Generating Magnetic Skyrmion Ground States with Generative Adversarial Networks

Bozhidar Trenchev
School of Electronics and Computer Science
University of Southampton
University Road, SO17 1BJ, UK
btrenchev1@gmail.com

Srinandan Dasmahapatra
School of Electronics and Computer Science
University of Southampton
University Road, SO17 1BJ, UK
sd@ecs.soton.ac.uk

Marijan Beg
Faculty of Engineering and Physical Sciences
University of Southampton
University Road, SO17 1BJ, UK
M.Beg@soton.ac.uk

Natalie Downing
Faculty of Engineering and Physical Sciences
University of Southampton
University Road, SO17 1BJ, UK
N.Downing@soton.ac.uk

Ondrej Hovorka
Faculty of Engineering and Physical Sciences
University of Southampton
University Road, SO17 1BJ, UK
O.Hovorka@soton.ac.uk

Abstract

Magnetic skyrmions are an exciting new area of magnetism research. They are nanosized quasi-particles with a vortex-like magnetization. They are expected to be introduced in new-generation magneto-electronics due to their small size and low energy consumption. The magnetic samples hosting skyrmions can have many different equilibrium magnetization states and it is thus of great importance to identify as many diverse equilibrium states as possible. Existing methods include exhaustive search by relaxing random initial states, but this is a very time consuming and inefficient process. In this work, we investigate the use of generative adversarial networks for finding new physically-realistic equilibrium states. The advantage is that once the network is configured and trained, it can generate many diverse relaxed magnetization fields with skyrmions very fast. Furthermore, several variations of this generative model are proposed that incorporate some of the micromagnetic physical laws into their architectures via physics loss functions.

1 Introduction

Skyrmions are topologically protected non-trivial field configurations (1) that have been predicted (2) and later observed in chiral magnetic materials (3). The topological stability of these nanometer-scale configurations and their coherent dynamics (4) have made them attractive candidates for both data-storage and information-processing devices such as race-track-like memories (5) and reservoir computing (6).

The key for the emergence of skyrmionic states is the presence of the Dzyaloshinskii-Moriya interaction (7; 8). Skyrmionic states are explored numerically by solving partial differential equations within the framework of micromagnetics or via Monte Carlo methods, where an initial magnetisation

Third Workshop on Machine Learning and the Physical Sciences (NeurIPS 2020), Vancouver, Canada.

configuration is relaxed to minimise the total magnetic energy. The computational exploration of parameter space with random initial configurations is far from exhaustive and it is computationally expensive. For the purposes of this paper, we chose bulk FeGe as the skyrmion-hosting material because of its near-room ordering temperature (9). We then seek to generate viable equilibrium configurations consistent with the micromagnetic model of FeGe using a Generative Adversarial Network (GAN) (10). Our system, implemented in TensorFlow (11) using Keras (12) and called PhysGAN, can generate many new field configurations resembling real physical states and are at a local energy minimum. Skyrmions have been studied using fully connected neural networks (13), but here we generate configurations to tailor diverse samples similar to given configurations.

2 PhysGAN Design

PhysGAN is a novel neural network, based on the popular Generative Adversarial Network (GAN) design by Goodfellow et al. (10) and its deep convolutional version (DCGAN) (14). The GAN's architecture includes two neural networks - the discriminator $\mathbf{D}$ and the generator $\mathbf{G}$. The generator takes as an input a vector drawn from a multivariate normal distribution and generates fake 2D images – the textured field configurations with 3 magnetisation components represented as 3 channels. The discriminator is a standard convolutional network trained on both real training data and on generated data. Its job is to map 2D input to a probability between 0 and 1 of being either fake of real. Those two networks are interlocked in a minimax game where the generator learns to create more realistic images in order to make the discriminator unable to distinguish between real and fake ones. They play this game until convergence and the outcome is that the generator generates more realistic data. Building on DCGAN, in PhysGAN (high-level architecture shown in Fig. 1) we introduce domain-specific "physics-aware" losses, a new training strategy, and new generator and discriminator architectures. The main modifications of the generator are the inclusion of batch norm layers (15) between its 2D convolutional layers (16) and leaky ReLU activations (17). The discriminator's initial convolutional layers have larger, 5x5 kernels to capture spatial spin correlations, followed by 3x3 kernels. We also randomly dropout 30% of the neurons for regularization (18). The discriminator learns to assign an unbounded negative number to fake images and a positive number to real images.

