Applied Physics Letters ARTICLE scitation.org/journal/apl

Reservoir computing with the frequency, phase, and amplitude of spin-torque nano-oscillators

Cite as: Appl. Phys. Lett. **114**, 012409 (2019); doi: 10.1063/1.5079305
Submitted: 30 October 2018 · Accepted: 26 December 2018 · Published Online: 10 January 2019

View Online Export Citation CrossMark

D. Marković,[1,a)] N. Leroux,[1] M. Riou,[1] F. Abreu Araujo,[2] J. Torrejon,[1] D. Querlioz,[3] A. Fukushima,[4] S. Yuasa,[4] J. Trastoy,[1] P. Bortolotti,[1] and J. Grollier[1]

AFFILIATIONS

[1] Unité Mixte de Physique, CNRS, Thales, University Paris-Sud, Université Paris-Saclay, 91767 Palaiseau, France
[2] Institute of Condensed Matter and Nanosciences, UCLouvain, Place Croix du Sud 1, 1348 Louvain-la-Neuve, Belgium
[3] Centre de Nanosciences et de Nanotechnologies, CNRS, University Paris-Sud, Université Paris-Saclay, 91405 Orsay, France
[4] National Institute of Advanced Industrial Science and Technology (AIST), Spintronics Research Center, Tsukuba, Ibaraki 305-8568, Japan

[a)] Author to whom correspondence should be addressed: danijela.markovic@cnrs-thales.fr

ABSTRACT

Spin-torque nano-oscillators can emulate neurons at the nanoscale. Recent works show that the non-linearity of their oscillation amplitude can be leveraged to achieve waveform classification for an input signal encoded in the amplitude of the input voltage. Here, we show that the frequency and phase of the oscillator can also be used to recognize waveforms. For this purpose, we phase-lock the oscillator to the input waveform, which carries information in its modulated frequency. In this way, we considerably decrease the amplitude, phase, and frequency noise. We show that this method allows classifying sine and square waveforms with an accuracy above 99% when decoding the output from the oscillator amplitude, phase, or frequency. We find that recognition rates are directly related to the noise and non-linearity of each variable. These results prove that spin-torque nano-oscillators offer an interesting platform to implement different computing schemes leveraging their rich dynamical features.

Spin-torque nano-oscillators are promising for neuromorphic computing.[1–5] These magnetic tunnel junctions can indeed emulate important properties of artificial neurons through the non-linearity and relaxation properties of current-induced magnetization dynamics. It has been shown recently that a time-multiplexed, single oscillating junction can enable or improve classification of different waveforms, distinguishing sines from squares and even spoken digits.[6,7] In these experiments, the input waveform was encoded in the amplitude of the input voltage and the quantity used for computing was the amplitude of voltage oscillations across the junction. Other dynamical variables are interesting to leverage for computing, such as the frequency[8] or phase[9] of the oscillators, offering a compelling platform to implement and compare different neuromorphic computing approaches. The variable used for computing should vary non-linearly with the input in a way that is easy to detect, and the signal to noise ratio should be large. However, the frequency and phase of spin-torque nano-oscillators tend to be highly noisy,[10–12] which has been shown to be detrimental to pattern classification.[7] Spin-torque induced magnetization dynamics indeed takes place in nanoscale magnetic volumes, which makes them sensitive to thermal fluctuations. In addition, phase noise is enhanced by amplitude noise due to the inherent coupling between the phase and the amplitude of magnetization oscillations.[13] In this work, we show that these issues can be circumvented by working in a regime where the oscillator is synchronized to the input waveform that it has to process which considerably reduces magnetization fluctuations.[14] For this purpose, we use a sinusoidal input waveform that carries information encoded in its modulated frequency, chosen close to the spin-torque oscillator frequency.

