

Machine Learning in Ultrafast Magnetism

*New horizons for the fastest, smallest
and most energy-efficient
brain-inspired computing*

Johan H. Mentink

Radboud University



“Machine Learning for Quantum Matter”, Nordita, Stockholm, Aug. 27 2019

Acknowledgements

Radboud University



A.V. Kimel
Th. Rasing
C. Davies
K. Yamada
A. Chakravarty
M. Barbeau
G. Fabiani



Max-Planck-Institut für
Struktur und Dynamik der Materie

Max Planck Institute for the Structure and Dynamics of Matter



M. Eckstein
K. Balzer



THE UNIVERSITY OF TOKYO



D. Bossini
E.V. Gomonyay



Netherlands Organisation for Scientific Research

Challenges for (Brain-Inspired) Computing

Operation
conditions



$T=300\text{K}$
 $P=10^5\text{ Pa}$



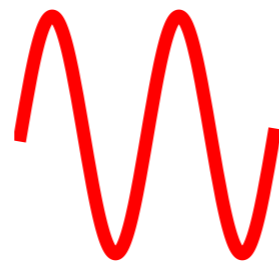
Dimensions
(synapse, neuron)



$L \sim \text{nm}$



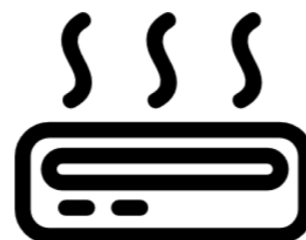
Speed
(adaptation time,
oscillation period)



$f \sim \text{GHz}$
 $\rightarrow \text{THz} \rightarrow \text{PHz}$



Energy
(dissipated)



$E \sim \text{nJ}$
 $\rightarrow \text{fJ} \rightarrow \dots \rightarrow \text{zJ}$
Landauer limit

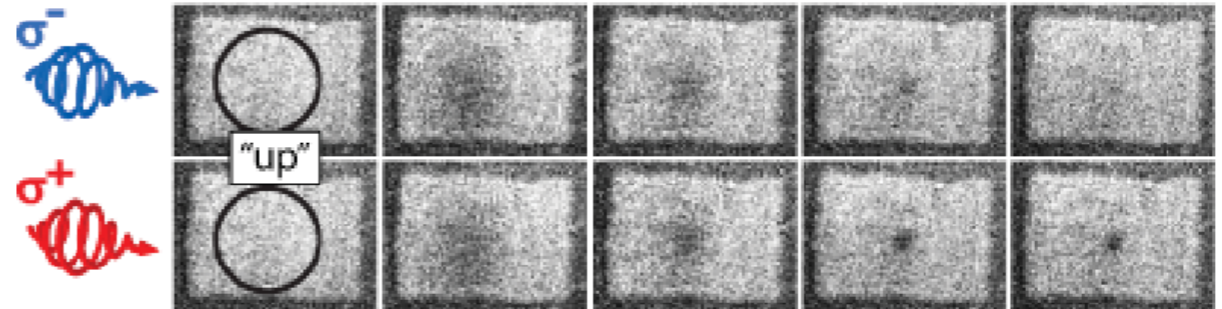


Potential of Ultrafast Magnetism

→ Faster than using electricity

All-optical switching of ferrimagnets as fast as 30 ps

Vahaplar *et al.*, PRL 2009

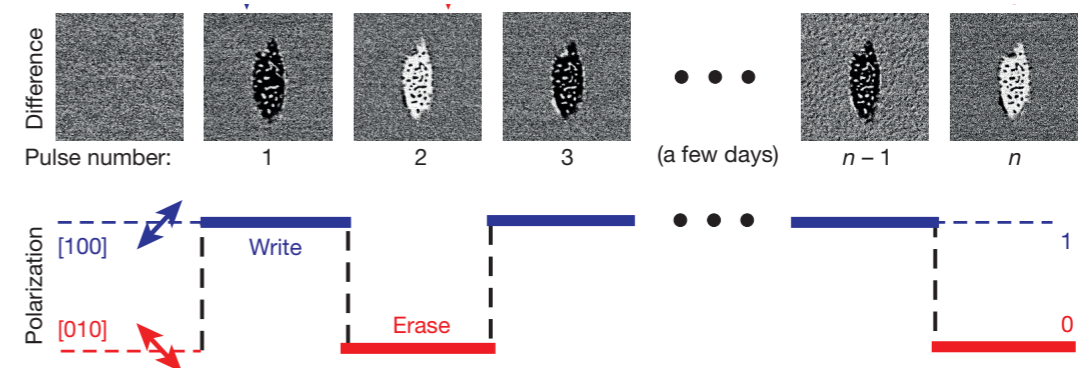


→ More energy efficient

Photomagnetic recording in oxides with projected

heat load down to 22 aJ/bit

Stupakiewicz *et al.*, Nature 2017



→ Possible even in technologically relevant materials

All-optical control in Co/Pt systems

Lambert *et al.*, Science 2014

First results

Experimental demonstration of artificial neural network using ultrafast optical control of Co/Pt thin films

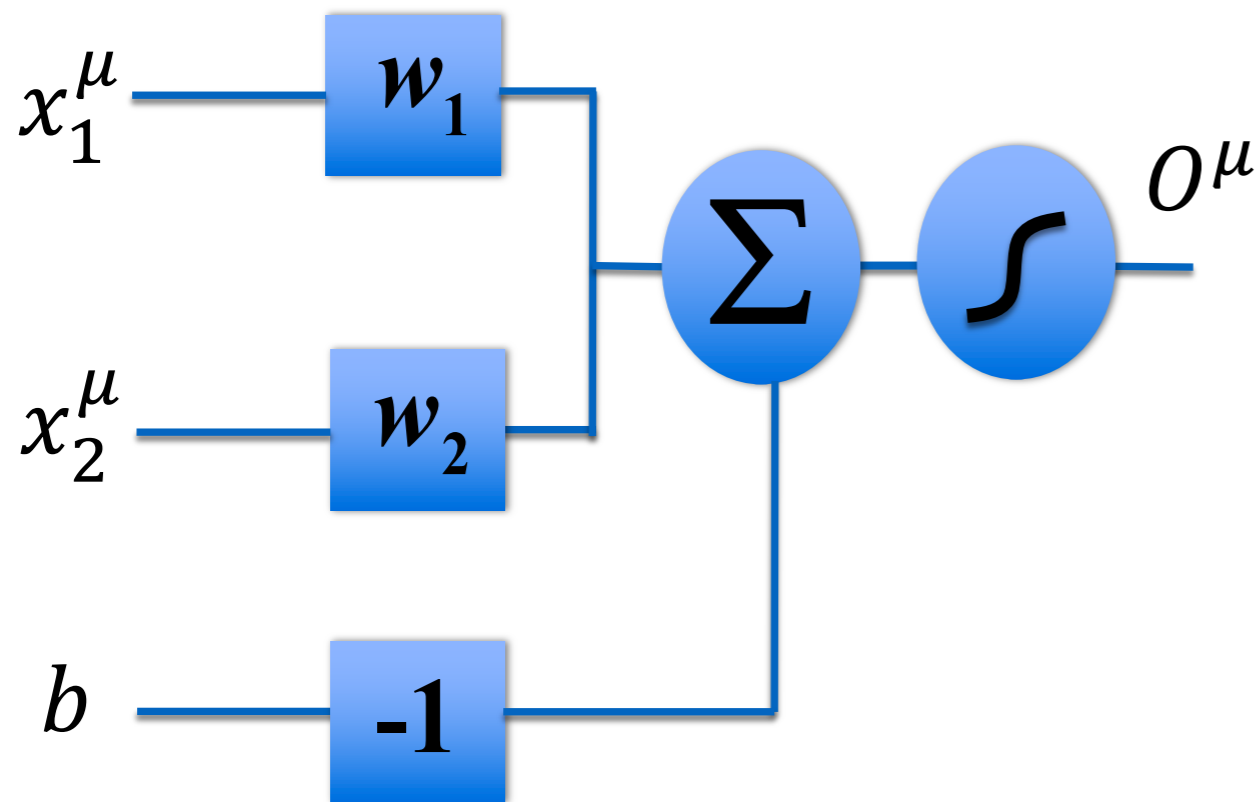
[A. Chakravarty, JHM, C.S. Davies, K. Yamada, A.V. Kimel and Th. Rasing](#)
Appl. Phys. Lett. 114, 192407 (2019)

Theoretical demonstration that artificial neural networks can simulate otherwise unsolvable nonequilibrium quantum dynamics of magnetic materials

[G. Fabiani and JHM](#)
SciPost Physics 7, 004 (2019)

Supervised learning

Perceptron model

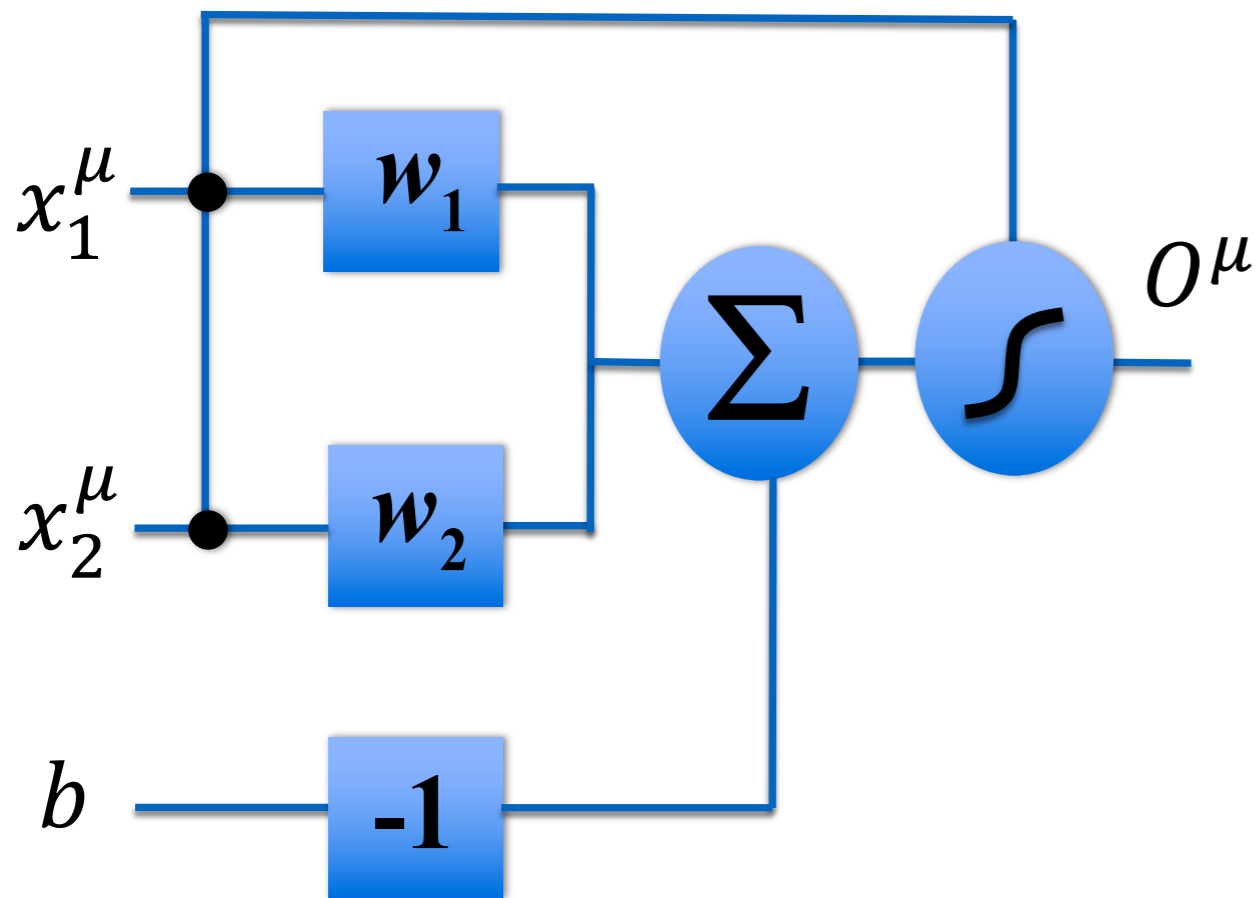


$$O^\mu = \text{sign} \left(\sum_i y_i^\mu - b \right)$$
$$y_i^\mu = w_i x_i^\mu$$