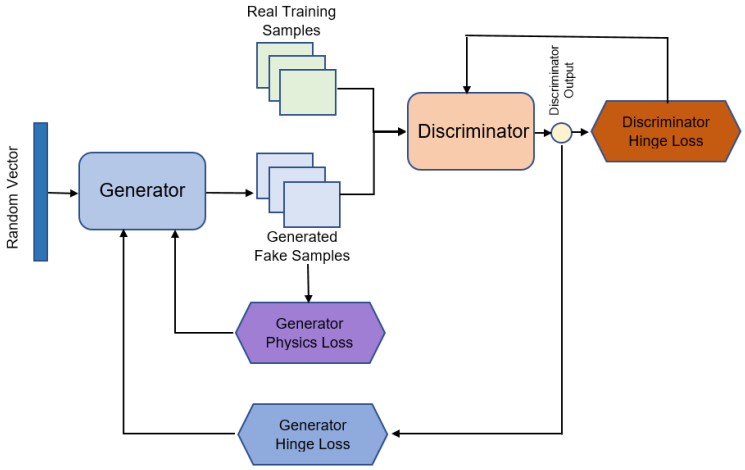

Figure 1: Diagram summarizing the PhysGAN architecture

2.1 Computational micromagnetics

The magnetization field $\mathbf{M}(\mathbf{r}) = M_s\hat{\mathbf{m}}(\mathbf{r}) \in \mathbb{R}^2$ (a hat denotes a unit length vector) is computed using a finite-difference method where $\mathbf{r}$ is restricted to a 2D $L_x \times L_y$ lattice grid. This way, the PhysGAN sees a magnetization as a tensor of shape [L_x, L_y, 3] with (L_x, L_y) either $(48, 32)$ or $(144, 96)$. Every discretized cell in the grid has width/height of $3\,\mathrm{nm}$. We compute the total magnetic energy as the sum of individual energy terms

$$E[\widehat{\mathbf{m}}] = \int \left[w_{\mathrm{ex}}(\widehat{\mathbf{m}}) + w_{\mathrm{dmi}}(\widehat{\mathbf{m}}) + w_{\mathrm{z}}(\widehat{\mathbf{m}}) \right] \mathrm{d}^2\mathbf{r}, \tag{1}$$

where $w_{\mathrm{ex}}(\widehat{\mathbf{m}}) = -A\widehat{\mathbf{m}} \cdot (\nabla^2 \widehat{\mathbf{m}})$, $w_{\mathrm{dmi}}(\widehat{\mathbf{m}}) = D\widehat{\mathbf{m}} \cdot (\nabla \times \widehat{\mathbf{m}})$, and $w_{\mathrm{z}}(\widehat{\mathbf{m}}) = \mu_0 M_{\mathrm{s}} \widehat{\mathbf{m}} \cdot \mathbf{H}$ are the exchange, DMI, and Zeeman energy densities, respectively. For material parameters M_{s}, A, and D we choose the values to model FeGe (19; 20; 21) and $\mathbf{H}$ is the external magnetic field. Since magnetocrystalline anisotropy and demagnetisation energies are not crucial for the emergence of skyrmionic states (19; 22), we omit them for simplicity. The first variational derivative of the energy functional $E[\widehat{\mathbf{m}}]$ with respect to $\widehat{\mathbf{m}}$ is called the effective field $\mathbf{H}_{\mathrm{eff}}(\mathbf{r})$. The magnetisation field dynamics is governed by the Landau-Lifshitz-Gilbert (LLG) equation (23):

$$\frac{d\widehat{\mathbf{m}}}{dt} = -\frac{\gamma_0^*}{1+\alpha^2} \widehat{\mathbf{m}} \times \mathbf{H}_{\mathrm{eff}} - \frac{\gamma_0^* \alpha}{1+\alpha^2} \widehat{\mathbf{m}} \times (\widehat{\mathbf{m}} \times \mathbf{H}_{\mathrm{eff}}) =: \ell(\widehat{\mathbf{m}}, \mathbf{H}_{\mathrm{eff}}(\widehat{\mathbf{m}})), \tag{2}$$

