Supplementary Material for "Machine Learning Assisted Prediction of Dynamics in Current-Driven Nested Skyrmion Bags"

Rui Li, Yuge Zhu, Xinyu Zhang, Mengting Li, Xingqiang Shi, Ruining Wang, Jianglong Wang, Hu Zhang, Penglai Gong, and Chendong Jin*

Hebei Research Center of the Basic Discipline for Computational Physics, Hebei Key Laboratory of High-precision Computation and Application of Quantum Field Theory, College of Physics Science and Technology, Hebei University, Baoding 071002, P. R. China.

I. Relaxation of S(2, S(2)) under varying *D*

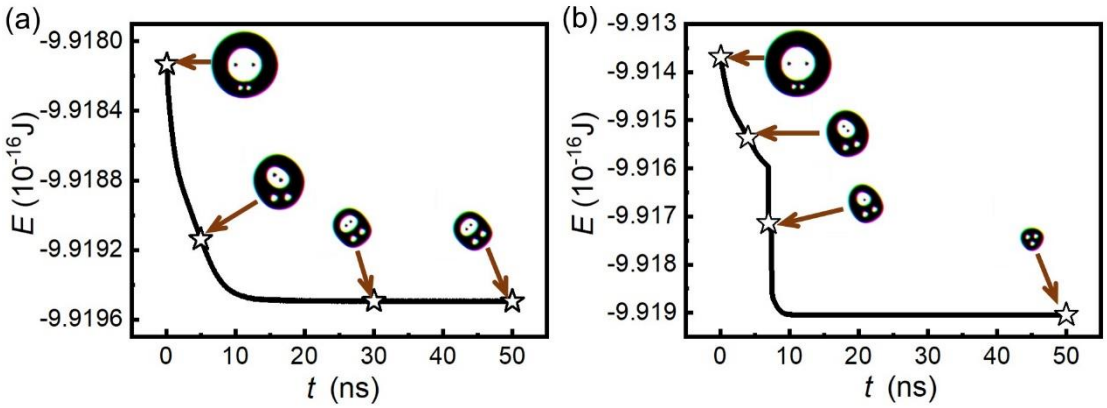

FIG. S1. Relaxation dynamics of the initial *S*(2, *S*(2)) structure over time under a DMI strength *D* of (a) 3.5 mJ/m^2 and (b) 3.4 mJ/m^2.

In Fig. 2 of the main text, we investigated the ground states of nested skyrmion bags under different nesting levels and DMI strengths *D*, and observed that structures with the same initial nesting level may collapse into simpler skyrmion bags or a single skyrmion under lower *D*. Here, we select an example with nesting level 2 (initial state configured as *S*(2, *S*(2))) and provide a detailed comparison of the relaxation processes at *D* = 3.5 mJ/m^2 and *D* = 3.4 mJ/m^2, as shown in Fig. S1. For *D* = 3.5 mJ/m^2, the system undergoes energy minimization through relaxation. The initial *S*(2, *S*(2)) configuration gradually reduces in size during the first 10 ns of relaxation and remains stable thereafter up to 50 ns, as shown in Fig. S1(a). In contrast, the same initial state evolves differently during energy minimization for *D* = 3.4 mJ/m^2, as shown in Fig. S1(b). Due to the reduced *D*, the two internal skyrmions are insufficient to overcome the energy barrier of the domain wall and annihilate at 7 ns and 10 ns, respectively. As a result, the initial *S*(2, *S*(2)) nested skyrmion bag eventually transitions into an *S*(3) skyrmion bag. These results demonstrate that *D* plays a critical role in determining the structural stability of nested skyrmion bags.