Supervised learning

Perceptron model

$$E = \text{sign}(O_d^\mu - O_d)$$



$$O^\mu = \text{sign} \left(\sum_i y_i^\mu - b \right)$$
$$y_i^\mu = w_i x_i^\mu$$

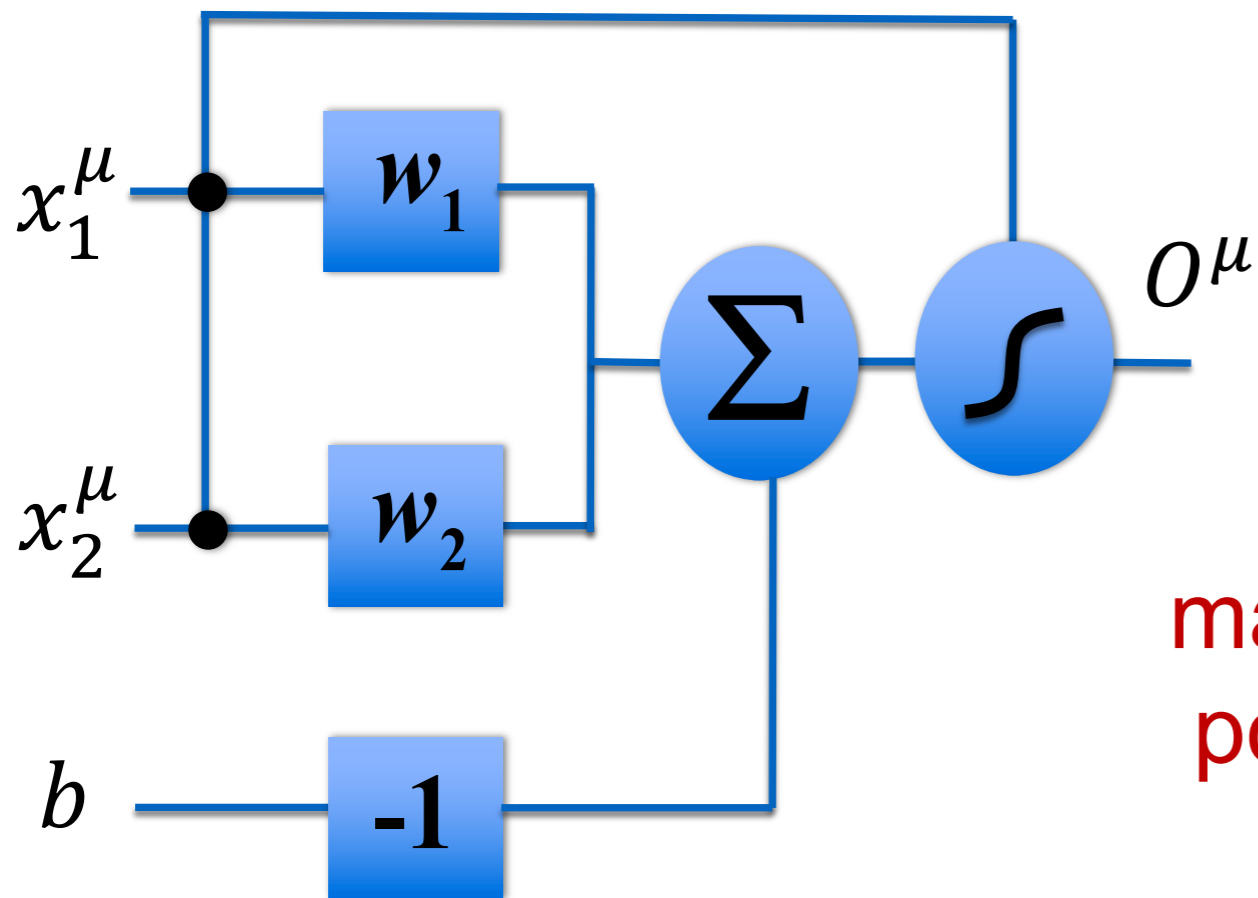
$$\Delta w_i = \eta E x_i^\mu$$

Learning requires global feedback only!