where the constant $\gamma_0^* = 2.211 \times 10^5 \,\mathrm{s\,m^{-1}A^{-1}}$ is the modified gyromagnetic ratio and α is the Gilbert damping constant dependent on the material, for FeGe it is $\alpha = 0.1$. From the LLG equation, a dynamically steady state (equilibrium state) is reached when $\widehat{\mathbf{m}} \times \mathbf{H}_{\mathrm{eff}} = 0$. It is this condition that we exploit in the learning dynamics of the PhysGAN and we describe it in the next section. To enable tight integration between the updates in the neural network layers and the LLG dynamics, all the ingredients of equations (1) and (2) were implemented with Tensorflow primitive functions, instead of using third-party packages. We need the total energy density and the effective field functions to be embedded in custom generator loss functions so that Tensorflow can correctly compute the gradient with respect to the generator weights needed for the backpropagation. The correctness of those custom implementations was checked by comparing their output with the output of micromagnetic simulation software packages such as OOMMF (24) and mumax^3 (25) using Ubermag (26).

2.2 PhysGAN Losses

In GANs, a game is played between generator that produces a sample $G(z)$ from a 50-dimensional noise input vector drawn from a normal random distribution $z \sim p_z$ and discriminator $D(\cdot)$ that rejects/certifies the samples by comparing $D(G(z))$ with $D(x)$ from training samples $x \sim p_d$. A minimax equilibrium of a suitable loss function gives optimal weights for the networks in D and G. Compared to the cross-entropy loss (10) and Wasserstein distance (27) we found that the hinge loss (28):

$$L_D = \mathbb{E}_{x\sim p_d}[\max(0, 1 - D(x))] + \mathbb{E}_{z\sim p_z}[\max(0, 1 + D(G(z)))],\ L_G = -\mathbb{E}_{z\sim p_z}[D(G(z))] \tag{3}$$

gave the best results.

In training with simulated field configurations that were equilibrium states of the skyrmion lattice, the goal of the generator $G(z)$ was to create "fake" field configurations that satisfied the checks from the discriminator, D. However, these configurations, although visually indistinguishable from the "real" data, were found not to be consistent with the steady state of the LLG equation (2) and we added to the learning configuration a "physics-aware" loss term in order to generate physically meaningful spin fields. To this end, we require the mean energy density to be as low as possible, and that $\mathbf{L}_\times := \widehat{\mathbf{m}}(\mathbf{r}) \times \mathbf{H}_{\mathrm{eff}}$ is close to $\mathbf{0}$. This guarantees that the magnetization is at or is close to an equilibrium state and it cannot change any more as time progresses. Using the Adam optimization algorithm (29), we have experimented with losses involving 4 different versions, that we collectively refer to as L_{phys}: (1) E loss - the average energy density of a mini-batch of generated samples reduced to a single number; (2) cross product $\mathbf{L}_\times$ loss which has 3 variants - the average of $(\mathbf{L}_\times)^2$ or the norm $|\mathbf{L}_\times|$ or $|\mathbf{L}_\times|^2$ of a mini-batch of generated samples reduced to a single number; (3) $\alpha E + \beta |\mathbf{L}_\times|^p$ for $p = 1, 2$, which was difficult to choose α and β for, as $\mathbf{L}_\times$ ignores the magnitude M_{s} of $\mathbf{M} = M_{\mathrm{s}}\widehat{\mathbf{m}}$, making it not directly comparable to E; (4) LLG loss, derived from Eq. (2), whose vanishing ensures a stationary field configuration. It has 3 variants - the average of $\ell(\widehat{m}, \mathbf{H}_{\mathrm{eff}}(\widehat{\mathbf{m}}))^2$ or the norm $|\ell(\widehat{\mathbf{m}}, \mathbf{H}_{\mathrm{eff}}(\widehat{\mathbf{m}}))|$ or $|\ell(\widehat{\mathbf{m}}, \mathbf{H}_{\mathrm{eff}}(\widehat{\mathbf{m}}))|^2$ of a mini-batch of generated samples reduced to a single number. Before calculating its gradient for the backpropagation, if the chosen L_{phys} is (2) or (4), it is multiplied by 3×10^{-9} which is the time step in seconds set in the evolver of the simulation software used to generate the training magnetization samples.

3 Training Data, Strategies and Results

The training samples of grid size 48×32 were generated using Ubermag (26) with OOMMF as a computational backend (24) with external field $\mathbf{B} = \mu_0\mathbf{H} = 0.1\,\mathrm{T}\,\widehat{\mathbf{z}}$ and also samples of size 144×86 with $\mathbf{B} = \mu_0\mathbf{H} = 0.01\,\mathrm{T}\,\widehat{\mathbf{z}}$ we generated using mumax^3 (25) for our experiments. Those samples form in bulk FeGe. The skyrmion-hosting material can be changed so that the PhysGAN generates samples that would form on the new material. This is achieved by training on samples formed on the new material. The LLG loss' Gilbert damping α needs to be tuned accordingly. As training with 144×86 magnetizaiton samples is quite computationally intensive, all of the PhysGAN data presented here is about 48×32 samples, except for Fig. 5.