II. Field-like torque in nested skyrmion bag dynamics

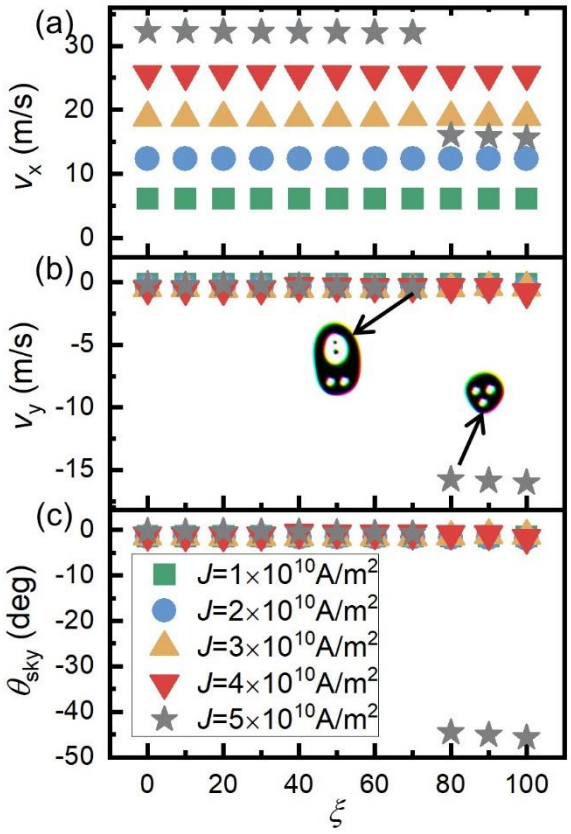

FIG. S2. (a) v_x, (b) v_y, and (c) θ_{sky} of $S(2, S(2))$ structure under different J, as functions of the field-like torque strength ξ.

The presence of a field-like torque induces an in-plane tilting of the magnetization. We therefore investigated its influence on the dynamics of the nested skyrmion bag $S(2, S(2))$ for different field-like torque strengths ξ, as shown in Fig. S2. For $J < 5 \times 10^{10}$ A/m^2, the values of v_x, v_y and θ_{sky} remain nearly unchanged with increasing ξ. However, at $J = 5 \times 10^{10}$ A/m^2, the $S(2, S(2))$ structure transitions into a skyrmion bag $S(3)$ when ξ reaches a critical value of 80, as illustrated in the insets of Fig. S2(b). In practice, the field-like torque in real materials is typically very small and significantly lower than the critical threshold of $\xi = 80$ observed in our calculations. For instance, in systems such as Pd/Fe bilayers on Ir (111) substrates [1], Permalloy layers in spin valves [2], and magnetic multilayers [3], the field-like torque strength ξ is smaller than 0.1. Even in Pt/Co/MgO thin films [4], a relatively high measured value of ξ is only about 0.45. Therefore, we have neglected the effect of the field-like torque in the analysis of nested skyrmion bag dynamics in the main text.

III. Dependence of the tensors on α, D, and J

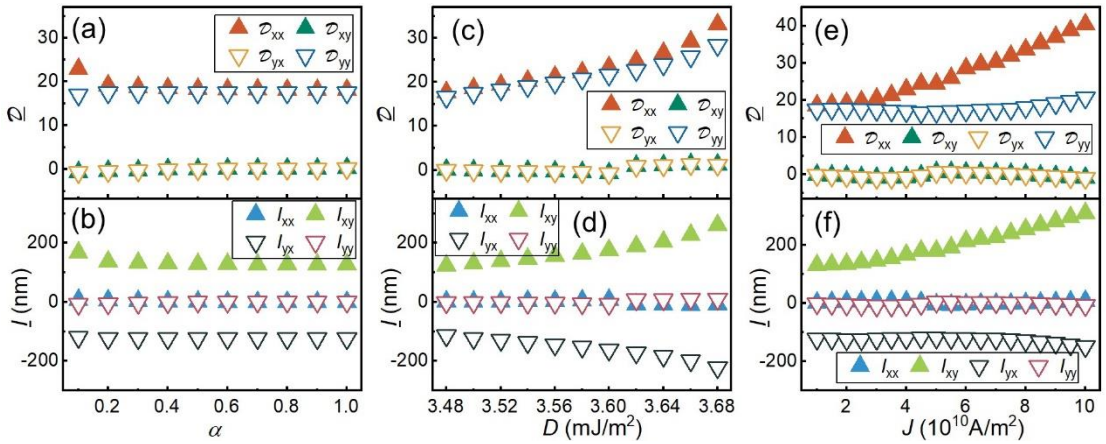

FIG. S3. Dependence of the tensors on α, D, and J for the $S(2, S(2))$ structure. (a) Dissipative tensor $\underline{\mathcal{D}}$ and (b) tensor $\underline{I}$ as functions of α with $J = 2 \times 10^{10}$ A/m^2 and $D = 3.5$ mJ/m^2. (c) Dissipative tensor $\underline{\mathcal{D}}$ and (d) tensor $\underline{I}$ as functions of D with $J = 2 \times 10^{10}$ A/m^2 and $\alpha = 0.2$. (e) Dissipative tensor $\underline{\mathcal{D}}$ and (f) tensor $\underline{I}$ as functions of J with $D = 3.5$ mJ/m^2 and $\alpha = 0.2$.