We first explain in detail our computing strategy and describe the experimental set-up used to implement it. We then show experimentally that sine and square waveforms can be classified by exploiting the frequency, phase, or amplitude of the oscillations. We highlight the correlation between recognition rates and the non-linearity of these dynamical variables as a

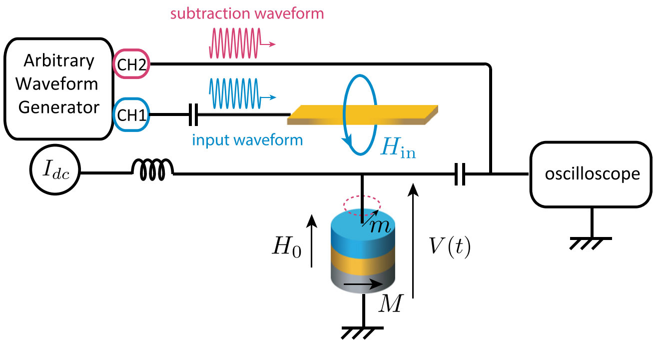

FIG. 1. Schematic of the measurement set-up. The spin-torque nano-oscillator is composed of two magnetic layers, of fixed magnetization M (gray) and free magnetization m (blue), separated by a thin insulating layer. At an external magnetic field of $H_0 = 2000$ Oe, a direct current of $I_{dc} = 5$ mA is injected in order to induce magnetization precessions. The microwave signal encoding the input data in its frequency (blue) is injected into a strip line above the oscillator, thus generating a microwave magnetic field interacting with the free layer. The microwave voltage $V(t)$ emitted by the oscillator is added to a microwave signal (subtraction waveform) that compensates for the residual input signal and then is measured using an oscilloscope.

function of the input signal to classify. Our work shows that it is possible to take full advantage of magnetization dynamics by computing through all the dynamical variables describing a spin-torque nano-oscillator. In addition, since the input waveform and the oscillator output are sinusoidal waveforms with close frequencies, our scheme should allow chain-connecting the oscillators to build large neural networks.

Spin-torque nano-oscillators[15,16] are composed of two ferromagnets separated by a thin non-magnetic layer. The magnetization of the bottom ferromagnet is pinned, whereas that of the top one is free. The spin-torque oscillator used in this experiment is a nano-pillar of 350 nm diameter and composed of a 1.6 nm thick CoFeB layer with a pinned magnetization, a 1 nm thick MgO insulating barrier, and a 4 nm thick FeB layer whose ground state is a magnetic vortex. Such nano-pillars can be fabricated with diameters down to 10 nm,[17] which is adapted for building large scale neural networks. When a direct current is injected into this magnetic tunnel junction in the presence of a magnetic field perpendicular to the layer stack, it induces magnetization precessions in the free layer through the effect of spin torque. Magnetoresistance effects convert magnetic oscillations into resistance oscillations, such that a microwave voltage is emitted by the oscillator and can be detected using an oscilloscope. The experimental set-up is shown in Fig. 1.

Spin-torque nano-oscillators have the ability to synchronize their voltage oscillations to external microwave signals at frequencies close to their natural frequency.[18–21] The frequency and phase of the oscillator lock to the external signal, and its amplitude is modified. Importantly, noise is reduced in frequency and phase so that these variables are well defined in this regime. We choose to work in this regime where the input signal synchronizes the oscillator in order to reduce the noise. The range of input microwave frequencies that synchronize the oscillator is called the injection locking range. In the following, we also take advantage of the frequency pulling regime, by setting the frequency of the input signal just outside of the locking range, such that the oscillator does not get phase locked, but its frequency gets pulled towards the frequency of the input signal.