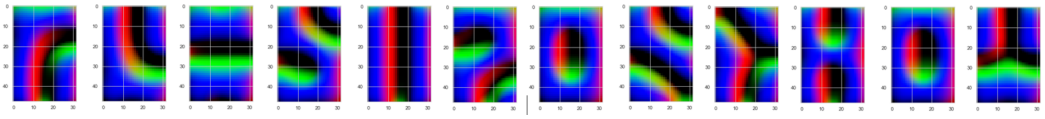

Figure 2: Example skyrmions samples of size 48×32 from the training data.

Training is achieved by maintaining two separate gradients to guide the updates to the generator G at every training step – one coming from the hinge loss with the discriminator D, and the other one is for the physics-aware L_{phys}. There are three strategies for training with physics losses: (1) Train G on only a chosen L_{phys}. (2) For each training step, update D and G using the hinge loss and update G again but with L_{phys}. This takes longer to train, but generated samples are more diverse, albeit with mixed results; (3) Either train the network only with hinge loss for several epochs and then train G only on L_{phys} , or execute several rounds of the PhysGAN cycle from (2) followed by training G only on L_{phys}. This results in smoothed generated samples, removing any artifacts, rendering the fake skyrmions more 'physical'. However, partial mode collapse (14) (30) appears if it is over-trained.

To evaluate the results, for every strategy, samples were generated from a newly-instantiated network with weights initialized with a uniform random distribution and trained for 10^3 epochs or more. Then, they were visually compared to the training data (31). Furthermore, two quantitative measures were calculated for a random batch of generated samples: the average energy density (AED) computed the same way as the E loss and the average cross product (ACP) defined as the average of $\mathbf{L}_{\times}$ for a batch of samples reduced to single number. It was checked how close they were to the reference values computed from the training dataset with samples of size 48×32, which are AED: -8.19×10^4 and ACP: $[1.76, 6.78, 1.83] \times 10^{-4}$.

As mentioned, hinge loss performance was far superior to cross entropy which suffered from mode collapse (32; 30). A basic only-hinge loss PhysGAN proved to be stable as testing it by running many newly-instantiated networks showed that the generated fake images were diverse enough and almost visually indistinguishable from the training set. The generator and discriminator losses also maintained Nash equilibrium (see Fig. 4). The stability is retained thanks to the custom GAN architecture, Adam optimizer and hinge loss. However, the physical quantitative measures deviated a lot from the reference values, although not by too much as shown in Fig. 4. The combination of physics and hinge loss does improve the 'physicality' of the artificial magnetizations as showcased in Fig. 3 and Fig. 4. In both GANs on Fig. 4 the generator and discriminator losses maintain Nash equilibrium and remain stable. The energy density E of the generated samples by the GAN on the left fluctuates a lot without any trend of being minimized, whilst the other GAN with integrated E loss actively minimizes the energy density E, so the AED is closer to the reference value at the 1000th epoch than the only-hinge-loss GAN. Similar effect is present with the other physics losses.

Testing of the four physics losses with training strategy (1) demonstrated that this method leads to imminent mode collapse in 100% of the test cases because the loss naturally guides the generator towards a single optimal point, thus generating almost the same magnetizations for any input vector. Although, this strategy had the worst performance it proved to be useful for fine-tuning the variable parameters of the physics losses for the other experiments. Fig. 3 is an example of strategy (2). The chart clearly illustrates that the L_{phys} can be minimized simultaneously with the hinge loss. G generated mostly diverse magnetizations which are visually indistinguishable from the training data and the AED is also very close to its reference training value. The drawback is that the ACP

