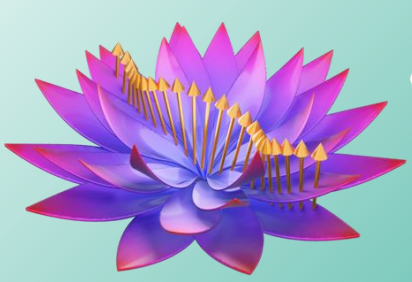

NIMFEIA

Deliverable D7.6

Report on implementation of adaptive time integration methods for machine learning simulations

Project number	101070290
Project name	Nonlinear Magnons for Reservoir Computing in Reciprocal Space
Project acronym	NIMFEIA
Work package	WP 7.2 Machine learning simulations of magnon processes in antiferro-magnets
Type	Report
Dissemination level	Public
Lead beneficiary	Radboud University
Due date of delivery	Month 18

Authors: Hrvoje Vrcan, Johan Mentink

Disclaimer:

The NIMFEIA project has received funding by the European Union's Research and Innovation Programme Horizon Europe under grant agreement No 101070290. However, views and opinions expressed in this document are those of the authors only and do not necessarily reflect those of the European Union. The European Union cannot be held responsible for them.

1. Introduction

1.1. Motivation

To investigate nonlinear magnon dynamics at the highest energy and shortest time scales, magnons with the shortest possible wavelength should be exploited. Classical approaches such as micromagnetism or atomistic spin dynamics are no longer good approximations to capture the physics at the shortest length and fastest time scales. The very nature of spins as quantum entities makes a quantum approach unavoidable at these scales. We approach the study of magnons, at scales corresponding to the edge of the Brillouin zone, from the fundamental principles of quantum mechanics.

However, from this approach, a quantum many-body problem emerges, making any kind of exact solutions for decently-sized systems quickly unviable. To tackle this difficult obstacle, we follow a novel approximation method based on using machine learning in the framework of the variational principle of quantum mechanics [1]. This approach was shown to successfully capture the features of many physical systems, including the magnetic models of quantum spins on the lattice.

Still, this method has been applied in only the linear regimes of magnetic dynamics so far [2]. For nonlinear dynamics, the numerical solution of the equations of motion derived from variational principles causes a breakdown [3], limiting the simulation of nonlinear magnonics at the edge of the BZ. To progress beyond this limitation, we explore more efficient time-integration methods to obtain accurate and stable dynamics, using exact solutions of minimal models with simple driving as a benchmark.

1.2. 2D magnetism

A minimal model to study magnons at the edge of the Brillouin zone is the 2D Heisenberg antiferromagnet. Taking into account only the nearest-neighbour exchange interactions of N spins on a lattice, such a system is described by the Heisenberg Hamiltonian:

$$H = \sum_{i,\boldsymbol{\delta}} (J_0 + J(t)\boldsymbol{e}\cdot\boldsymbol{\delta})\boldsymbol{S}_i \cdot \boldsymbol{S}_{i+\boldsymbol{\delta}}$$

Here, the $\boldsymbol{S}_i = ({S_i}^x, {S_i}^y, {S_i}^z)$ is the total spin vector operator of the spin on the lattice site i. The summation is done over the four lattice vectors $\boldsymbol{\delta} = (\pm 1, 0), (0, \pm 1)$ that define the four lattice directions. This defines the summation over all the nearest-neighbour pairs in the lattice.

Of particular interest are the systems with antiferromagnetic coupling, where the exchange constant J_0 is positive. Dynamics are induced by perturbing the exchange bonds along some (e.g. vertical) lattice direction with an external field polarized in that direction $\boldsymbol{e}$, whose strength is governed by $J(t)$.

2. Method

2.1. Neural quantum states

In 2017, Carleo and Troyer demonstrated that neural networks are a good candidate for a variational ansätz for quantum systems, pioneering the field of *neural quantum states* (NQS) [1]. Following this approach, we represent the spin system by a neural network called Restricted Boltzmann Machine (RBM):

$$\Psi(s) = \prod_{j}^{M} 2\cosh\left(b_j + w_{ij}s_i^z\right).$$

Here, we represent the system by the values of z-components $s_i^z = \pm 1$ of spins at sites $i = 1, \dots, N$, and parameters b_j and w_{ij} at each network layer $j = 1, \dots, M$. The number of spins and layers are related by the network density parameter $\alpha = M/N$, which can be tuned to control the expressibility and accuracy of the neural network representation.