We apply a perpendicular magnetic field $H = 2000$ Oe to the oscillator and inject a direct current $I_{dc} = 5$ mA, which induces voltage oscillations of amplitude $|V| = 13$ mV at a frequency of 232.1 MHz. The corresponding emitted power and linewidth are close to 1 μW and 1 MHz, respectively. We choose these field and current bias parameters in order to have a large locking range which is important for the frequency encoding and to minimize the linewidth and maximize the output signal. We use an arbitrary waveform generator (AWG) to generate input microwave signals. These are injected in a strip line patterned 350 nm above the oscillator rather than in the oscillator itself in order to facilitate the extraction of the oscillator response from the overall measured signal. The injected signal induces a microwave magnetic field on the oscillator, as well as a microwave current in the oscillator due to capacitive coupling with the strip line. In order to synchronize the oscillator, amplitudes of $\approx$350 mV of the injected signal need to be applied, such that the total voltage detected by the oscilloscope is dominated by a residual capacitive microwave tone rather than the oscillator voltage. We compensate for this residual tone by adding the output voltage in a power combiner to an exactly opposed microwave signal waveform (subtraction signal in Fig. 1) delayed by the time t_0 that the input signal takes to travel through the lines and that we calibrate prior to the measurement.

In order to characterize the synchronization of the spin-torque oscillator with an external source, we send 5 μs long waveforms modulated at different frequencies in a 20 MHz range

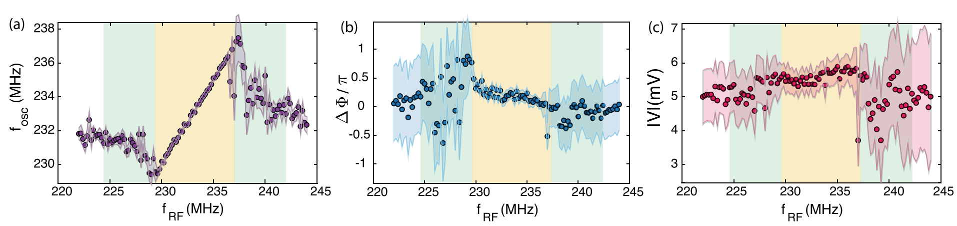

FIG. 2. (a) Frequency f_{osc}, (b) phase $\Delta\Phi$, and (c) amplitude $|V|$ of the oscillator as a function of the frequency f_{RF} of the injected microwave signal. The phase is determined with respect to that of the input waveform. Measurement uncertainties, determined on 5 μs time intervals on which the mean is calculated, are shown in the lighter color shaded area. Yellow and green shaded areas designate the locking range and the frequency pulling range, respectively.

within the natural frequency of the oscillator. We apply the Hilbert transform[22,23] to the detected voltage in order to extract frequency, amplitude, and phase relative to that of the injected microwave signal, which we average over the entire 5 μs waveform. The oscillator frequency, phase, and amplitude as a function of the frequency of the injected microwave signal are shown in Figs. 2(a)–2(c). As the injection signal frequency approaches the natural oscillator frequency, the oscillator frequency first gets pulled towards the injected signal and then becomes identical to it in the locking range. Noise is reduced in all three variables in the locking range. Due to the subtraction of the residual microwave signal performed using a power combiner, the detected amplitude of the oscillator voltage is divided by two. This results in a low signal-to-noise ratio even in the locking range [note large error bars in Fig. 2(c)]. The locking range, highlighted in yellow in Fig. 2, is experimentally determined from the standard deviation of the phase that strongly decreases in this range and is found to be 7.2 MHz. As expected, the measured frequency of the oscillator is equal to the injected frequency in the locking range [Fig. 2(a)]. The phase difference between the oscillator and the input signal roughly follows the arcsin dependence on the input frequency predicted by theory[13] [Fig. 2(b)].

An oscillator can only achieve good performance in neuromorphic computing if it transforms the input signal in a non-linear manner.[7,24–26] In the pulling regime (green in Fig. 2), the oscillator frequency, phase, and amplitude are all highly non-linear. The oscillator frequency is linear over the entire locking range, whereas the phase difference and the oscillator amplitude are non-linear at the edges [Figs. 2(b) and 2(c)].

The fact that the frequency, amplitude, and phase are all non-linear functions of input frequency enables us to use them for neuromorphic computing. We now demonstrate this capability on a task that consists in classification of sine and square waves of equal periods but different amplitudes. For this, we use a method called single node reservoir computing.[6,7,24,25] This method uses time multiplexing in order to emulate a reservoir with a single nano-oscillator that plays a role of a different effective virtual neuron at each time step.