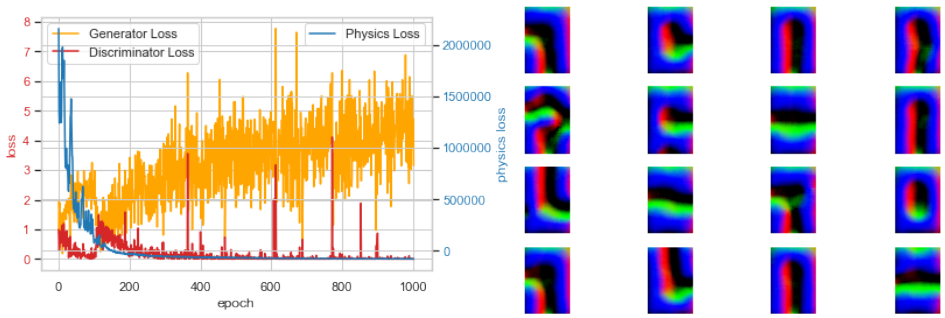

Figure 3: Left: The losses for PhysGAN trained simultaneously with hinge and E loss for 10^3 epochs. Right: Generated samples of size 48×32 with AED: -7.73×10^4; ACP: $[-2623, 1820, 829]$.

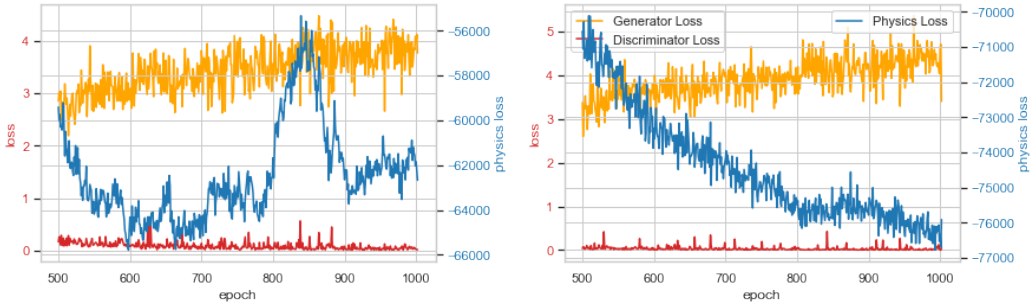

Figure 4: Left: The average hinge losses and E physics loss for the last 500 out of 1000 epochs of the training of 10 PhysGANs with only hinge loss, the E physics loss is calculated only as a score and not applied to the generators' gradients. One such PhysGAN had AED: -6.61×10^4, ACP: $[-2268, 160, -6512]$; Right: The average hinge losses and E losses of the training of 10 PhysGANs for the last 500 out of 1000 epochs.

significantly deviates from the training ACP. Those statements are further confirmed by additional testing with the other physics losses.

Training strategy (3) performed better with the ACP compared to (2), while maintaining the AED close to its reference value. This type of training works best combined with $\mathbf{L}_\times$ losses (a PhysGAN trained on $\mathbf{L}_\times^2$ loss for 10^3 epochs with strategy (3) had AED: -8.1×10^4; ACP: $[-937, -484, 1678]$) and LLG losses (a PhysGAN trained on $\ell(\widehat{\mathbf{m}}, \mathbf{H}_{\text{eff}}(\widehat{\mathbf{m}}))^2$ loss with strategy (3) for 2×10^3 epochs had AED: -8.2×10^4; ACP: $[-1552, -586, 12]$). See Appendix for more data. However, the partial mode collapse is more prevalent if the network is over-trained. The physics losses tend to smooth out the G output and remove out-of-place artifacts in strategies (2) and (3). Furthermore, it is noticeable in Fig. 5 how the generator finds a more optimal state during the L_{phys} only training.

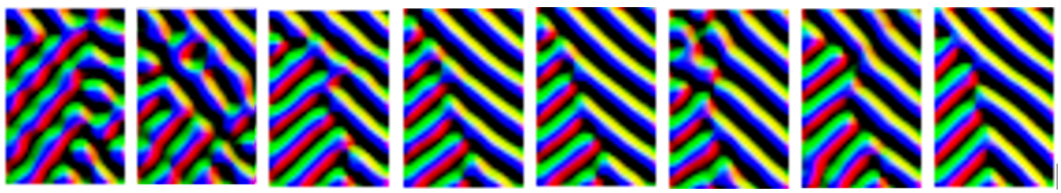

Figure 5: 144x96 PhysGAN output ($H = 0.01$, AED: -10^5); first trained only on hinge loss, then on E loss for 100 epochs. Output from left to right are from epoch 6 to 41 with step of 6.

This work demonstrated the power of the novel generative adversarial networks for physics data generation. The proposed PhysGAN variations show promising although not yet perfect results regarding the generation of new skyrmion magnetization states while trying to maintain them close to an equilibrium state.