The RBM representation has, in practice, a drastically smaller number of parameters than the dimension of Hilbert space, which encapsulates the power of this approach. The ground-state wave function can be obtained by optimizing the network for minimal energy. The dynamical behaviour of the wave function is acquired from the time dependence of network parameters.

2.2. Time-dependent variational principle

To access the system dynamics from NQS, we use a time-dependent extension of the variational principle – *time-dependent variational principle* (TDVP). This approach consists of propagating the ansatz parameters, in our case the network's weights and biases, by a fundamental equation of motion:

$$S_{kk'}\dot{W}_{k'} = -iF_k.$$

Here, the vector $\boldsymbol{W} = (W_1, W_2, \dots, W_M)$ includes all the network parameters, while $\boldsymbol{F} = (F_1, F_2, \dots, F_M)$ is the energy gradient. The covariance matrix S contains the metric properties of the parameter space [4]. Further details on the calculation of these observables can be found in the supplementary material of [1].

Solved in imaginary time, this equation finds the ground state parameters, akin to optimising the neural network. In real time, it recovers the full dynamical description of the wave function through time dependence of the parameters [5]. Since, in general, the equation is very non-linear, a numerical integration scheme is used to obtain the solution.

2.3. Numerical time integration

The TDVP equation of motion is the fundamental principle that governs network parameters' time evolution. However, the energy gradient $\boldsymbol{F}$ and covariance matrix S are both complicated functions of network parameter vector $\boldsymbol{W}$, which makes numerical integration the only way to access the equation's solutions.

We use Heun's explicit scheme, a second-order time-integration algorithm, to acquire the values of parameters at each point in time. The scheme consists of the following update:

$$\boldsymbol{W}_{n+1} = \boldsymbol{W}_n + \frac{\mathrm{d}t}{2}\Big(\boldsymbol{f}(\boldsymbol{W}_n) + \boldsymbol{f}\big(\boldsymbol{W}_n + \mathrm{d}t\boldsymbol{f}(\boldsymbol{W}_n)\big)\Big).$$

Here, $\boldsymbol{W}_n$ and $\boldsymbol{W}_{n+1}$ are parameter values at times t_n and $t_{n+1} = t_n + \mathrm{d}t$ respectively, and the update function is $\boldsymbol{f} = (f_1, f_2, \dots, f_M) = -iS^{-1}\boldsymbol{F}$. It should be noted that the covariance matrix is generally ill-defined due to overdetermination, so the inverse S^{-1} must be obtained by a regularization method. A standard practice is to introduce a small diagonal regulation factor $S \to S + \varepsilon I$ [6], where the number ε typically has a value between 10^{-4} and 10^{-3} [2].

However, Heun's method with diagonal regularization has been shown to produce numerical instabilities for nontrivial lattices and strong driving [3]. We therefore present a modified version of the second-order Heun's algorithm:

$$\boldsymbol{W}_{n+1} = \boldsymbol{W}_n + \frac{\mathrm{d}t}{2}\left(\frac{\boldsymbol{f}(\boldsymbol{W}_n)}{1+\mathrm{d}t|\boldsymbol{f}(\boldsymbol{W}_n)|} + \frac{\boldsymbol{f}\big(\boldsymbol{W}_n + \mathrm{d}t\boldsymbol{f}(\boldsymbol{W}_n)\big)}{1+\mathrm{d}t\big|\boldsymbol{f}\big(\boldsymbol{W}_n + \mathrm{d}t\boldsymbol{f}(\boldsymbol{W}_n)\big)\big|}\right).$$

This update scheme is dubbed *tamed* Heun's algorithm [7], where the purpose of the absolute values in the denominators is to stabilize the time integration, therefore "taming" the evolution of network parameters. We will compare the standard Heun's update, which we dub *classic*, with the tamed integrator, and benchmark both against the exact solution.

3. Calculations

3.1. Overview

We performed calculations of a Heisenberg lattice with antiferromagnetic coupling represented with the RBM ansätz, using the TDVP method to get the network dynamics. To keep the results of exact diagonalization available, we opted for a small 4×4 lattice and restricted the Hilbert space to include only configurations with zero net spin, explicitly holding the system in an antiferromagnetic state. The Hilbert space averages necessary for the calculation of S and $\boldsymbol{F}$ are calculated with Monte Carlo sampling.