The input data encoding procedure is shown in Fig. 3. The input data are a sequence of 100 waveforms randomly chosen between sines and squares of equal frequencies and different amplitudes: the amplitude of the square wave is 50% larger than that of the sine wave. Half of these data are used for training and the other half for testing the performance. Each waveform is discretized into 8 points [see Fig. 3(a)], and the task consists in determining which of the two waveform types each point belongs to. This is a non-linearly separable task and thus represents a good benchmark for neuromorphic computing.[6,7,25]

Time multiplexing is achieved by preprocessing the input data as illustrated in Fig. 3(b). The detailed procedure can be found in previous work.[7] Each input point is multiplied by the same binary valued sequence called mask, whose length N determines the size of the emulated reservoir. Figure 3(b) shows an illustrative schematic for a reservoir containing N = 6 neurons. In our experiment, we have used N = 25 virtual neurons which we found was the minimum number of neurons required for good performance. The output of the neural network for each input time step is a weighted sum of the outputs of each virtual neuron corresponding to this input

$$y = \sum_{i=1}^{N} W_i f_{NL}(x_i), \tag{1}$$

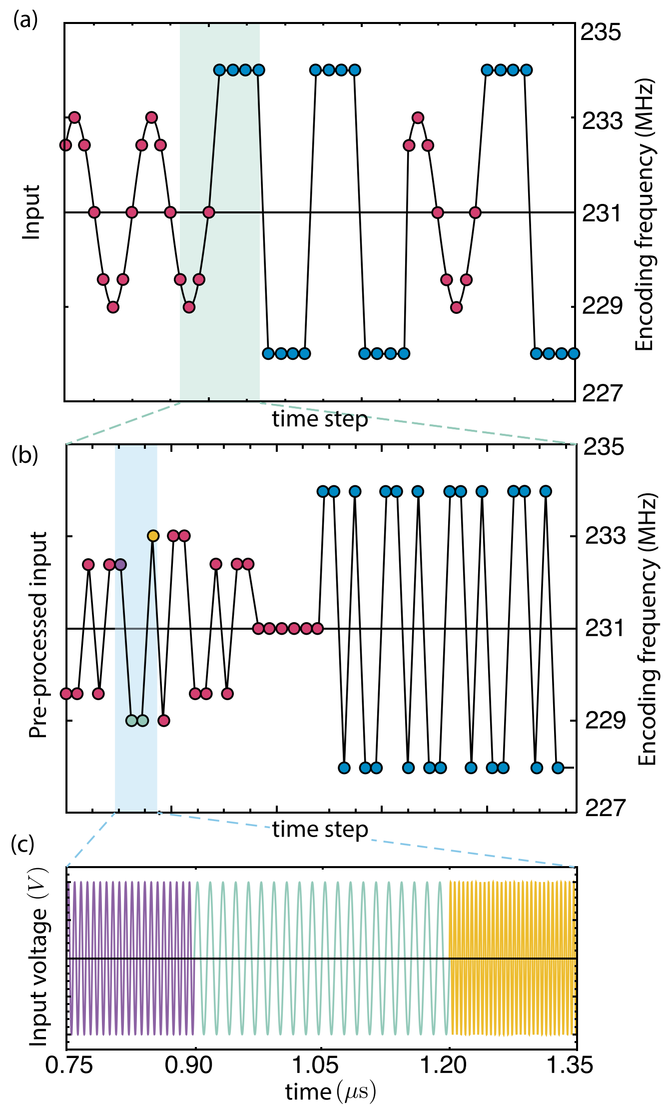

FIG. 3. (a) The input data are a sequence of random sine and square waves of equal periods and different amplitudes discretized in 8 points. (b) Pre-processed data corresponding to half of a sine wave followed by half of a square one. In this example, the mask maps the problem to six virtual neurons. The Y axis corresponds to one example of encoding frequencies. (c) Sketch of the input voltage corresponding to four neuron entries for a sine wave. Different input values are presented in different colors. The waveform amplitudes are encoded in the frequency of the microwave voltage which is then injected into the strip line for 150 ns for each data point.