Broader Impact

The PhysGAN allows for lots of room for improvement and additional features such as making it conditional where it can generate different classes of skyrmions depending on the external field $\mathbf{H}$ or some other energy term. Another possibility for future work is experimenting with different energy equations and parameters. Although most of those parameters depend on the material, an interesting experimentation would be to vary them as well and check if skyrmions emerge. This would assist the process of finding new helimagnetic materials. This work can also serve as a stepping stone for other physics or any scientific field research which requires a similar generative model where a physics/biology/chemistry rule, invariance or equation is transformed into custom loss functions that guide the generator output.

References

[1] T. H. Skyrme, "A unified field theory of mesons and baryons," *Nuclear Physics*, vol. 31, no. C, pp. 556–569, 1962.

[2] A. N. Bogdanov and D. A. Yablonskii, "Thermodynamically stable "vortices" in magnetically ordered crystals. the mixed state of magnets," *Sov. Phys. JETP*, vol. 68, pp. 101–3, Jan 1989.

[3] S. Mühlbauer, B. Binz, F. Jonietz, C. Pfleiderer, A. Rosch, A. Neubauer, R. Georgii, and P. Böni, "Skyrmion lattice in a chiral magnet," *Science*, vol. 323, no. 5916, pp. 915–919, 2009. [Online]. Available: https://science.sciencemag.org/content/323/5916/915

[4] N. Nagaosa and Y. Tokura, "Topological properties and dynamics of magnetic skyrmions," *Nature Nanotechnology*, vol. 8, no. 12, pp. 899–911, 2013. [Online]. Available: https://doi.org/10.1038/nnano.2013.243

[5] R. Tomasello, E. Martinez, R. Zivieri, L. Torres, M. Carpentieri, and G. Finocchio, "A strategy for the design of skyrmion racetrack memories," *Scientific Reports*, vol. 4, no. 1, p. 6784, 2014. [Online]. Available: https://doi.org/10.1038/srep06784

[6] D. Pinna, F. Abreu Araujo, J.-V. Kim, V. Cros, D. Querlioz, P. Bessiere, J. Droulez, and J. Grollier, "Skyrmion gas manipulation for probabilistic computing," *Phys. Rev. Applied*, vol. 9, p. 064018, Jun 2018. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevApplied.9.064018

[7] I. E. Dzyaloshinskii, "Theory of helicoidal structures in antiferromagnets. i. nonmetals." *J. Exp. Theor. Phys.*, vol. 19, p. 960, 1964.

[8] T. Moriya, "Anisotropic superexchange interaction and weak ferromagnetism," *Physical Review*, vol. 120, no. 1, pp. 91–98, 1960.

[9] X. Z. Yu, N. Kanazawa, Y. Onose, K. Kimoto, W. Z. Zhang, S. Ishiwata, Y. Matsui, and Y. Tokura, "Near room-temperature formation of a skyrmion crystal in thin-films of the helimagnet FeGe," *Nature Materials*, vol. 10, no. 2, pp. 106–109, feb 2011. [Online]. Available: http://www.nature.com/nmat/journal/v10/n2/full/nmat2916.html

[10] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative Adversarial Networks," jun 2014. [Online]. Available: http://arxiv.org/abs/1406.2661

[11] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, X. Zheng, and G. Research, "TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems," Tech. Rep., 2016. [Online]. Available: www.tensorflow.org.

[12] F. Chollet, "Keras: The Python Deep Learning library," *Keras.Io*, 2015.

[13] H. Y. Kwon, N. J. Kim, C. K. Lee, and C. Won, "Searching magnetic states using an unsupervised machine learning algorithm with the Heisenberg model," *Physical Review B*, vol. 99, no. 2, p. 024423, jan 2019. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevB.99.024423

[14] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," in *4th International Conference on Learning Representations, ICLR 2016 - Conference Track Proceedings*, 2016.

[15] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *32nd International Conference on Machine Learning, ICML 2015*, vol. 1. International Machine Learning Society (IMLS), 2015, pp. 448–456.