Supervised learning

Perceptron model

$$E = \text{sign}(O_d^\mu - O_d)$$



$$O^\mu = \text{sign} \left(\sum_i y_i^\mu - b \right)$$

$$y_i^\mu = w_i x_i^\mu$$

magneto-optical probing
polarization microscope

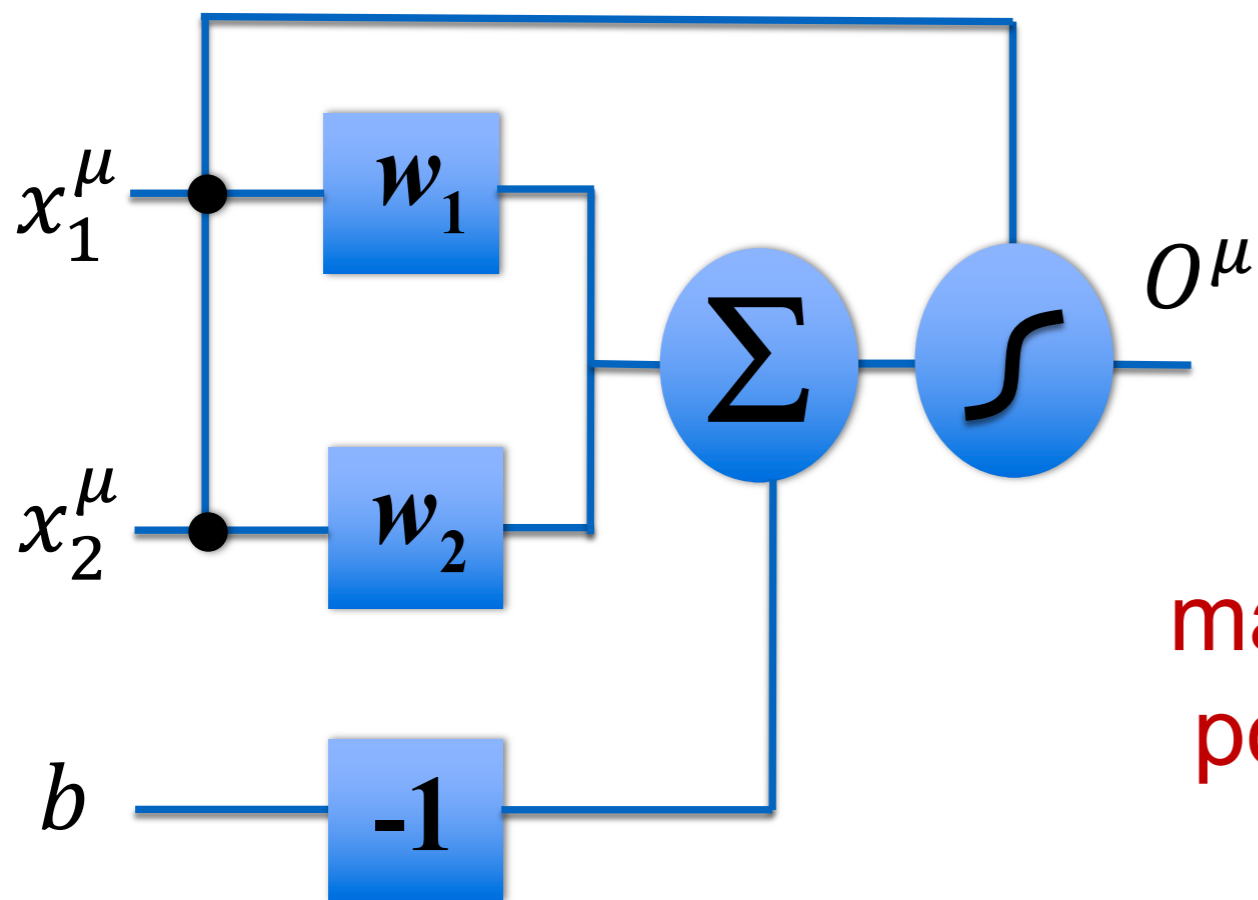
$$\Delta w_i = \eta E x_i^\mu$$

x_i^μ light/no light

Supervised learning

Perceptron model

$$E = \text{sign}(O_d^\mu - O_d)$$



$$O^\mu = \text{sign} \left(\sum_i y_i^\mu - b \right)$$

$$y_i^\mu = w_i x_i^\mu$$

magneto-optical probing
polarization microscope

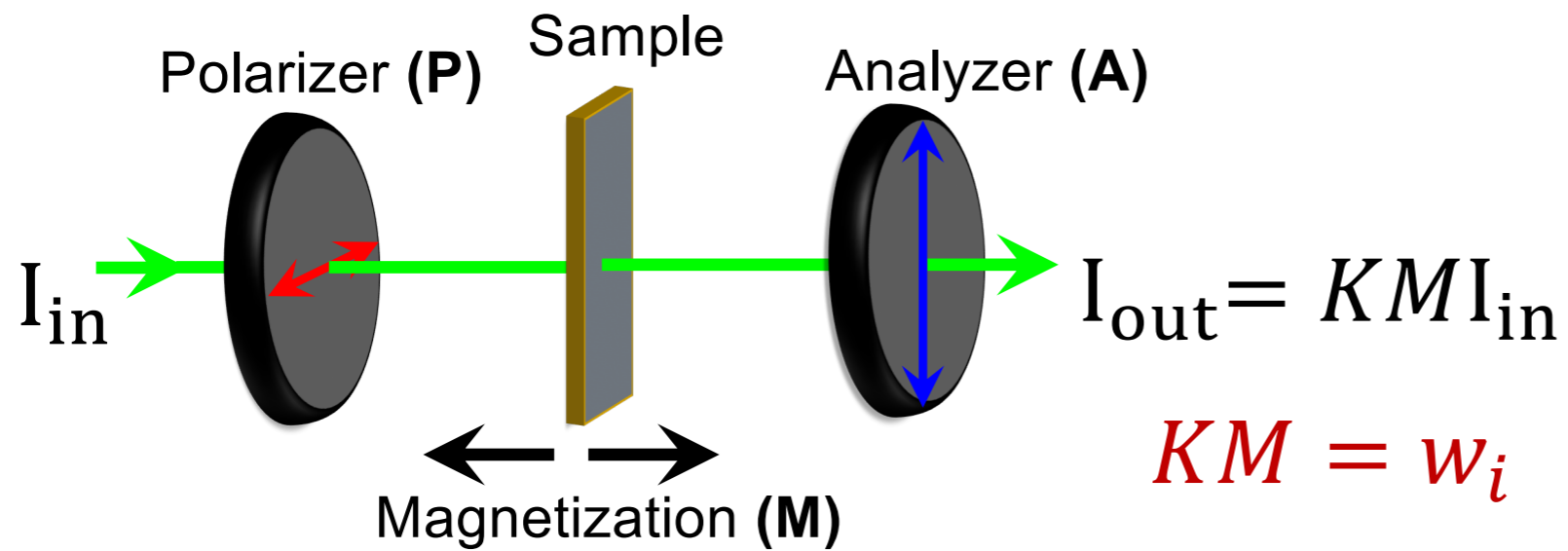
$$\Delta w_i = \eta E x_i^\mu$$

x_i^μ light/no light

η number of pump pulses
 E right/left circular polarization

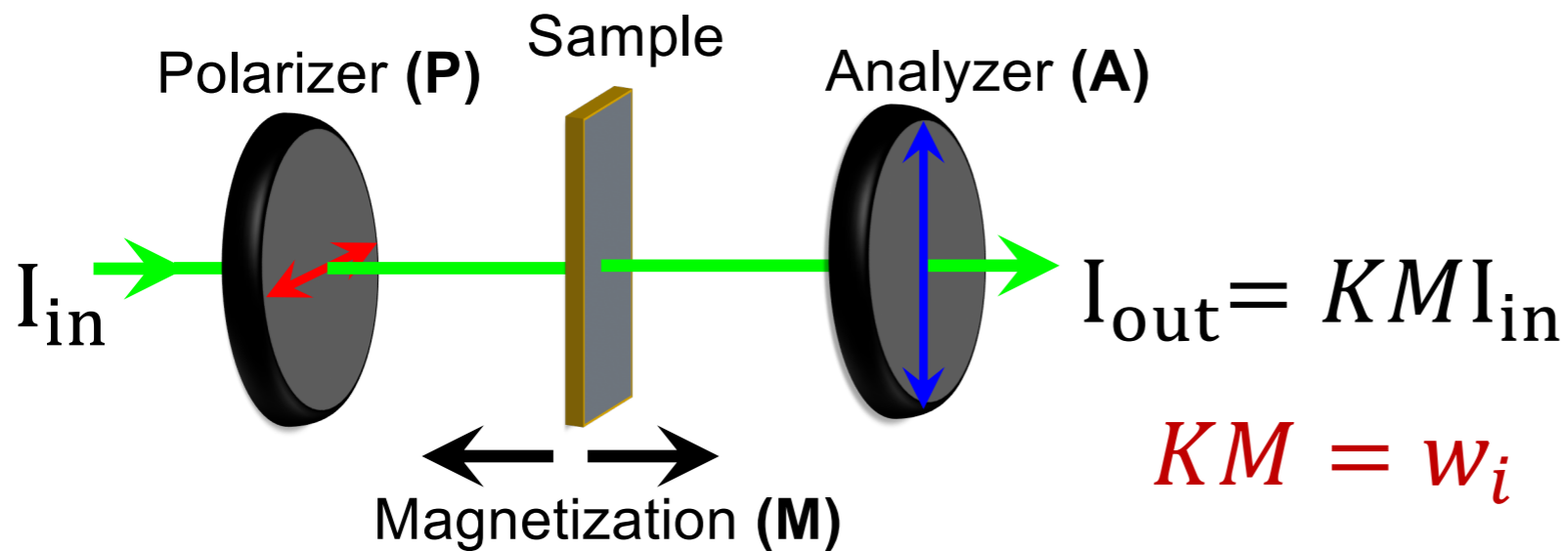
Magneto-optical synapses

Multiplication realized by polarization microscope



Magneto-optical synapses

Multiplication realized by polarization microscope



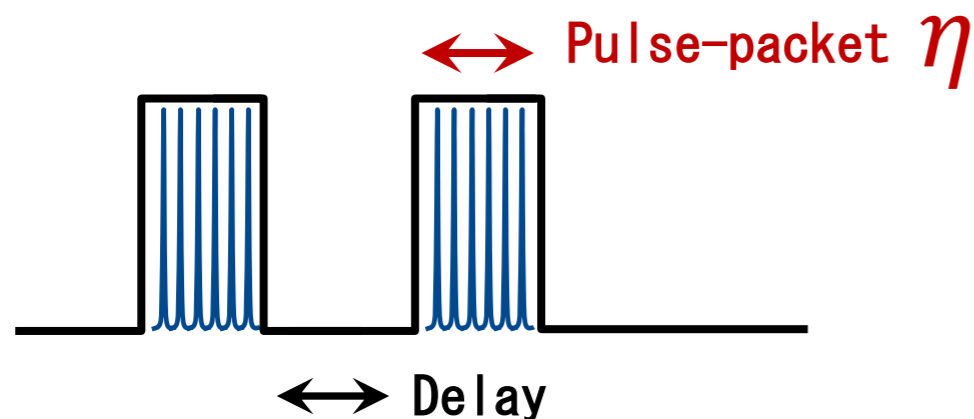
Writing



Erasing



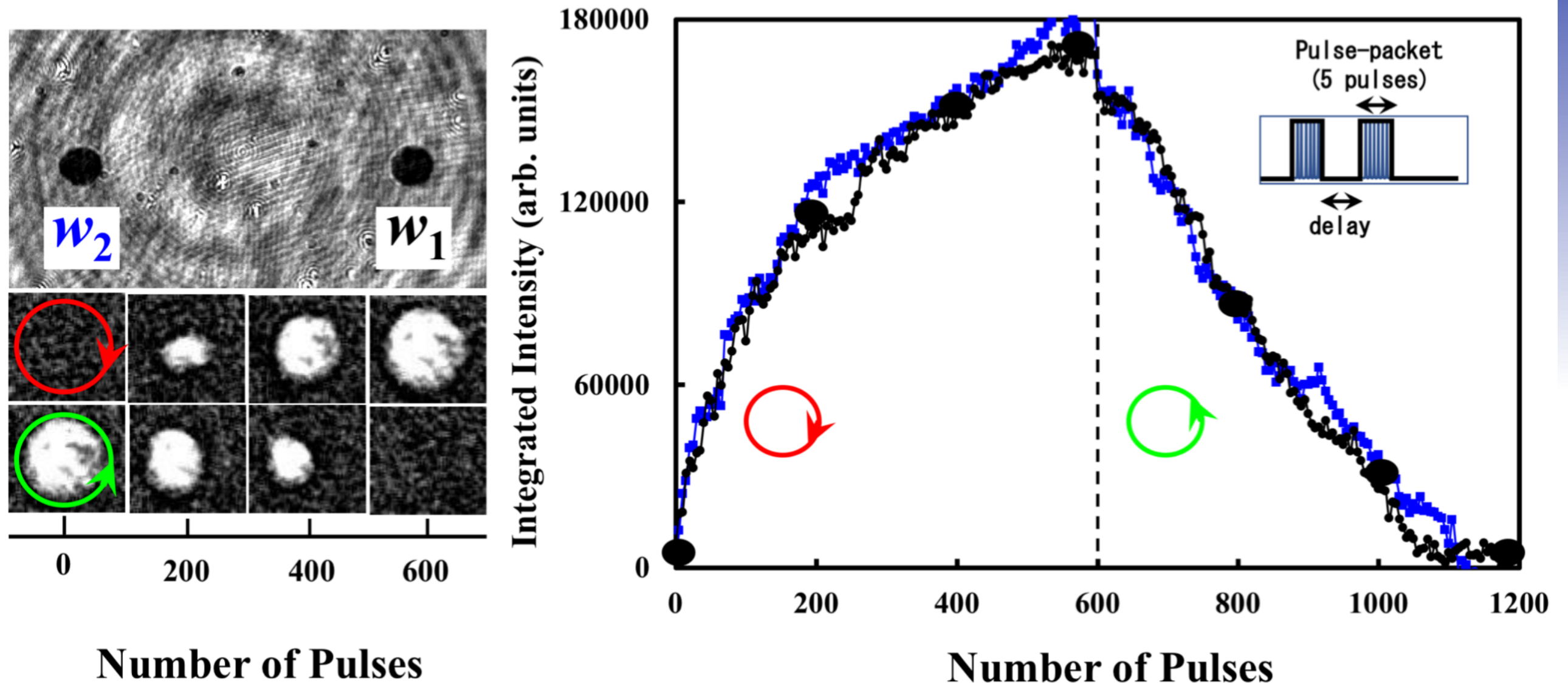
Gradual changes of M needed: Co/Pt thin films



Fullerton, Magnin, Aeschlimann et al., Science 2014

R. Medapalli et al, Phys. Rev.B 96, 224421 (2017)

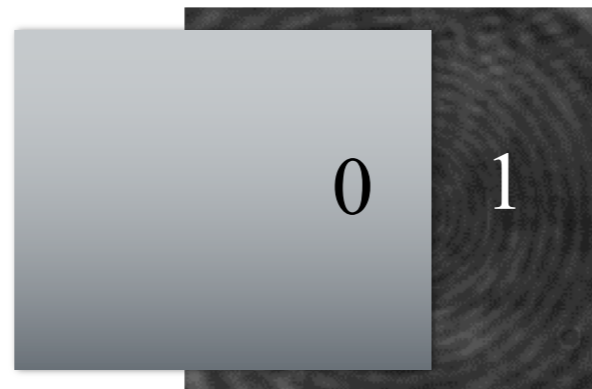
Continuously variable weights



Reproducible adaptation of weights with
circularly polarized laser pulses
pulse width 4 ps, 5 pulses/package fluence 1.3mJ/cm²