In Fig. 5 of the main text, we investigated the influence of various parameters (α, D, and J) on the dynamics (v_x, v_y, and θ_{sky}) of the nested skyrmion bag $S(2, S(2))$ using a combination of micromagnetic simulations and analysis based on the Thiele equation. The variations of the tensors $\underline{\mathcal{D}}$ and $\underline{I}$ used in the theoretical analysis with these parameters are provided in Fig. S3. Figures S3(a) and (b) show that $\mathcal{D}_{xx}$ and I_{xy} exhibit larger values at low α, which leads to a non-zero θ_{sky}. Figures S3(c) and (d) demonstrate that the magnitudes of $\mathcal{D}_{xx}$, $\mathcal{D}_{yy}$, I_{xy}, and I_{yx} all increase with the DMI strength D, resulting in a corresponding increase in v_x for $S(2, S(2))$. Figures S3(e) and (f) reveal that $\mathcal{D}_{xx}$ and I_{xy} increase significantly with J, implying that the $S(2, S(2))$ structure undergoes deformation due to the increasing Magnus force.

IV.Comparison of confinement potentials of $S(9, S(9))$

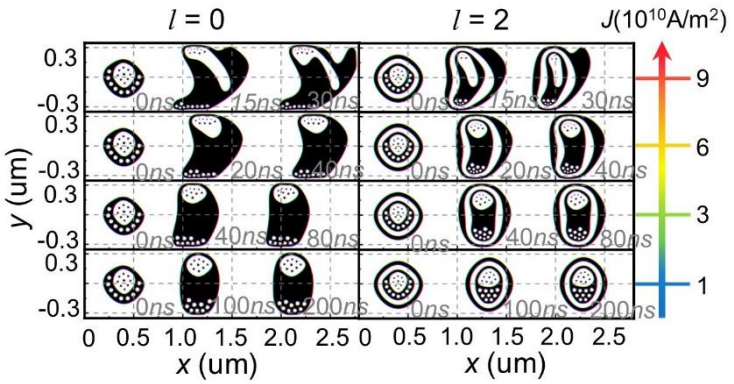

FIG. S4. Comparison of the motion of $S(9, S(9))$ surrounded by zero ($l = 0$) and two ($l = 2$) concentric domain walls under different J. The width of the nanotrack (y-direction) is 700 nm.

In addition to the suppression of transverse elongation in nested skyrmion bags by the outer domain walls proposed in the main text, the boundaries of the nanotrack can also provide a certain degree of confinement, as shown in Fig. S4. We replaced the original square simulation system with a side length of 2800 nm with a nanotrack structure with a width of 700 nm. Under the combined confinement effects of both the nanotrack boundaries and the domain walls, the $S(9, S(9))$ structure remains mobile under current densities of $J = 1 \times 10^{10}$ A/m^2, 3×10^{10} A/m^2, and 6×10^{10} A/m^2. It should be noted that the deformation of $S(9, S(9))$ is less pronounced when two additional domain walls are added ($l = 2$) compared to the case where confinement is provided solely by the nanotrack boundaries ($l = 0$). However, when the current density J increases to 9×10^{10} A/m^2, both the $l = 0$ and $l = 2$ configurations exhibit significant deformation accompanied by the annihilation of internal skyrmions.