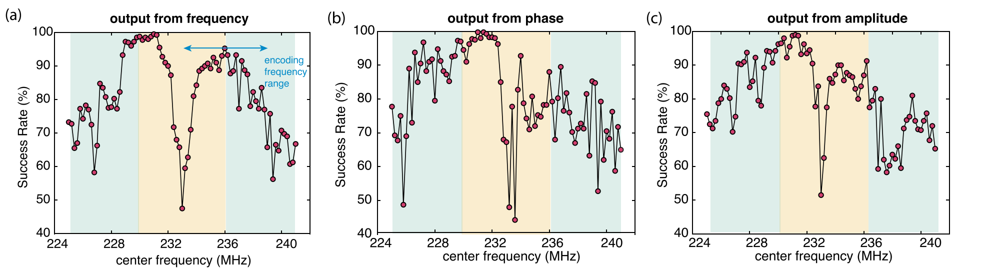

FIG. 4. Success rates obtained when decoding from the frequency, phase, and amplitude of the oscillator, as a function of the center of the frequency range chosen for encoding the input data. The frequency range used for encoding is indicated by a blue double arrow for two measurement points. Yellow and green shaded areas designate the locking range and the frequency pulling range, respectively.

where N = 25 is the number of neurons, W_i is the weight matrix element corresponding to the ith neuron, f_{NL} is the non-linear function implemented by the nanodevice, and x_i is the input of the ith neuron, which is the corresponding microwave frequency. The weight matrix is calculated on a computer in order to match the target $\tilde{y} = 0$ or 1 for sines or squares, respectively. For a target vector $\tilde{Y}$ containing targets $\tilde{y}$ for all the training examples, the weight matrix is calculated as $W = \tilde{Y}F^{\dagger}$, where $F^{\dagger}$ is the Moore-Penrose pseudo-inverse of matrix F containing outputs $f_{NL}(x_i)$ of all neurons and for all training examples.[24]

Classification performance is highly dependent on the frequency window chosen for input data encoding. We choose this window in a partial or full overlap with the oscillator locking range. We fix the window width such that sine and square waves always take values between 4 MHz and 6 MHz, respectively. We repeat the encoding and measurement procedures for center frequencies of the encoding window varying between 225 MHz and 241 MHz, and we calculate the success rate. Recognition rates obtained when decoding neuron outputs from the frequency, phase, and amplitude are shown in Figs. 4(a)–4(c) as a function of the center frequency of the sliding window.

Changing the center frequency has a double impact on output variables, which is the presence of noise, and the non-linear dependence on the input frequency. Noise is minimized in the middle of the locking range, but the output in this regime is a linear function of the input (see Fig. 2), which results in a disability of the neural network to learn the task. Indeed, as can be seen in Figs. 4(a)–4(c), the success rate for the frequency encoding window centered in the middle of the locking range is close to 50% for all the three output variables, which for this task corresponds to random choice. The linear regime is larger for frequency than for the amplitude and phase, which is reflected in the bad performance for a larger number of center frequencies in the middle of the locking range.

We find the best performance for the center frequency on the edge of the locking range, with some of the frequencies used for encoding lying in the highly non-linear frequency pulling regime. The best recognition rates are obtained when neuron outputs are decoded from the phase of the oscillations [99.75%, Fig. 4(b)], as the phase is both more non-linear than frequency [best recognition rate of 99.5%, Fig. 4(a)] and less noisy than the amplitude [best recognition rate of 99%, Fig. 4(c)]. In addition, higher recognition rates are obtained on the left-hand side of the locking range compared to the right-hand side due to lower frequency and amplitude noise on this side [see Figs. 2(a) and 2(c)] as well as higher phase non-linearity [see Fig. 2(b)].