Artificial neural network

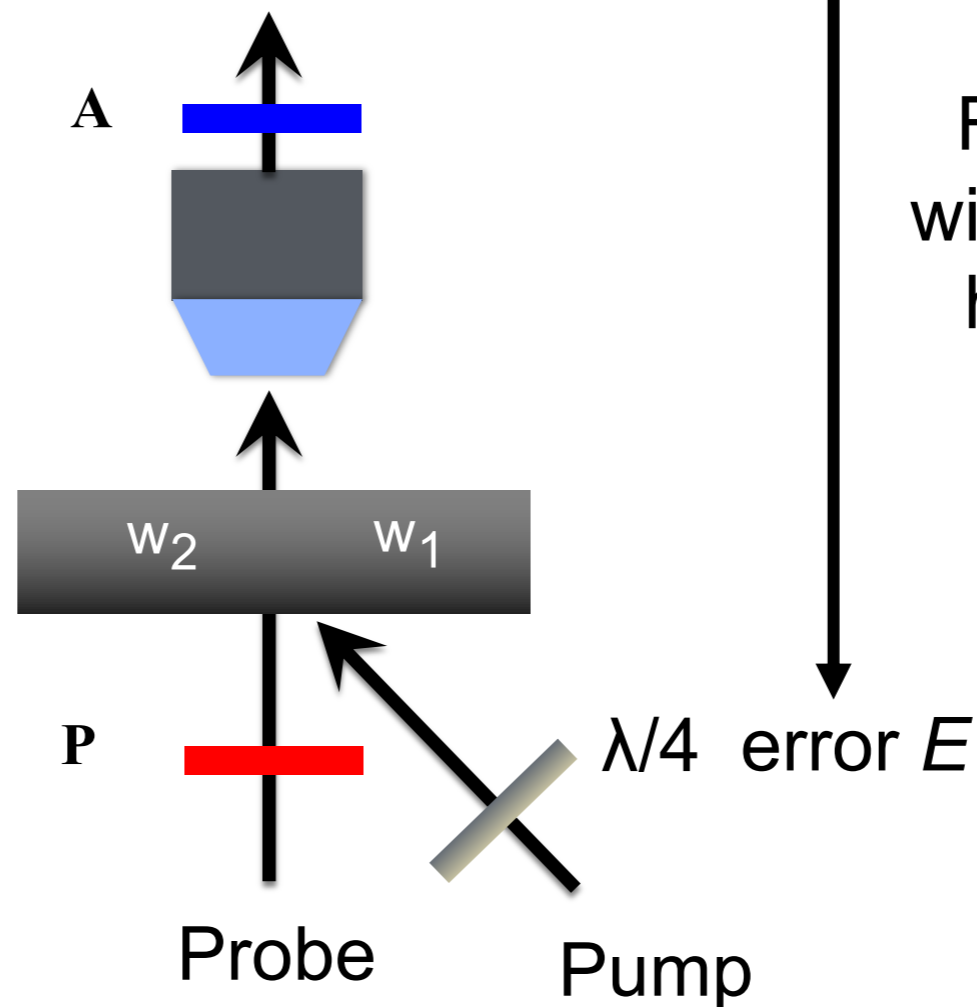
Translational Shutter:
inputs 0/1
(light/no light)



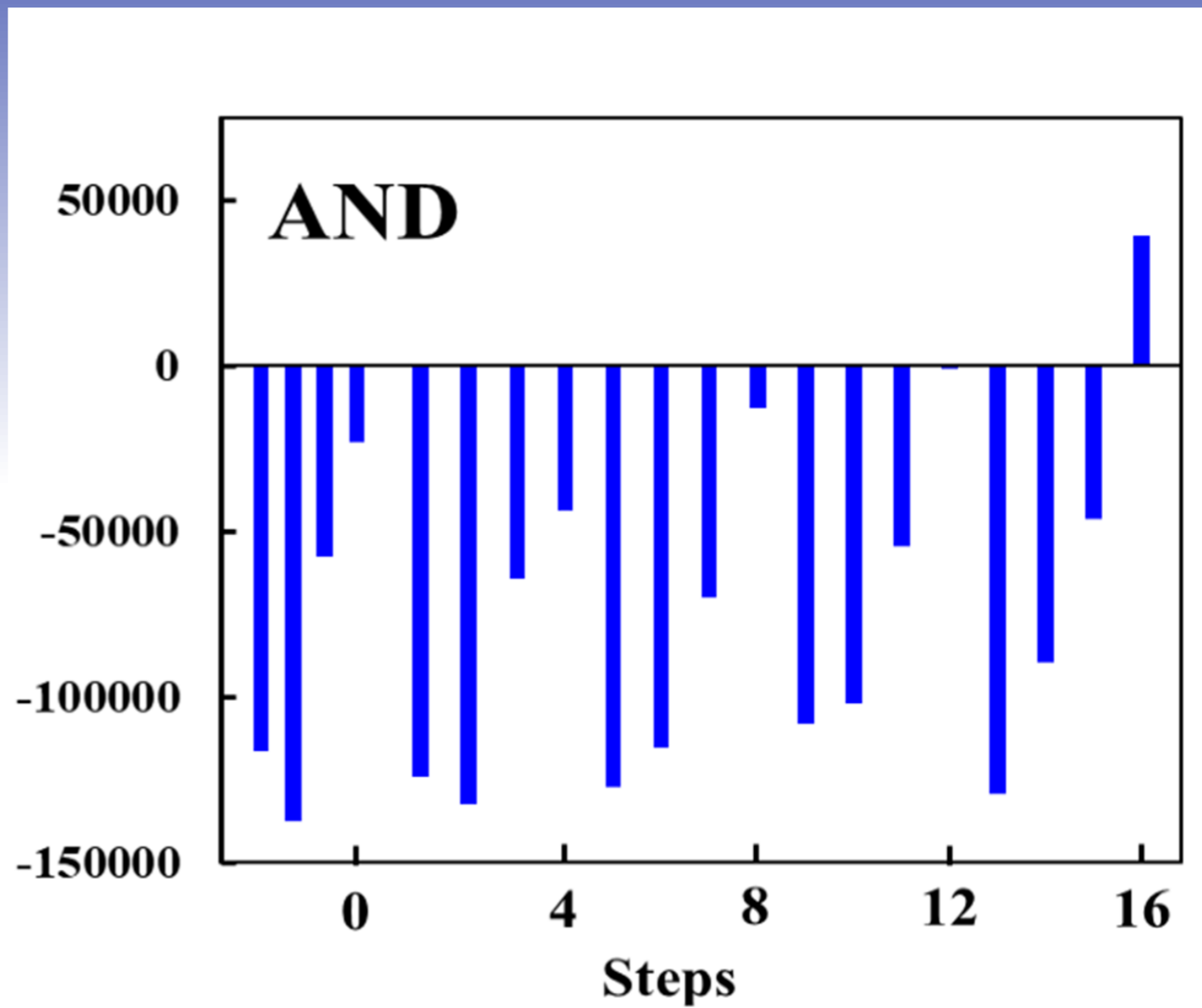
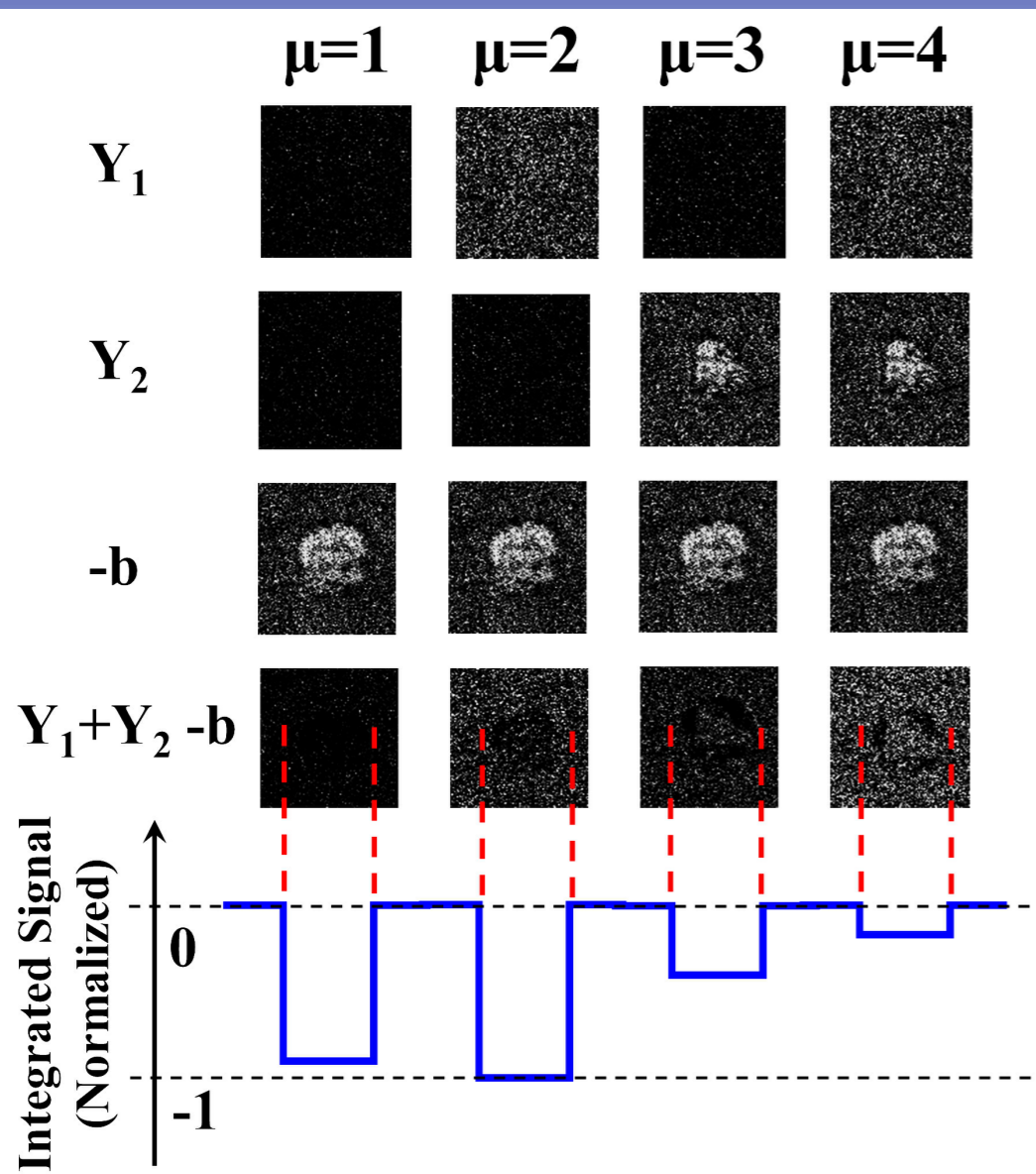
CCD camera