[16] V. Dumoulin, F. Visin, and G. E. P. Box, "A guide to convolution arithmetic for deep learning," Tech. Rep., 2018. [Online]. Available: http://ethanschoonover.com/solarized

[17] B. Xu, N. Wang, T. Chen, and M. Li, "Empirical Evaluation of Rectified Activations in Convolutional Network," may 2015. [Online]. Available: http://arxiv.org/abs/1505.00853

[18] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *Journal of Machine Learning Research*, vol. 15, no. 56, pp. 1929–1958, 2014. [Online]. Available: http://jmlr.org/papers/v15/srivastava14a.html

[19] M. Beg, R. Carey, W. Wang, D. Cortés-Ortuño, M. Vousden, M.-A. Bisotti, M. Albert, D. Chernyshenko, O. Hovorka, R. L. Stamps, and H. Fangohr, "Ground state search, hysteretic behaviour, and reversal mechanism of skyrmionic textures in confined helimagnetic nanostructures," *Scientific Reports*, vol. 5, no. October, p. 17137, nov 2015. [Online]. Available: http://www.nature.com/articles/srep17137

[20] M. Beg, M. Albert, M.-A. Bisotti, D. Cortés-Ortuño, W. Wang, R. Carey, M. Vousden, O. Hovorka, C. Ciccarelli, C. S. Spencer, C. H. Marrows, and H. Fangohr, "Dynamics of skyrmionic states in confined helimagnetic nanostructures," *Physical Review B*, vol. 95, no. 1, p. 014433, jan 2017. [Online]. Available: http://link.aps.org/doi/10.1103/PhysRevB.95.014433

[21] M. Beg, R. A. Pepper, D. Cortés-Ortuño, B. Atie, M.-A. Bisotti, G. Downing, T. Kluyver, O. Hovorka, and H. Fangohr, "Stable and manipulable Bloch point," *Scientific Reports*, vol. 9, no. 1, p. 7959, dec 2019. [Online]. Available: http://www.nature.com/articles/s41598-019-44462-2

[22] D. Cortés-Ortuño, M. Beg, V. Nehruji, L. Breth, R. Pepper, T. Kluyver, G. Downing, T. Hesjedal, P. Hatton, T. Lancaster, R. Hertel, O. Hovorka, and H. Fangohr, "Proposal for a micromagnetic standard problem for materials with Dzyaloshinskii–Moriya interaction," *New Journal of Physics*, vol. 20, no. 11, p. 113015, nov 2018. [Online]. Available: http://arxiv.org/abs/1803.11174http://stacks.iop.org/1367-2630/20/i=11/a=113015?key=crossref.ea259c795a4bf527ead3b7eaf0cee2d2

[23] T. L. Gilbert, "A phenomenological theory of damping in ferromagnetic materials," *IEEE Transactions on Magnetics*, vol. 40, no. 6, pp. 3443–3449, nov 2004. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1353448

[24] M. J. Donahue and D. G. Porter, "OOMMF User Guide," Cambridge, pp. xiii–xxii, 2012. [Online]. Available: http://ebooks.cambridge.org/ref/id/CBO9780511841538A006

[25] A. Vansteenkiste, J. Leliaert, M. Dvornik, M. Helsen, F. Garcia-Sanchez, and B. Van Waeyenberge, "The design and verification of MuMax3," *AIP Advances*, vol. 4, no. 10, p. 107133, oct 2014. [Online]. Available: http://aip.scitation.org/doi/10.1063/1.4899186

[26] M. Beg, R. A. Pepper, and H. Fangohr, "User interfaces for computational science: A domain specific language for OOMMF embedded in Python," *AIP Advances*, vol. 7, no. 5, p. 056025, may 2017. [Online]. Available: http://aip.scitation.org/doi/10.1063/1.4977225

[27] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein GAN," Tech. Rep., 2017.

[28] J. Zhao, M. Mathieu, and Y. LeCun, "Energy-based generative adversarial network," 2017.

[29] D. P. Kingma and J. L. Ba, "Adam: A method for stochastic optimization," in *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*. International Conference on Learning Representations, ICLR, 2015.

[30] T. Salimans, I. Goodfellow, W. Zaremba, V. Cheung, A. Radford, and X. Chen, "Improved Techniques for Training GANs," Tech. Rep. [Online]. Available: https://github.com/openai/

[31] A. Borji, "Pros and Cons of GAN Evaluation Measures," Tech. Rep., 2018.

[32] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," 2016.

