we concentrate on power consumption estimation. Assuming zero power SW propagation through waveguides as SW doesn't require electron movement and just electron spinning, we can estimate the energy consumption as $E_{\mathrm{SW}} = TN \times PT \times \mathrm{Delay}$, where TN is the number of transducers in the implementation, PT the power consumed by one transducer, and Delay the time necessary to excite a SW. Given that in order to outperform CMOS $E_{SW} < E_{CMOS}$, the transducer power consumption upper bound can be determined as $PT = E_{CMOS}/(TN \times Delay)$. TN is determined by circuit topology and for each design we account one per primary input, one per primary output, and (if the case) the appropriate number of repeaters or converters necessary to interconnect the gates forming the prefix adder circuit, which results in 98, 1040, 262, and 203 transducers for S1, S2, S3, and S4, respectively. It was assumed that each amplifier consumes $\sqrt{n}$, where n is the amplification level.

The actual E_{SW} value is dependent on the SW operation mode, which defines the $Delay$ value in its evaluation expression. In Continuous Mode Operation (CMO)[29] the transducers

TABLE I. Transducer Power Upper Bound.

Implementation	Maximum Power (nW)	
	Continuous Mode Operation (CMO)	Pulse Mode Operation (PMO)
Ideal Case	17	64.9
Normalizer	0.18	6.1
Normalizer and Converter	0.51	24
Hybrid	0.54	31

are active as long as the SWs are propagating through the circuit, i.e., from SW excitation till the output detection. This means that *Delay* equals the overall adder delay, i.e., 20 ns, 12.3 ns, and 24.3 ns for S2, S3, and S4, respectively. In Pulse Mode Operation (PMO)[29], transducers are active only for a very short period of time required to initiate their output, which we assume to be 0.42 ns for all the implementations. Based on this reasoning, we determined the maximum allowable transducer power consumption PT for the CoFeB implementations under CMO and PMO scenarios as presented in Table I.

As one can observe in the Table, CMO puts a high pressure on the transducer performance whereas PMO relaxes it by 1-2 orders of magnitude. Moreover, regardless of the operation mode, the hybrid-based implementation is the most energy effective and allows for the highest PT value. Therefore, our preliminary evaluation indicates that the hybrid-based pulse mode operation approach potentially allows spin wave technology to outperform 7 nm CMOS, assuming that transducers with maximal 31 nW power consumption are achievable.

IV. CONCLUSIONS

In this paper, we assessed Magnonic circuits potential to outperform functionally equivalent CMOS counterparts in terms of energy consumption. We based our analysis on the fact that SW circuits energy consumption is determined by the energy spent by transducers to generate the input SWs and sense the output SWs, as SWs propagation through ferromagnetic waveguides do not consume noticeable energy. While it has been suggested that magneto-electric (ME) cells would be capable to excite and detect SWs while consuming ultra-low power, they have not been experimentally demonstrated and no figures of merit are available. Thus instead of performing a traditional benchmarking we carried on a re-

verse engineering investigation in an attempt de determine the ME power consumption upper bound that still make Magnonic circuits outperform CMOS counterparts. To this end, we assumed a 32-bit Brent-Kung prefix adder as discussion vehicle and determined the maximum transducer power consumption that still make the SW implementation outperform its 7 nm CMOS counterpart. We evaluated different SW implementations that rely on conversion- or normalization-based gate cascading and under continuous or pulse SW generation scenarios. Our evaluations indicated that 31 nW is the maximum transducer power consumption for which the 32-bit Brent-Kung SW implementation outperforms its 7 nm CMOS counterpart in term of energy.

ACKNOWLEDGEMENT

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No. 801055 "Spin Wave Computing for Ultimately-Scaled Hybrid Low-Power Electronics" – CHIRON. It has also been partially supported by imec's industrial affiliate program on beyond-CMOS logic. F.V. acknowledges financial support from the Research Foundation—Flanders (FWO) through grant No. 1S05719N.

REFERENCES

[1]N. D. Shah *et al.*, JAMA (2018).

[2]N. Z. Haron *et al.*, in *Design and Test Workshop, IDT* (IEEE, 2008) pp. 98–103.

[3]Y. Jiang *et al.*, IEEE Trans. Nanotechnology **18**, 287 (2019).

[4]H. A. D. Nguyen *et al.*, J. Emerg. Technol. Comput. Syst. **16** (2020), 10.1145/3365837.

[5]S. Agarwal *et al.*, "International roadmap of devices and systems 2017 edition: Beyond cmos chapter." Tech. Rep. (Sandia National Lab.(SNL-NM), United States, 2018).

[6]A. Mahmoud *et al.*, J. Appl. Phys. **128**, 161101 (2020).

[7]A. N. Mahmoud *et al.*, IEEE Transactions on Circuits and Systems I: Regular Papers **68**, 536 (2021).

[8]A. Mahmoud *et al.*, in *DATE* (2020) pp. 642–645.

[9]A. Mahmoud *et al.*, in *ISVLSI* (2020) pp. 60–65.

[10]M. P. Kostylev *et al.*, Appl. Phys. Lett. **87**, 153501 (2005).

[11]A. Mahmoud *et al.*, AIP Advances **10**, 035119 (2020).

[12]A. N. Mahmoud *et al.*, IEEE Transactions on Magnetics **57**, 1 (2021).

[13]A. Mahmoud *et al.*, in *2020 IEEE 38th International Conference on Computer Design (ICCD)* (2020) pp. 332–335.

[14]A. Khitun *et al.*, Journal of Applied Physics **110**, 034306 (2011).

[15]A. Mahmoud *et al.*, IEEE Transactions on Emerging Topics in Computing , 1 (2021).

[16]A. Mahmoud *et al.*, in *2021 28th IEEE International Conference on Electronics, Circuits, and Systems (ICECS)* (2021) pp. 1–4.

[17]A. N. Mahmoud *et al.*, IEEE Nanotechnology Magazine **16**, 47 (2022).

[18]F. Gertz *et al.*, IEEE Trans. Magn. **51**, 1 (2015).

[19]D. E. Nikonov and I. A. Young, Proceedings of the IEEE **101**, 2498 (2013).

[20]C. Pan and A. Naeemi, IEEE J. Explor. Solid-State Comput. Devices Circuits **3**, 101 (2017).

[21]D. E. Nikonov and I. A. Young, IEEE J. Explor. Solid-State Comput. Devices Circuits **1**, 3 (2015).

[22]O. Zografos *et al.*, in *IEEE-NANO* (2015) pp. 686–689.

[23]A. V. Chumak *et al.*, Journal of Physics D: Applied Physics **50**, 244001 (2017).

[24]F. Vanderveken *et al.*, J. Phys. D: Applied Physics **53**, 495006 (2020).

[25]M. Naik V, International Journal of Electrical Electronics and Data Communication **3** (2015).

[26]V. Pudi *et al.*, IEEE Transactions on Computers **66**, 1824 (2017).

[27]A. Mahmoud *et al.*, in *2021 Design, Automation Test in Europe Conference Exhibition (DATE)* (2021) pp. 948–953.

[28]G. Talmelli *et al.*, Science Advances **6** (2020).

[29]A. Mahmoud *et al.*, in *2021 22nd International Symposium on Quality Electronic Design (ISQED)* (2021) pp. 54–59.