Feedback
with external
hardware

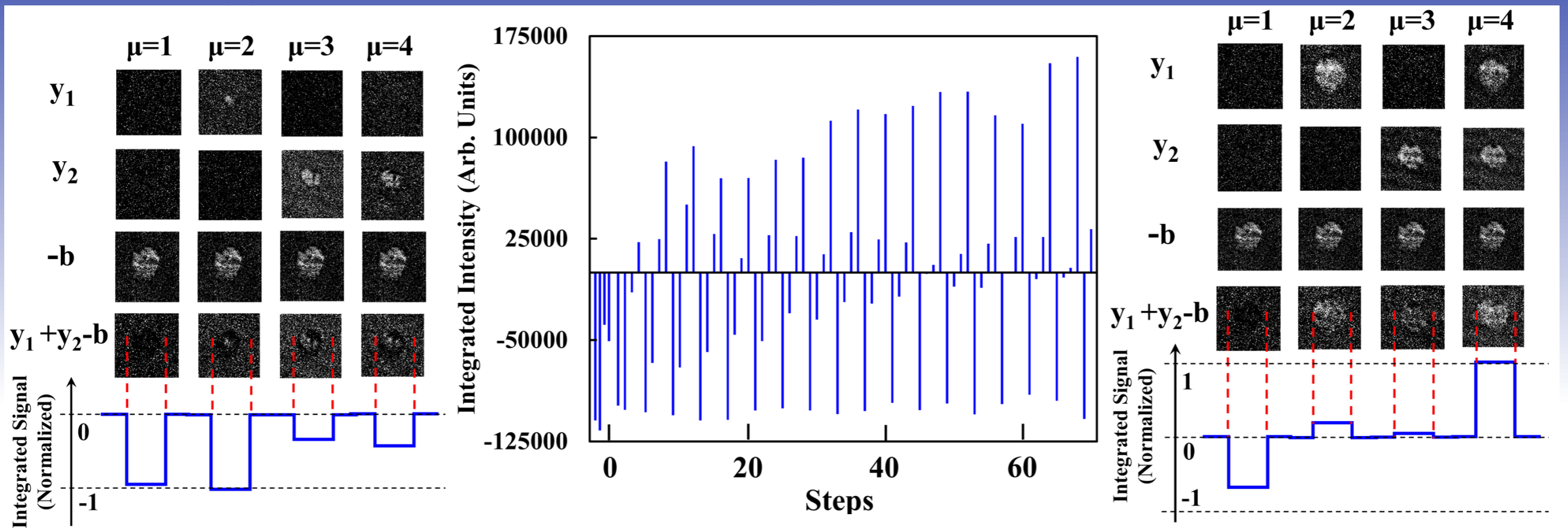
Translation stage



AND function



OR function



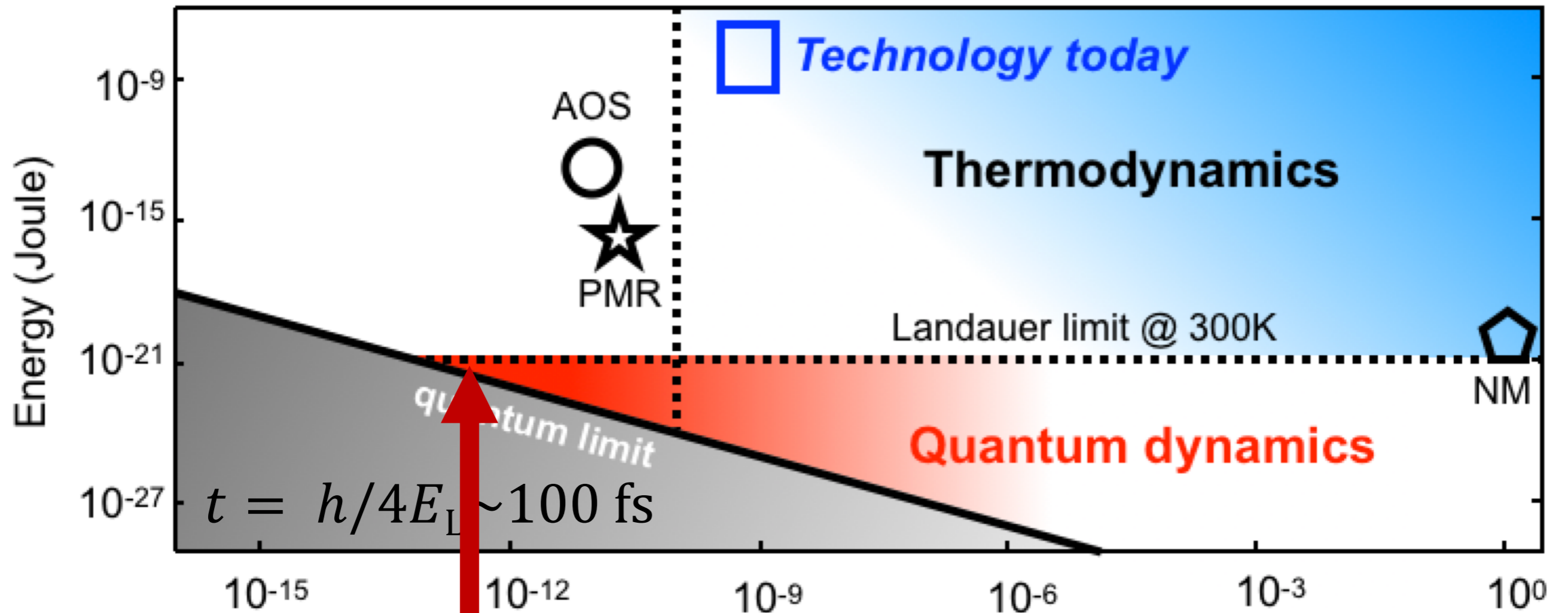
Opto-magnetic neural network

- ✓ Realization of opto-magnetic synapses using ultrashort laser pulses on Co/Pt films
- ✓ Supervised learning with opto-magnetic synapses
- ✓ Optimization with global feedback only
 - No external storage needed
- ✓ Energy absorbed: 65 pJ/synapse/step (1.125 μm)
 - Extrapolates to 20 fJ/synapse/step (20 nm)

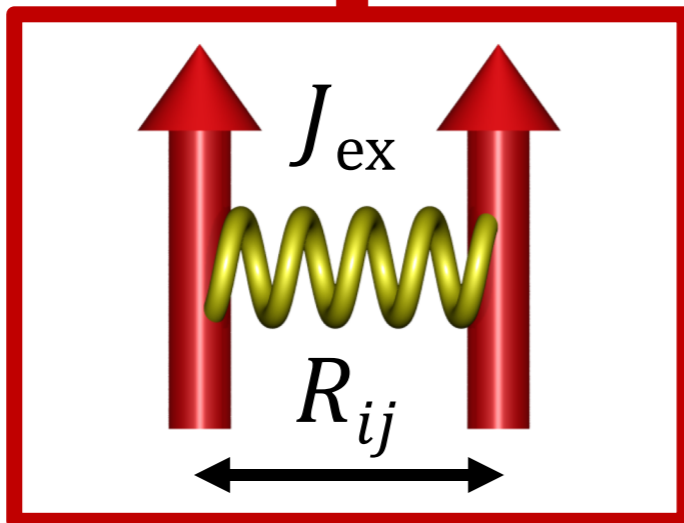
Appl. Phys. Lett. 114, 192407 (2019)

Next steps: more/smaller, implement backpropagation

Physical limits of computing



AOS: Stanciu
 PMR: A. Stuyve
 NM: J. Hong



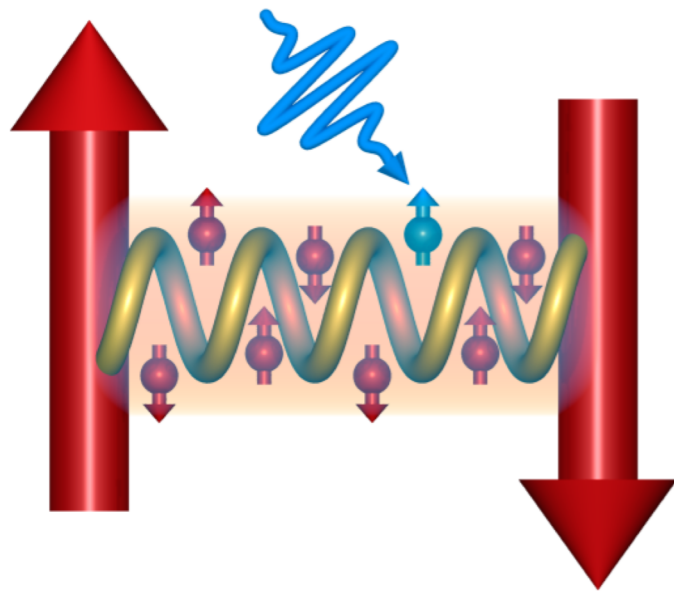
$$E_L = k_B T \ln 2 = 3 \text{ zJ} = 20 \text{ meV}$$

$$E_{\text{ex}} = J_{\text{ex}} \vec{S}_1 \vec{S}_2 \sim 20 \text{ meV}$$

interaction range $R_{ij} \sim \text{nm}$

Manipulating magnetism by ultrafast control of the exchange interaction

J H Mentink 2017 *J. Phys.: Condens. Matter* **29** 453001



Control Jex in Mott-Hubbard systems

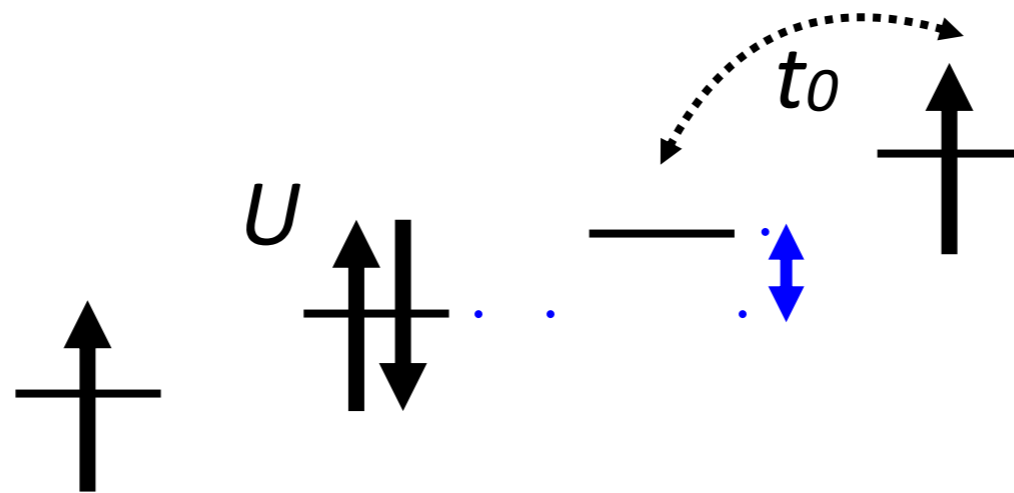
- Photo-doping
- *Non-resonant driving*

Manipulation of magnetism

- Excitation of spin precession
- Ultrafast cooling`
- Effective time-reversal
- *Time-resolved two-magnon dynamics*

Single-band Hubbard model

$$\hat{H} = \underbrace{-t_0 \sum_{\langle ij \rangle \sigma} e^{-i[\Phi_i(t) - \Phi_j(t)]} \hat{c}_{i\sigma}^\dagger \hat{c}_{j\sigma}}_{H_{\text{kin}}(t)} + \underbrace{U \sum_i \hat{n}_{i\uparrow} \hat{n}_{i\downarrow}}_V$$



$$\mathcal{E} = \frac{eaE_0}{\hbar\omega}$$

$$E_0(t) = E_0 \cos(\omega t)$$

$$\Phi_i(t) - \Phi_j(t) = -\mathcal{E}(\hat{e} \cdot \hat{r}_{ij}) \sin(\omega t)$$

Effective low-energy Hamiltonian

$$H = H_{\text{kin}}(t) + V$$

Hilbert space $\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_1$

Low-energy

High-energy

V has only

Projectors

$$\mathcal{P}_0$$

$$\mathcal{P}_1 = 1 - \mathcal{P}_0$$

$$\mathcal{P}_1 V \mathcal{P}_1 = V_{11}$$

Construct $S(t)$ such that no mixing takes place at each time

$$\tilde{H} = e^{iS(t)} (H - i\partial_t) e^{-iS(t)}$$

$$S(t) = S^{(1)}(t) + S^{(2)}(t) + \dots$$

$$H_{\text{kin}}(t) + [iS^{(1)}(t), V] - \partial_t S^{(1)}(t) = 0$$

Bukov et al., *PRL* **116**, 125301 (2016)

Canovi et al., *PRE* **93** 012130 (2016)

Eckstein et al., arXiv:1703.03269 (2017)

Light-perturbation to J_{ex}

$$\tilde{H}^{(2)} = e^{i\hat{S}^{(1)}(t)} \left(\hat{H} - i\partial_t \right) e^{-i\hat{S}^{(1)}(t)} \longrightarrow H_0 + \delta H$$

Time-averaging

Half-filling (one electron per site)

$$H_0 = J \sum_{\langle ij \rangle} \mathbf{S}_i \cdot \mathbf{S}_j, \quad J = 2t_0^2/U$$