4 Appendix

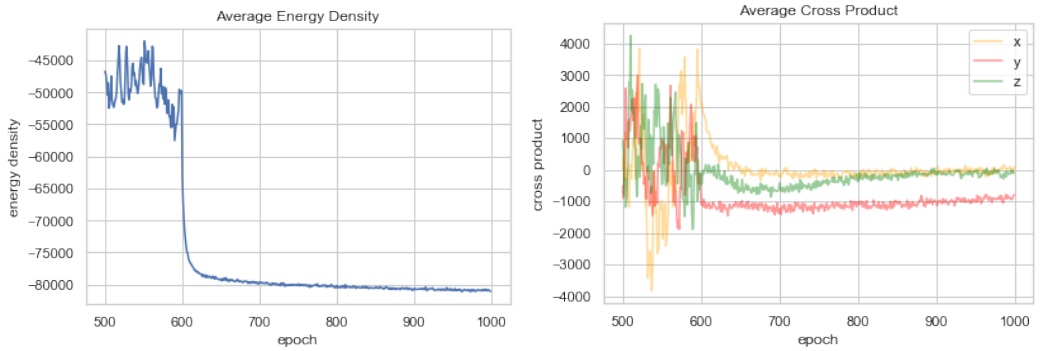

Figure 6: The average physical measures for the last 500 epochs of 10 PhysGANs trained with hinge and $|\ell(\widehat{m}, \mathbf{H}_{\text{eff}}(\widehat{\mathbf{m}}))|$ loss for 600 epochs, then only on $|\ell(\widehat{m}, \mathbf{H}_{\text{eff}}(\widehat{\mathbf{m}}))|$ loss for 400 epochs. Left: AED; Right: the three components of the ACP.

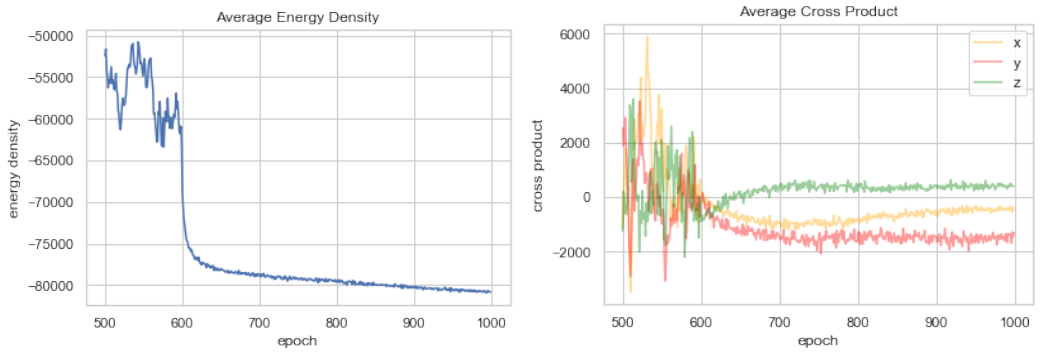

Figure 7: The average physical measures for the last 500 epochs of 10 PhysGANs trained with hinge and $|L_{\times}|^2$ loss for 600 epochs, then only on $|L_{\times}|^2$ loss for 400 epochs. Left: AED; Right: the three components of the ACP.

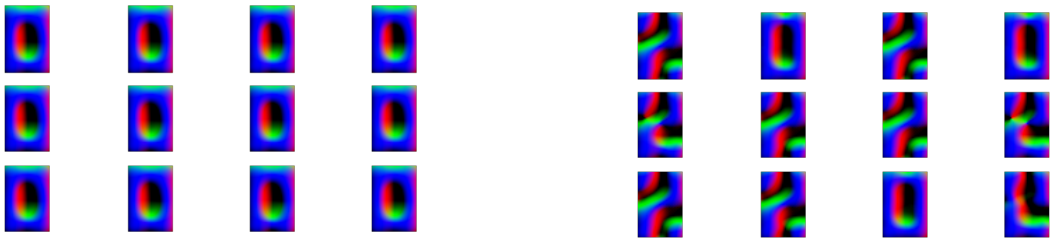

Figure 8: Left: samples generated at the last epoch of a PhysGAN with the same configuration as the one in Fig. 7; Right: samples generated at the last epoch of a PhysGAN with the same configuration as the one in Fig. 6.

In Fig. 6 and Fig. 7 it is noticeable how switching to physics-loss-only training drastically minimizes the physical measures and gets them much closer to the training reference values. Regarding the ACP, the PhysGAN from Fig. 6 performs much better than the other one when it comes to getting the x, y, z components close to reference values which are almost 0, although the y-component of the ACP deviates more than the other components. The final result at the 1000th epoch is that the AED is essentially almost equal to its reference value in both cases, but the generated magnetization samples lack diversity due to partial mode collapse evident in Fig 8.