V. Dynamics of nested skyrmion bags $S(m, S(5))$

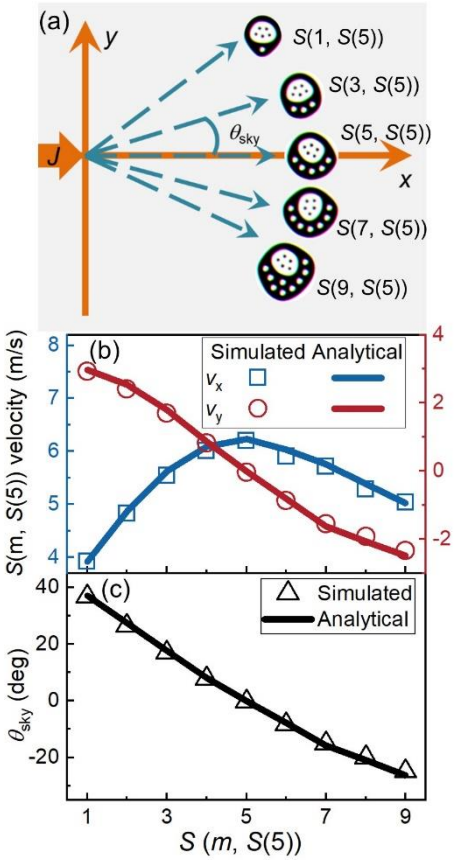

FIG. S5. Current-driven motion of nested skyrmion bags $S(m, S(5))$. (a) Schematics trajectories of five representative $S(m, S(5))$ (m = 1, 3, 5, 7, 9) driven by SOT. The initial position for all $S(m, S(5))$ is at the center of the two-dimensional plane. The solid orange arrow indicates the direction of current flow with a density of 1×10^{10} A/m^2. The dashed dark cyan arrows illustrate the trajectories of the different nested skyrmion bags. (b) Velocity and (c) θ_{sky} of $S(m, S(5))$ as functions of m. The Hollow symbols and solid lines represent simulated and theoretically analytical results, respectively.

In the main text, we investigated the dynamic processes of $S(5, S(n))$ structures. For comparison, we have also studied the case of $S(m, S(5))$, as shown in Fig. S5. Figure S5(a) displays the trajectories of five nested skyrmion bags: $S(1, S(5))$, $S(3, S(5))$, $S(5, S(5))$, $S(7, S(5))$, and $S(9, S(5))$, with topological charges Q of -4, -2, 0, 2, and 4, respectively (where $Q = m - 5$). Fig. S5(b) shows that as m increases from 1 to 9, v_x first increases and then decreases, while v_y decreases from 2.96 m/s to -2.49 m/s. Correspondingly, the skyrmion Hall angle θ_{sky} decreases from 37° to -26°. Therefore, variation in either m or n within nested skyrmion bags enables effective tuning of θ_{sky}.

VI. Dependence of MAPE on dataset size

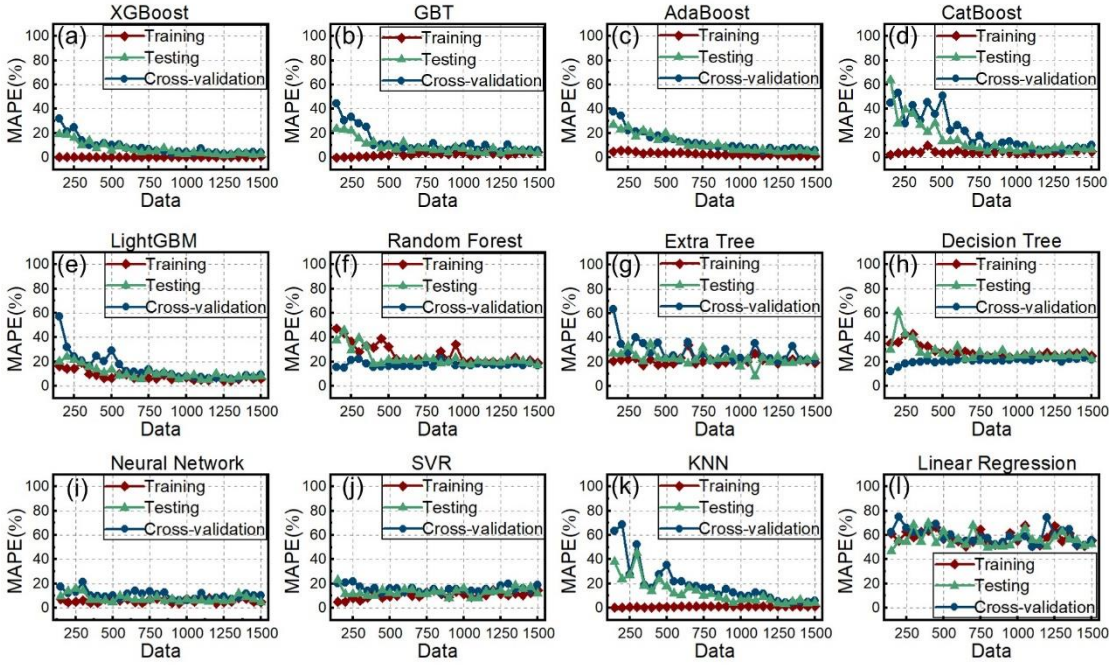

FIG. S6. Variation of the Training MAPE, Testing MAPE, and Cross-validation MAPE of different ML models as the dataset size increases from 150 to 1500.