Simple cubic lattice, weak fields $\mathcal{E} \ll 1$

$$\delta H = \Delta J \sum_{\langle ij \rangle} (\hat{e} \cdot \hat{r}_{ij})^2 \mathbf{S}_i \cdot \mathbf{S}_j, \quad \Delta J = \frac{t_0^2}{2U} \frac{\omega^2 \mathcal{E}^2}{U^2 - \omega^2}$$

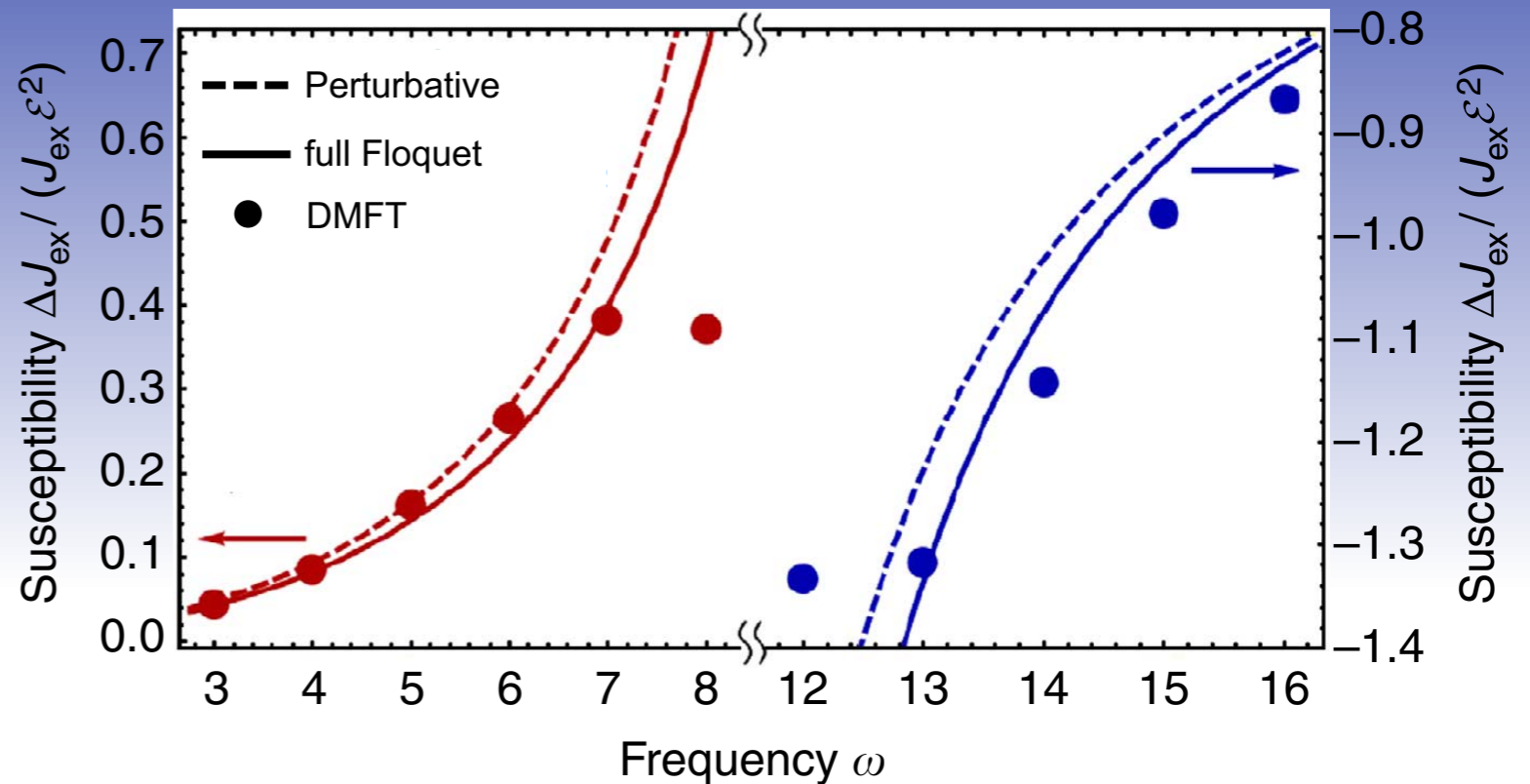
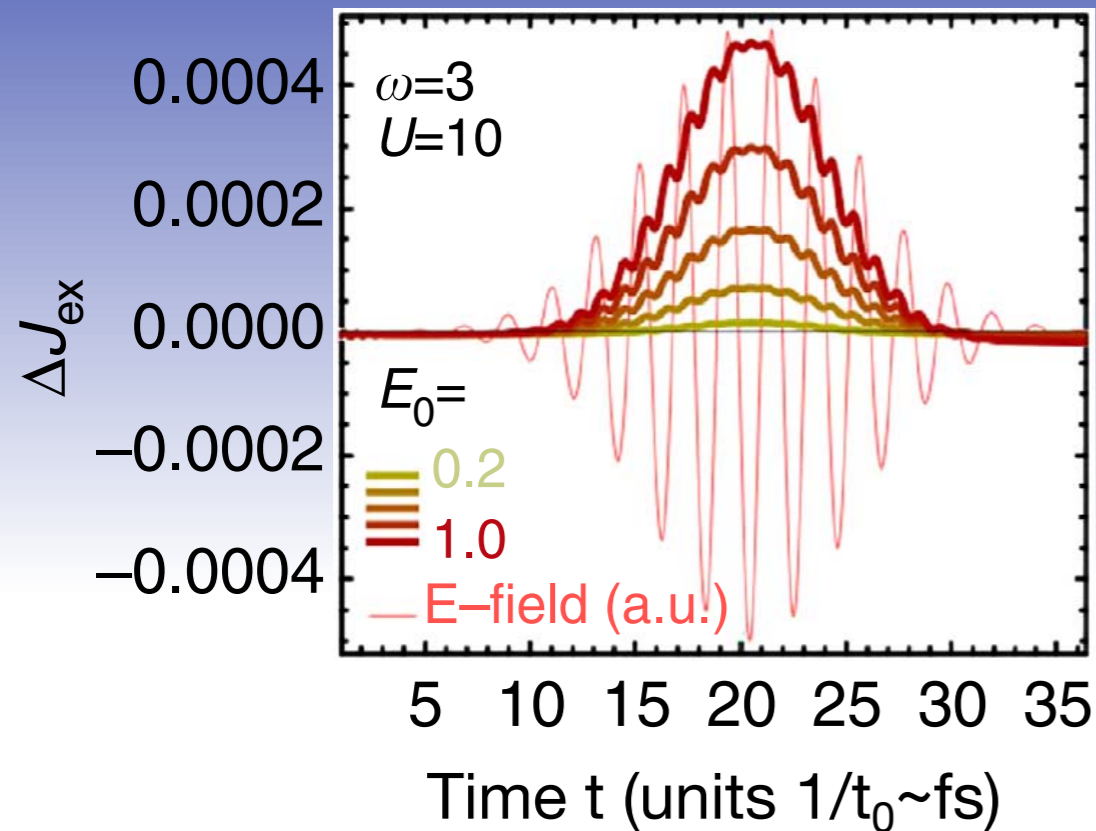
$$\mathcal{E} = \frac{eaE_0}{\hbar\omega} \quad \text{Only bonds along the field perturbed} \quad J_{\text{ex}} \text{ stronger/weaker below/above gap}$$