These high classification rates have been obtained by using relatively large input microwave amplitudes to drive the oscillator and reduce its noise. In this regime, the magnetization relaxation time, which decreases with the excitation amplitude in the locking range,[27] is very short, smaller than 4 ns in our case. Therefore, the emulated neural network performs best at tasks that do not require a memory, such as the classification of different inputs. When the waveforms to separate have identical input values that can only be recognized by keeping memory of past inputs, as is the case for sine and square waves with the same amplitude, the network performance is lower (82% recognition rate at maximum). In the future, it will be interesting to study the network intrinsic memory as a function of drive amplitude and oscillator noise. In addition, an external memory can be added to the system by using a time-delayed feedback loop and re-injecting the signal emitted by the oscillator together with the input data.[24–26]

As a conclusion, we have shown that spin-torque nano-oscillators synchronized to microwave signals can emulate artificial neural networks. The frequency, phase, and amplitude of the voltage emitted by the oscillator are all non-linear functions of the frequency of the input microwave signal and can be used as outputs of the network. Working with synchronized neurons has the advantage of decreasing the frequency and phase noise, which will be of particular importance when scaling down the size of nano-pillars. In addition, frequency encoding is a simple way to use the output of an oscillator to drive the input of another, thus paving the path for neural networks composed of chain-connected spin-torque or spin-hall nano-oscillators.[20,28–30]

This work was supported by the European Research Council ERC under Grant bioSPINspired 682955 and the French Ministry of Defense (DGA).

REFERENCES

[1] F. Macià, A. D. Kent, and F. C. Hoppensteadt, Nanotechnology **22**, 095301 (2011).
[2] M. R. Pufall, W. H. Rippard, G. Csaba, D. E. Nikonov, G. I. Bourianoff, and W. Porod, IEEE J. Explor. Solid-State Comput. Devices Circuits **1**, 76 (2015).
[3] D. E. Nikonov, G. Csaba, W. Porod, T. Shibata, D. Voils, D. Hammerstrom, I. A. Young, and G. I. Bourianoff, IEEE J. Explor. Solid-State Comput. Devices Circuits **1**, 85 (2015).