The system is first prepared in a ground state with no external driving. This is done by setting the driving function to $J(t) = 0$ and training the neural network for an optimal solution. We used a standard gradient descent training scheme for parameter updates:

$$\boldsymbol{W}_{n+1} = \boldsymbol{W}_n - \eta S^{-1}\boldsymbol{F},$$

where η is the learning rate. After the ground state optimization, an external perturbation is introduced in the shape:

$$J(t) = \Delta J \cdot J_0 \exp\left(-\frac{(t-1/J_0)^2}{2\cdot(0.4/J_0)^2}\right)\sin(8J_0 t).$$

This shape corresponds to an external laser pulse, polarized to perturb only the bonds along some lattice direction [3].

An observable of particular interest is the nearest-neighbour spin correlation function:

$$C(t) = \frac{1}{N}\sum_{i,\boldsymbol{\delta}} \langle \boldsymbol{S}_i \cdot \boldsymbol{S}_{i+\boldsymbol{\delta}} \rangle,$$

where the $i, \boldsymbol{\delta}$ once again denotes the summation over all the nearest neighbours pairs in the lattice. We can extract information about the propagation properties of the system through this observable, so it's essential to achieve its accurate calculation.

3.2. Results

We present the results for the correlation function calculated with both the classical and the tamed Heun integrator, benchmarked We present the results for the correlation function calculated with both the classical and the tamed Heun integrator, benchmarked against the ex-

act solutions. We compare the time dependence of four different values for the driving amplitude $\Delta J = \{0.1, 0.2, 0.3, 0.4\}$. The following calculations were done for network density of $\alpha = 4$ and time step $\mathrm{d}t = 0.0025$.

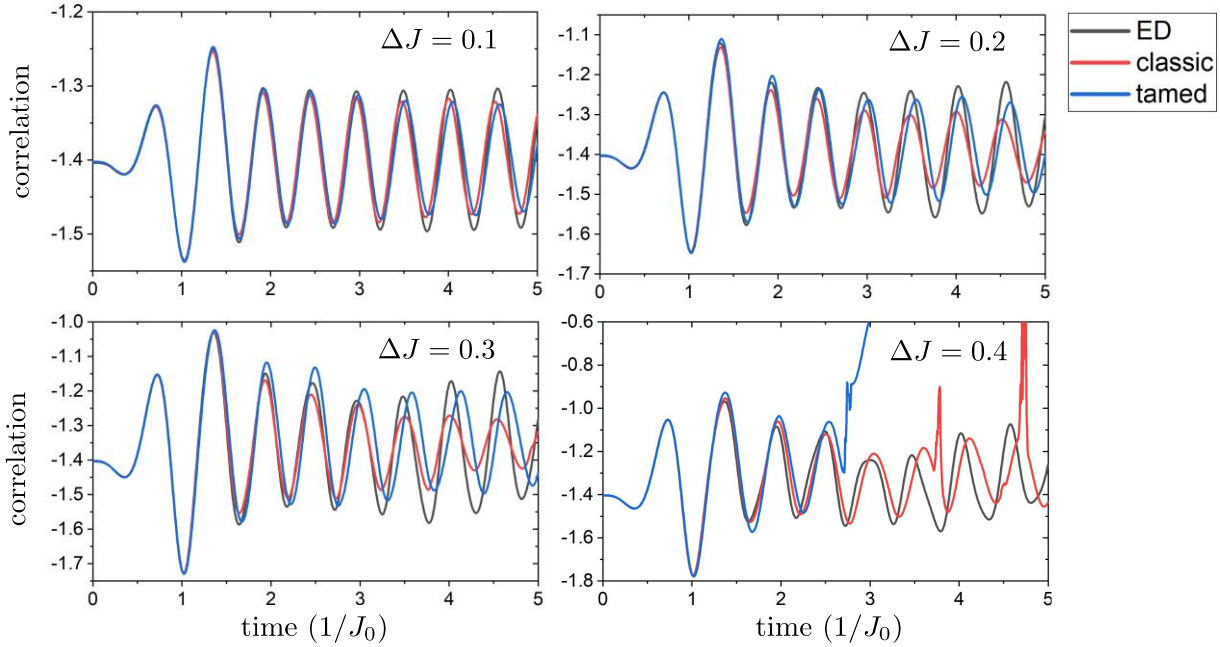

Figure 1 *Time dependence of the nearest-neighbour spin correlations function calculated by exact diagonalization (grey), and RBM simulations using classic Heun's integrator (red), and tamed Heun's integrator (blue). Results are shown for different values of the perturbation amplitude ΔJ of external driving. Simulated results agree with the exact benchmark only for the weakest driving, and a breakdown is observed for the strongest amplitude for both integrators.*