Floquet theory: Mentink et al., *Nat. Commun.* 2015

High-frequency expansion: Itin, *PRL* 2015

Time-dependent Schrieffer-Wolff: Bukov et al., *PRL* 2016

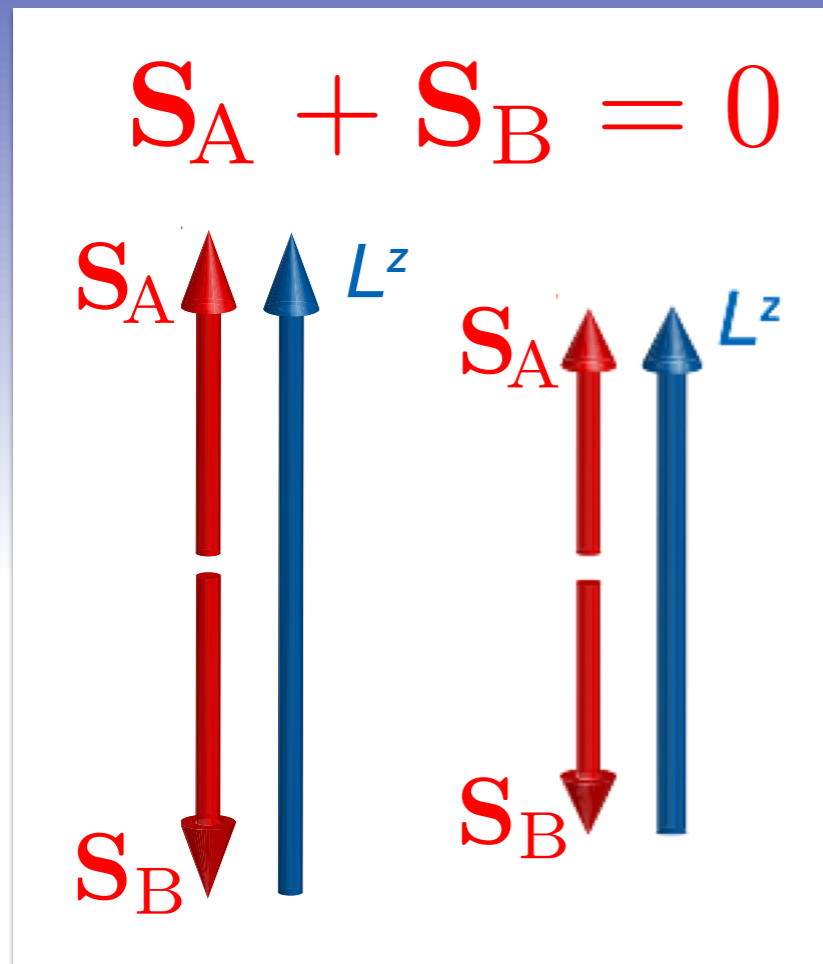
Nonequilibrium DMFT



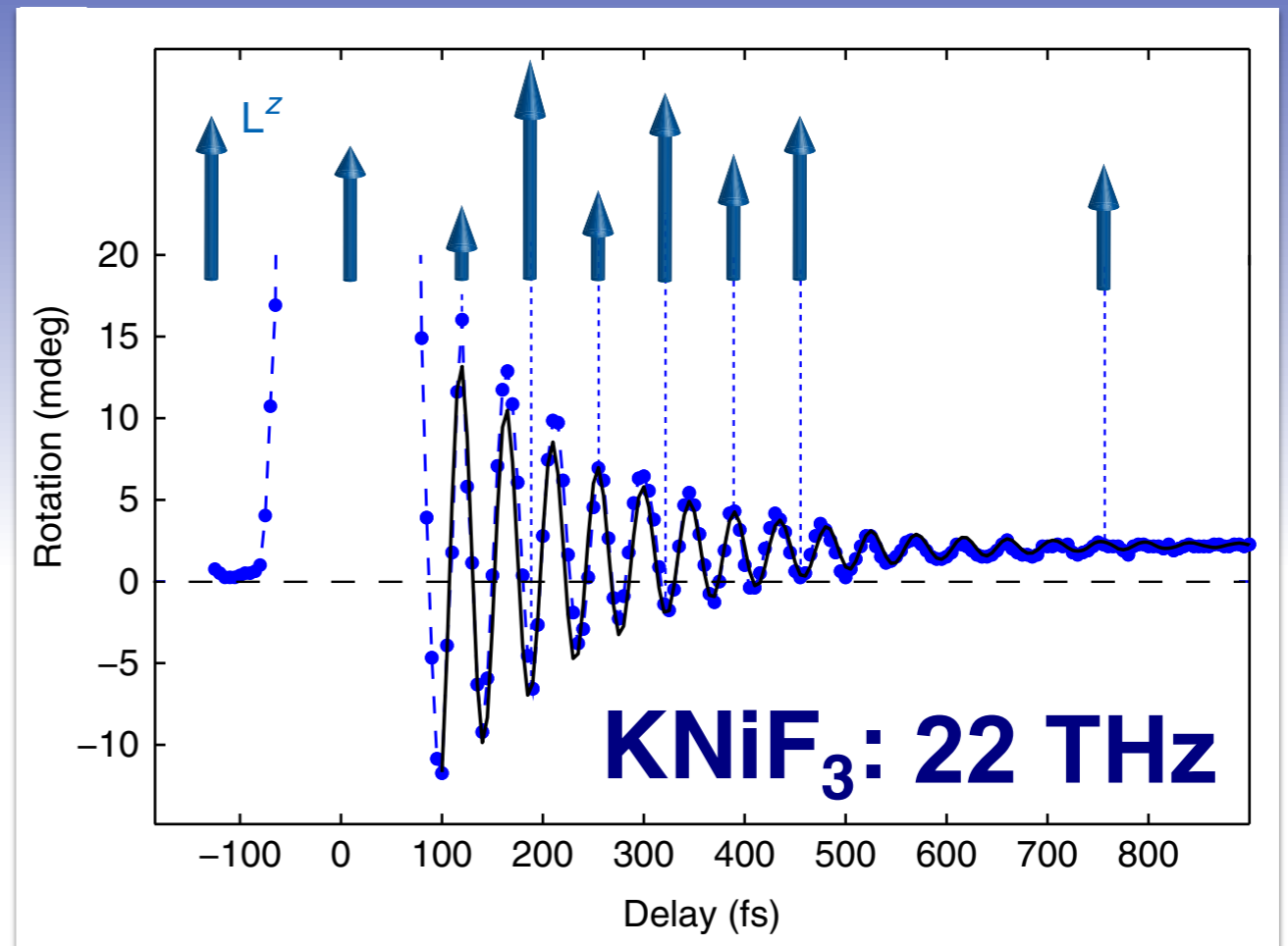
Non-equilibrium DMFT:
Aoki et al., RMP 2014

More general exchange formulas:
J.H. Mentink and M. Eckstein, PRL 2014
R.V. Mikhaylovskiy, et al. *Nat. Commun.* 2015

Quantum spin dynamics in solid state



$$L = S_A - S_B$$



- **Coherent** *longitudinal* dynamics of magnetic order parameter

Magnon-pair description

Holstein-Primakov + Bogoliubov transform

$$H_0 \sim \sum_{\mathbf{k}} \omega_{\mathbf{k}} \quad 2\hat{K}_{\mathbf{k}}^z$$

$$\omega_{\mathbf{k}} = zJS \sqrt{1 - \gamma_{\mathbf{k}}^2}$$

Simple cubic lattice

$$\gamma_{\mathbf{k}} = \frac{1}{z} \sum_{\delta} \exp(i\mathbf{k} \cdot \delta) = \frac{1}{3} \sum_{i=xyz} \cos(k_i a)$$

$$\delta H \sim \sum_{\mathbf{k}} \delta\omega_{\mathbf{k}} \quad 2\hat{K}_{\mathbf{k}}^z$$

$$\delta\omega_{\mathbf{k}} = z\Delta JS \frac{1 - \xi_{\mathbf{k}}\gamma_{\mathbf{k}}}{\sqrt{1 - \gamma_{\mathbf{k}}^2}}$$

$$V_{\mathbf{k}} = z\Delta JS \frac{\xi_{\mathbf{k}} - \gamma_{\mathbf{k}}}{\sqrt{1 - \gamma_{\mathbf{k}}^2}}$$

$$+ V_{\mathbf{k}} \left[\hat{K}_{\mathbf{k}}^- + \hat{K}_{\mathbf{k}}^+ \right]$$

$$\xi_{\mathbf{k}} = \frac{1}{z} \sum_{\delta} (\hat{e} \cdot \hat{\delta})^2 \exp(i\mathbf{k} \cdot \delta) = \frac{1}{3} \sum_{i=xyz} \hat{e}_i^2 \cos(k_i a)$$

Magnon-pair commutators

$$\hat{K}_{\mathbf{k}}^z = (\hat{a}_{\mathbf{k}}^\dagger \hat{a}_{\mathbf{k}} + \hat{\beta}_{-\mathbf{k}}^\dagger \hat{\beta}_{-\mathbf{k}} + 1)/2, \quad \text{Bose commutator relations}$$

$$\hat{K}_{\mathbf{k}}^+ = \hat{a}_{\mathbf{k}}^\dagger \hat{\beta}_{-\mathbf{k}}^\dagger, \quad [\alpha_{\mathbf{k}}, \alpha_{\mathbf{k}}^\dagger] = 1$$

$$\hat{K}_{\mathbf{k}}^- = \hat{a}_{\mathbf{k}} \hat{\beta}_{-\mathbf{k}}. \quad [\beta_{\mathbf{k}}, \beta_{\mathbf{k}}^\dagger] = 1$$

$$[\hat{K}_{\mathbf{k}}^z, \hat{K}_{\mathbf{k}}^\pm] = \pm \hat{K}_{\mathbf{k}}^\pm, \quad [\hat{K}_{\mathbf{k}}^-, \hat{K}_{\mathbf{k}}^+] = 2\hat{K}_{\mathbf{k}}^z$$

SU(1,1), hyperbolic / Perelomov operators

Casimir invariant:

$$\hat{Q} = \frac{1}{2} \left(\hat{K}_{\mathbf{k}}^+ \hat{K}_{\mathbf{k}}^- + \hat{K}_{\mathbf{k}}^- \hat{K}_{\mathbf{k}}^+ \right) - \left(\hat{K}_{\mathbf{k}}^z \right)^2 = \frac{1}{4} (1 - \Delta_{\mathbf{k}}^2)$$

$$\Delta_{\mathbf{k}} = \hat{a}_{\mathbf{k}}^\dagger \hat{a}_{\mathbf{k}} - \hat{\beta}_{-\mathbf{k}}^\dagger \hat{\beta}_{-\mathbf{k}} \quad \text{Only equal changes of sublattices}$$

Free dynamics ($\Delta J=0 \rightarrow \mathbf{V}_k=0$)

$$[\hat{K}_k^z, \hat{K}_k^\pm] = \pm \hat{K}_k^\pm, \quad [\hat{K}_k^-, \hat{K}_k^+] = 2\hat{K}_k^z$$

$$\hat{K}_k^\pm(t) = \hat{K}_k^\pm e^{\pm i2\omega_k t}, \quad \hat{K}_k^z(t) = \hat{K}_k^z$$

$$\sum_{\langle i,j \rangle} \hat{S}_i^z(t) \hat{S}_j^z(t) = -\frac{zN}{2} S(S+1) + zS \sum_k g_k 2\hat{K}_k^z - zS \sum_k \gamma_k g_k [\hat{K}_k^+(t) + \hat{K}_k^-(t)]$$

$$\frac{1}{2} \sum_{\langle i,j \rangle} \hat{S}_i^+(t) \hat{S}_j^-(t) + \hat{S}_i^-(t) \hat{S}_j^+(t) = zS \sum_k \gamma_k^2 g_k 2\hat{K}_k^z + zS \sum_k \gamma_k g_k [\hat{K}_k^+(t) + \hat{K}_k^-(t)]$$

Ising and spin-flip terms play role of kinetic and potential energy

$$\hat{L}_z(t) = \frac{NS}{2} - \frac{1}{zS} \sum_{\langle i,j \rangle} \hat{S}_i^z(t) \hat{S}_j^z(t).$$

Longitudinal dynamics at frequencies $2\omega_k$

Non-classical magnon dynamics

Interaction representation

$$|\Psi(t > 0)\rangle = e^{i\tau \sum_{\mathbf{k}} [\bar{V}_{\mathbf{k}} (\hat{K}_{\mathbf{k}}^+ (t) + \hat{K}_{\mathbf{k}}^- (t)) + \delta\bar{\omega}_{\mathbf{k}} \hat{K}_{\mathbf{k}}^z (t)]} |\Psi_G\rangle$$

Ground state wave function

$$|\Psi_G\rangle = \prod_{\mathbf{k}} |0_{\mathbf{k}}\rangle |0_{-\mathbf{k}}\rangle$$

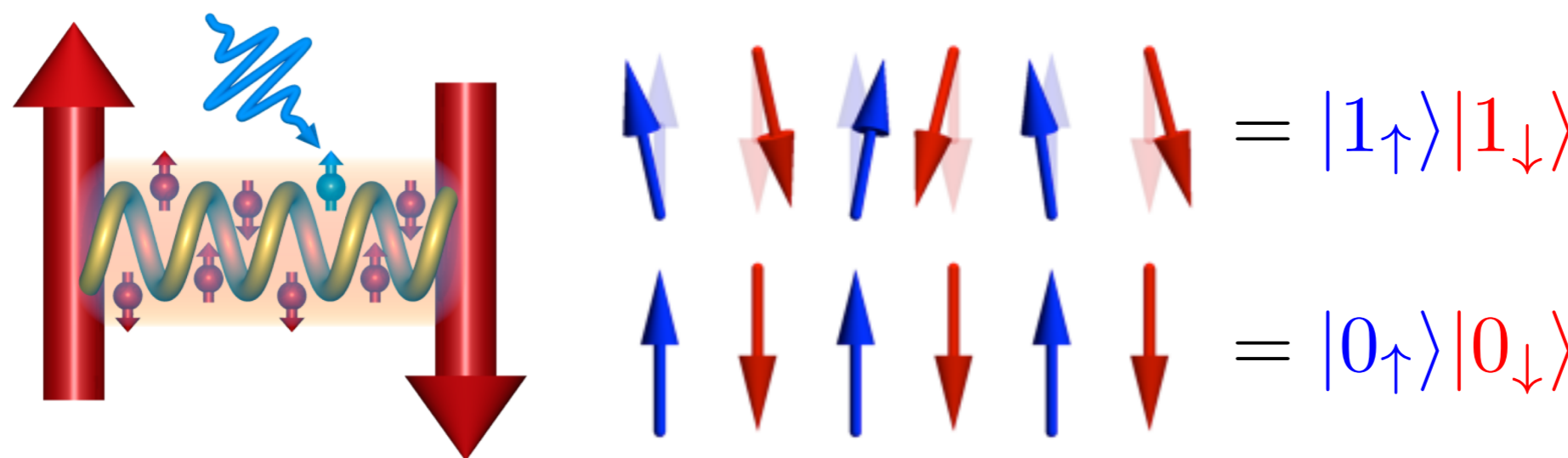
Two-particle coherence: **entangled magnons**

$$\Delta J \ll J$$

$$|\Psi_{\mathbf{k}}(t > 0)\rangle \approx |0_{\mathbf{k}}\rangle |0_{-\mathbf{k}}\rangle + iV_{\mathbf{k}} e^{i2\omega_{\mathbf{k}}t} |1_{\mathbf{k}}\rangle |1_{-\mathbf{k}}\rangle$$