[4]K. Yogendra, D. Fan, and K. Roy, IEEE Trans. Magn. **51**, 4003909 (2015).
[5]J. Grollier, D. Querlioz, and M. D. Stiles, Proc. IEEE **104**, 2024 (2016).
[6]M. Riou, F. A. Araujo, J. Torrejon, S. Tsunegi, G. Khalsa, D. Querlioz, P. Bortolotti, V. Cros, K. Yakushiji, A. Fukushima, H. Kubota, S. Yuasa, M. D. Stiles, and J. Grollier, in 2017 *IEEE International Electron Devices Meeting (IEDM)* (2017), pp. 36.3.1–36.3.4.
[7]J. Torrejon, M. Riou, F. A. Araujo, S. Tsunegi, G. Khalsa, D. Querlioz, P. Bortolotti, V. Cros, K. Yakushiji, A. Fukushima, H. Kubota, S. Yuasa, M. D. Stiles, and J. Grollier, Nature **547**, 428 (2017).
[8]E. Vassilieva, G. Pinto, J. A. De Barros, and P. Suppes, IEEE Trans. Neural Networks **22**, 84 (2011).
[9]F. C. Hoppensteadt and E. M. Izhikevich, Phys. Rev. Lett. **82**, 2983 (1999).
[10]M. Quinsat, D. Gusakova, J. F. Sierra, J. P. Michel, D. Houssameddine, B. Delaet, M. C. Cyrille, U. Ebels, B. Dieny, L. D. Buda-Prejbeanu, J. A. Katine, D. Mauri, A. Zeltser, M. Prigent, J. C. Nallatamby, and R. Sommet, Appl. Phys. Lett. **97**, 182507 (2010).
[11]M. W. Keller, M. R. Pufall, W. H. Rippard, and T. J. Silva, Phys. Rev. B: Condens. Matter Mater. Phys. **82**, 054416 (2010).
[12]E. Grimaldi, A. Dussaux, P. Bortolotti, J. Grollier, G. Pillet, A. Fukushima, H. Kubota, K. Yakushiji, S. Yuasa, and V. Cros, Phys. Rev. B: Condens. Matter Mater. Phys. **89**, 104404 (2014).
[13]A. Slavin and V. Tiberkevich, IEEE Trans. Magn. **45**, 1875 (2009).
[14]R. Lebrun, A. Jenkins, A. Dussaux, N. Locatelli, S. Tsunegi, E. Grimaldi, H. Kubota, P. Bortolotti, K. Yakushiji, J. Grollier, A. Fukushima, S. Yuasa, and V. Cros, Phys. Rev. Lett. **115**, 017201 (2015).
[15]S. I. Kiselev, J. C. Sankey, I. N. Krivorotov, N. C. Emley, R. J. Schoelkopf, R. A. Buhrman, and D. C. Ralph, Nature **425**, 380 (2003).
[16]W. H. Rippard, M. R. Pufall, S. Kaka, S. E. Russek, and T. J. Silva, Phys. Rev. Lett. **92**, 027201 (2004).
[17]H. Sato, E. C. Enobio, M. Yamanouchi, S. Ikeda, S. Fukami, S. Kanai, F. Matsukura, and H. Ohno, Appl. Phys. Lett. **105**, 062403 (2014).
[18]W. H. Rippard, M. R. Pufall, S. Kaka, T. J. Silva, S. E. Russek, and J. A. Katine, Phys. Rev. Lett. **95**, 067203 (2005).
[19]M. Quinsat, J. F. Sierra, I. Firastrau, V. Tiberkevich, A. Slavin, D. Gusakova, L. D. Buda-Prejbeanu, M. Zarudniev, J. P. Michel, U. Ebels, B. Dieny, M. C. Cyrille, J. A. Katine, D. Mauri, and A. Zeltser, Appl. Phys. Lett. **98**, 182503 (2011).
[20]A. Houshang, E. Iacocca, P. Dürrenfeld, S. Sani, J. Akerman, and R. Dumas, Nat. Nanotechnol. **11**, 280 (2016).
[21]M. Romera, P. Talatchian, S. Tsunegi, F. A. Araujo, V. Cros, P. Bortolotti, J. Trastoy, K. Yakushiji, A. Fukushima, H. Kubota, S. Yuasa, M. Ernoult, D. Vodenicarevic, T. Hirtzlin, N. Locatelli, D. Querlioz, and J. Grollier, Nature **563**, 230 (2018).
[22]B. Picinbono, IEEE Trans. Signal Process. **45**, 552 (1997).
[23]L. Bianchini, S. Cornelissen, J. V. Kim, T. Devolder, W. Van Roy, L. Lagae, and C. Chappert, Appl. Phys. Lett. **97**, 032502 (2010).
[24]L. Appeltant, M. C. Soriano, G. Van Der Sande, J. Danckaert, S. Massar, J. Dambre, B. Schrauwen, C. R. Mirasso, and I. Fischer, Nat. Commun. **2**, 468 (2011).
[25]Y. Paquot, F. Duport, A. Smerieri, J. Dambre, B. Schrauwen, M. Haelterman, and S. Massar, Sci. Rep. **2**, 287 (2012).
[26]L. Larger, A. Baylón-Fuentes, R. Martinenghi, V. S. Udaltsov, Y. K. Chembo, and M. Jacquot, Phys. Rev. X **7**, 011015 (2017).
[27]W. Rippard, M. Pufall, and A. Kos, Appl. Phys. Lett. **103**, 182403 (2013).
[28]V. E. Demidov, S. Urazhdin, A. Zholud, A. V. Sadovnikov, and S. O. Demokritov, Appl. Phys. Lett. **105**, 172410 (2014).
[29]A. A. Awad, P. Dürrenfeld, A. Houshang, M. Dvornik, E. Iacocca, R. K. Dumas, and J. Åkerman, Nat. Phys. **13**, 292 (2017).
[30]M. Zahedinejad, H. Mazraati, H. Fulara, J. Yue, S. Jiang, A. A. Awad, and J. Åkerman, Appl. Phys. Lett. **112**, 132404 (2018).