Figure S6 presents the variation in training, testing, and cross-validation MAPE of 12 ML models with increasing dataset size, offering insights into their generalization capability and data efficiency in predicting θ_{sky}. Gradient boosting models, including XGBoost, GBT, AdaBoost, CatBoost, and LightGBM (Figs. S6(a)-(e)), consistently demonstrate excellent data efficiency, with both testing and cross-validation MAPE rapidly decreasing and stabilizing near the training error as data size increases beyond approximately 1000 samples. This indicates their strong capability to learn the underlying physical relationships without overfitting, corroborating their superior performance shown in Table 1 and Fig. 10. The bagging methods, Random Forest and Extra Trees (Figs. S6(f)–(g)), show

intermediate performance with reasonable but slower convergence compared to gradient boosting methods. Decision Tree (Fig. S6(h)) exhibits overfitting, characterized by minimal training error but poor generalization. Notably, Neural Network (Fig. S6(i)) exhibits similar characteristics to the top gradient boosting models, with all three curves converging at large data sizes, confirming its capability to capture complex non-linear relationships when sufficient data is available. SVR (Fig. S6(j)) achieves moderate yet fluctuating performance, likely due to its sensitivity to hyperparameter selection. KNN (Fig. S6(k)) shows characteristically high initial error that decreases sharply with more data, indicating that sufficient data coverage is critical for this instance-based method to capture underlying patterns. In contrast, Linear Regression (Fig. S6(l)) shows fundamentally different behavior: despite perfect training fit, its testing and cross-validation errors remain persistently high even with increasing data, providing clear evidence that the relationship between input parameters and θ_{sky} is inherently non-linear and cannot be captured by linear models. The rapid convergence and stable performance of gradient boosting models and neural networks make them particularly suitable for modeling complex spintronic systems. Their data efficiency enables reliable prediction with feasible data collection efforts, facilitating accelerated exploration of skyrmion-based device parameter spaces.

VII. Description of four Supplementary videos

Supplementary video I corresponds to Fig. 4, showing the dynamic process of nested skyrmion bags $S(2, S(2))$ and $S(9, S(9))$ driven by SOT with $J = 1 \times 10^{10}$A/m^2, $\alpha = 0.2$, and $D = 3.5$ mJ/m^2.

Supplementary video II corresponds to Fig. 7, showing the dynamic process of nested skyrmion bags with different l driven by SOT with $J = 2 \times 10^{10}$A/m^2, $\alpha = 0.2$, and $D = 3.5$ mJ/m^2.

Supplementary video III corresponds to Fig. 9, showing the dynamic process of nested skyrmion bags $S(5, S(n))$ driven by SOT with $J = 1 \times 10^{10}$A/m^2, $\alpha = 0.2$, and $D = 3.5$ mJ/m^2.

Supplementary video IV corresponds to Fig. 11, showing the nested skyrmion bags-based 1-to-4 DEMUX.

---

*Corresponding author: Chendong Jin

References

[1] Ritzmann, U., Desplat, L., Dupé, B., Camley, R. E. & Kim, J.-V. Asymmetric skyrmion-antiskyrmion production in ultrathin ferromagnetic films. *Phys. Rev. B* **102**, 174409 (2020).

[2] Boone, C. T. *et al.* Rapid Domain Wall Motion in Permalloy Nanowires Excited by a Spin-Polarized Current Applied Perpendicular to the Nanowire. *Phys. Rev. Lett.* **104**, 097203 (2010).

[3] Abert, C. *et al.* Fieldlike and Dampinglike Spin-Transfer Torque in Magnetic Multilayers. *Phys. Rev. Appl.* **7**, 054007 (2017).

[4] Juge, R. *et al.* Current-driven skyrmion dynamics and drive-dependent skyrmion hall effect in an ultrathin film. *Phys. Rev. Appl.* **12**, 044007 (2019).