For the weakest driving amplitude of $\Delta J = 0.1$, the RBM calculations are consistent with the exact results for both the classic and the tamed Heun's schemes. However, deviations from the correct results are immediately visible for stronger driving amplitudes for both methods. Finally, the calculations break down in the case of $\Delta J = 0.4$, regardless of the integrator.

From the results presented, it's visible that neither the classic nor the tamed Heun's scheme can handle the time integration, as the strength of the driving increases.

4. Discussion

The results presented above show that, in the regime of challenging dynamics, both applied integration methods still have difficulties properly matching the exact solution. Numerical instabilities observed in [3] motivated a search for a more reliable integration method. Therefore, the results presented here call for a more thorough improvement in the adaptive integration method, beyond the tamed integrator.

To achieve both stability and accuracy of time integration, a completely symplectic method may be necessary, which in some cases is not possible using an explicit scheme. Implicit integrators such as the midpoint method [8] can be utilized to maintain the symplecticity of integration.

Finally, we address the inversion method used to enable solving the TDVP equation of motion. Regularization is a necessary procedure to enable to inversion of the S matrix, but the diagonal regulator itself is unphysical and may account for a source of instabilities amplified by the nonlinearities in the equation of motion. This can be especially problematic in the case

of strong driving. To tackle this problem, a different approach may be needed to obtain the solutions of the equation, such as convergent algorithms [9] that don't rely on direct inversion.

5. Bibliography

[1] G. Carleo and M. Troyer, "Solving the quantum many-body problem with artificial neural networks," *Science*, vol. 355, no. 6325, pp. 602–606, Feb. 2017, doi: 10.1126/science.aag2302.

[2] G. Fabiani, "Quantum dynamics of 2D antiferromagnets: predictions from theory and machine learning." Radboud University, PhD dissertatoion, Jun. 14, 2022. [Online] Available: hdl.handle.net/2066/250503

[3] D. Hofmann, G. Fabiani, J. Mentink, G. Carleo, and M. Sentef, "Role of stochastic noise and generalization error in the time propagation of neural-network quantum states," *SciPost Phys.*, vol. 12, no. 5, p. 165, May 2022, doi: 10.21468/SciPostPhys.12.5.165.

[4] J. Haegeman, J. I. Cirac, T. J. Osborne, I. Pizorn, H. Verschelde, and F. Verstraete, "Time-dependent variational principle for quantum lattices," *Phys. Rev. Lett.*, vol. 107, no. 7, p. 070601, Aug. 2011, doi: 10.1103/PhysRevLett.107.070601.

[5] K. Ido, T. Ohgoe, and M. Imada, "Time-dependent many-variable variational Monte Carlo method for nonequilibrium strongly correlated electron systems," *Phys. Rev. B*, vol. 92, no. 24, p. 245106, Dec. 2015, doi: 10.1103/PhysRevB.92.245106.

[6] R. Penrose, "A generalized inverse for matrices," *Math. Proc. Camb. Philos. Soc.*, vol. 51, no. 3, pp. 406–413, Jul. 1955, doi: 10.1017/S0305004100030401.

[7] C. Kelly and G. Lord, "Adaptive Euler methods for stochastic systems with non-globally Lipschitz coefficients." arXiv, Jun. 01, 2021. [Online] Available: http://arxiv.org/abs/1805.11137

[8] I. L. Gutiérrez and C. B. Mendl, "Real time evolution with neural-network quantum states," *Quantum*, vol. 6, p. 627, Jan. 2022, doi: 10.22331/q-2022-01-20-627.

[9] C. C. Paige and M. A. Saunders, "Solution of Sparse Indefinite Systems of Linear Equations," *SIAM J. Numer. Anal.*, vol. 12, no. 4, pp. 617–629, Sep. 1975, doi: 10.1137/0712047.