Non-classical magnon dynamics

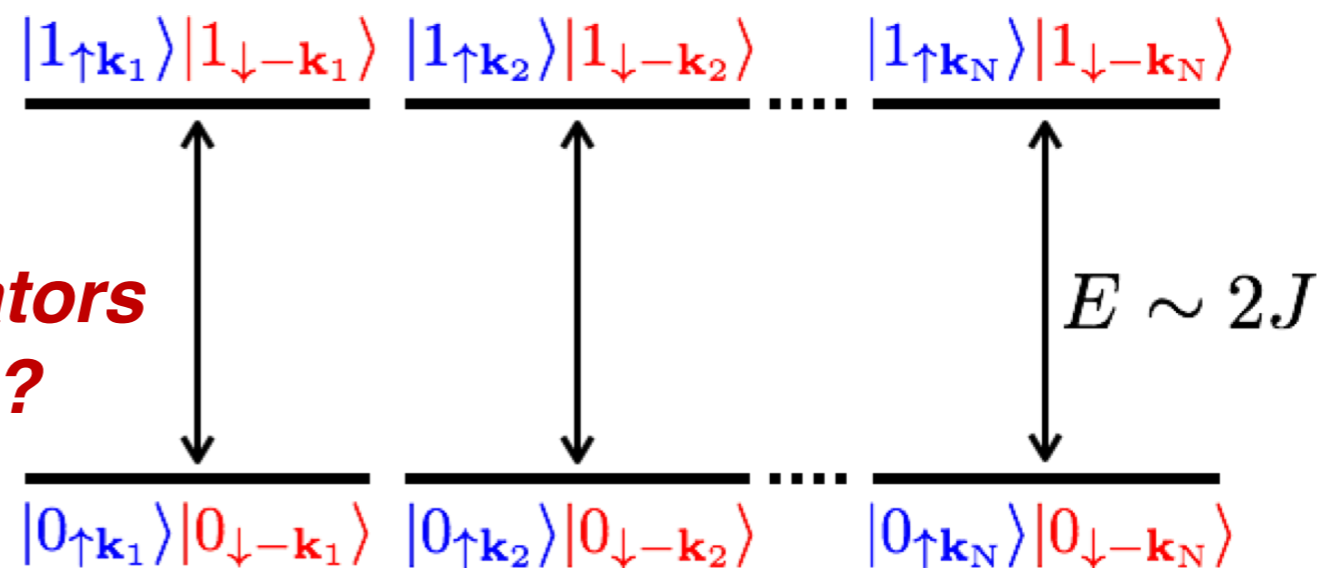
- Perturbation of J_{ex} causes excitation of **magnon-pairs**



- Femtosecond quantum spin dynamics in antiferromagnets

- $\hbar\omega_{2M} > k_B T$: survives at T_{ambient}

- **Quantum oscillators for every k**
 → **Large ensemble nano-oscillators**
 → **Non-linear? Synchronization?**



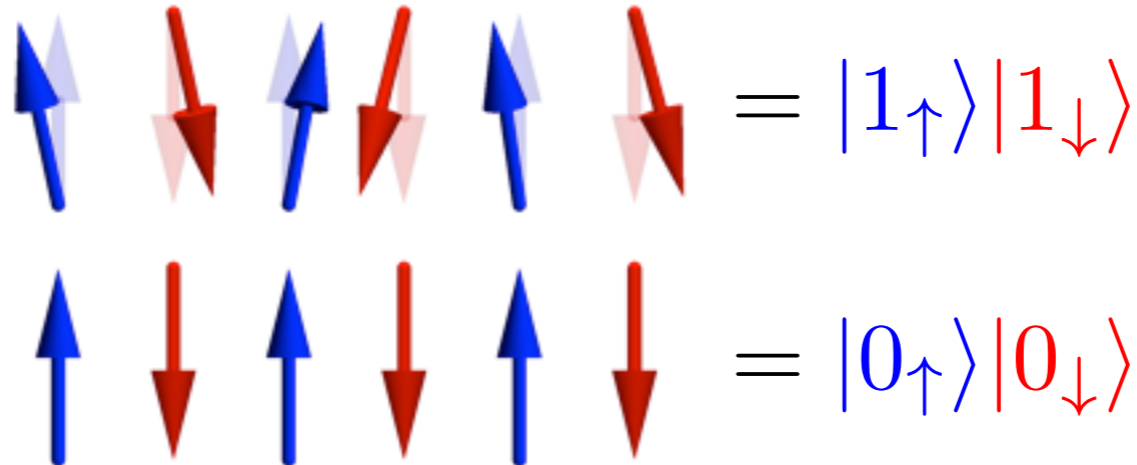
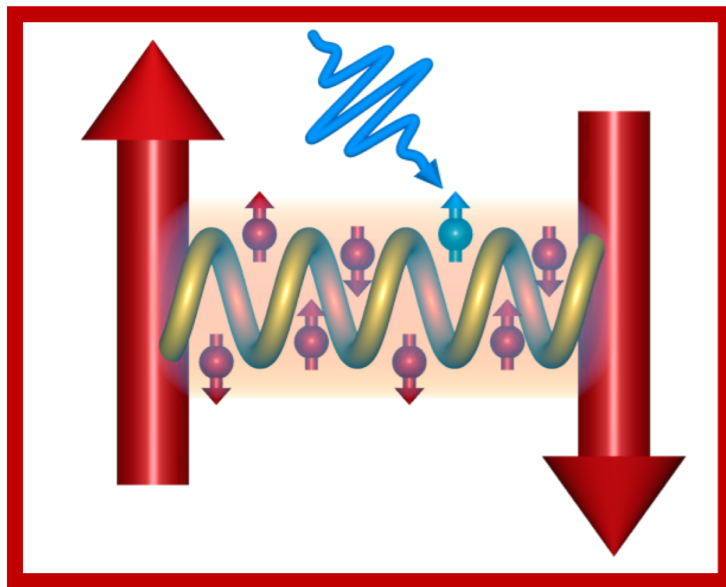
J.H. Mentink et al., Nat. Commun 2015

J.H. Mentink JPCM 2017

D. Bossini, E.V. Gomonay, J.H. Mentink, et al., PRB 100, 024428 (2019)

Non-classical magnon dynamics

- Perturbation of J_{ex} causes excitation of **magnon-pairs**



- **Nanoscale** quantum spin dynamics in antiferromagnets
- at T_{ambient}

$$|1_{\uparrow\mathbf{k}_1}\rangle |1_{\downarrow-\mathbf{k}_1}\rangle \quad |1_{\uparrow\mathbf{k}_2}\rangle |1_{\downarrow-\mathbf{k}_2}\rangle \quad |1_{\uparrow\mathbf{k}_N}\rangle |1_{\downarrow-\mathbf{k}_N}\rangle$$

- **Ultrafast Nanoscale Quantum Magnonics**

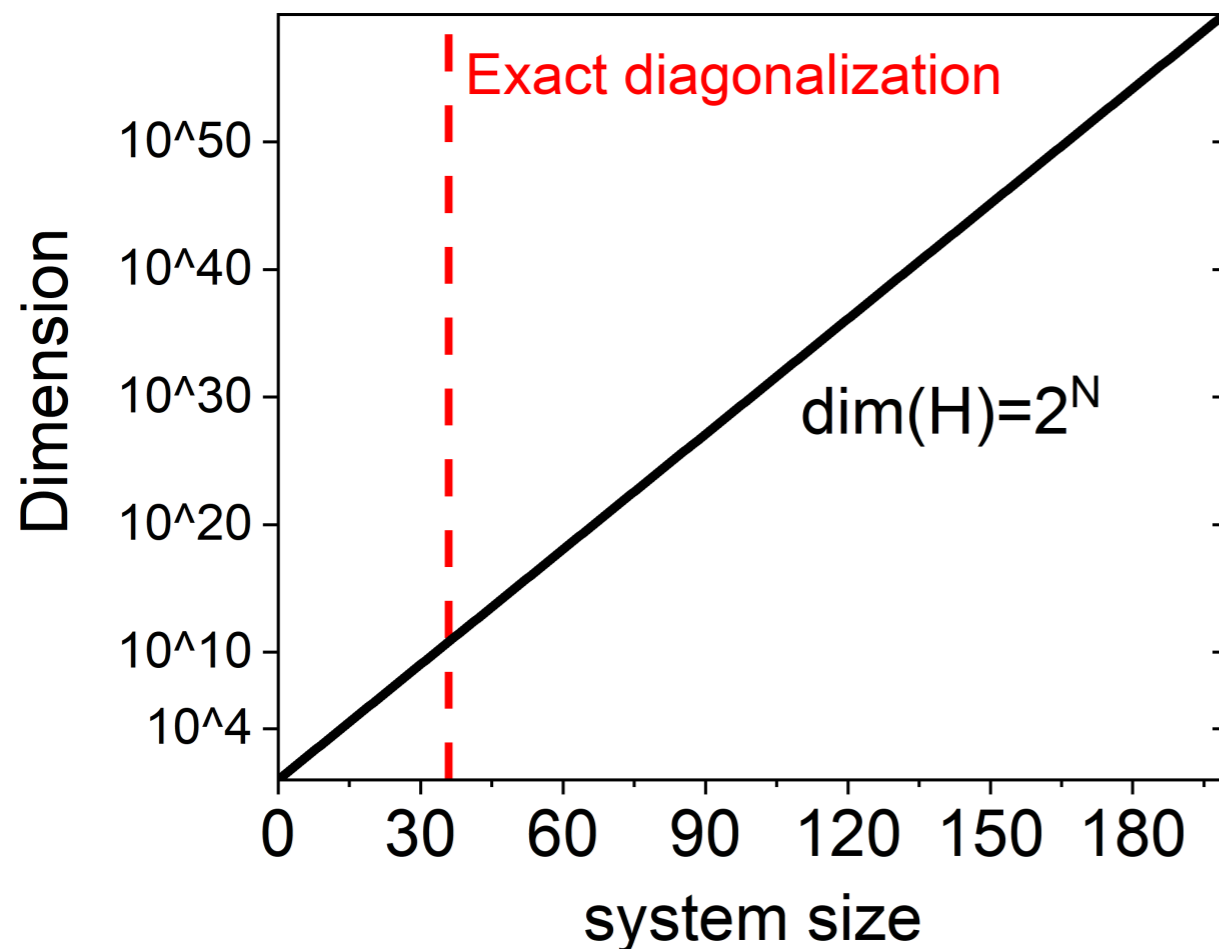
$$\langle S_i(t) S_j(t) \rangle$$

Quantum Many-Body Problem

Minimal model: 2D Heisenberg model $H_0 = J \sum_{\langle ij \rangle} \mathbf{S}_i \cdot \mathbf{S}_j$

Challenge: non-local correlations both in space and time!

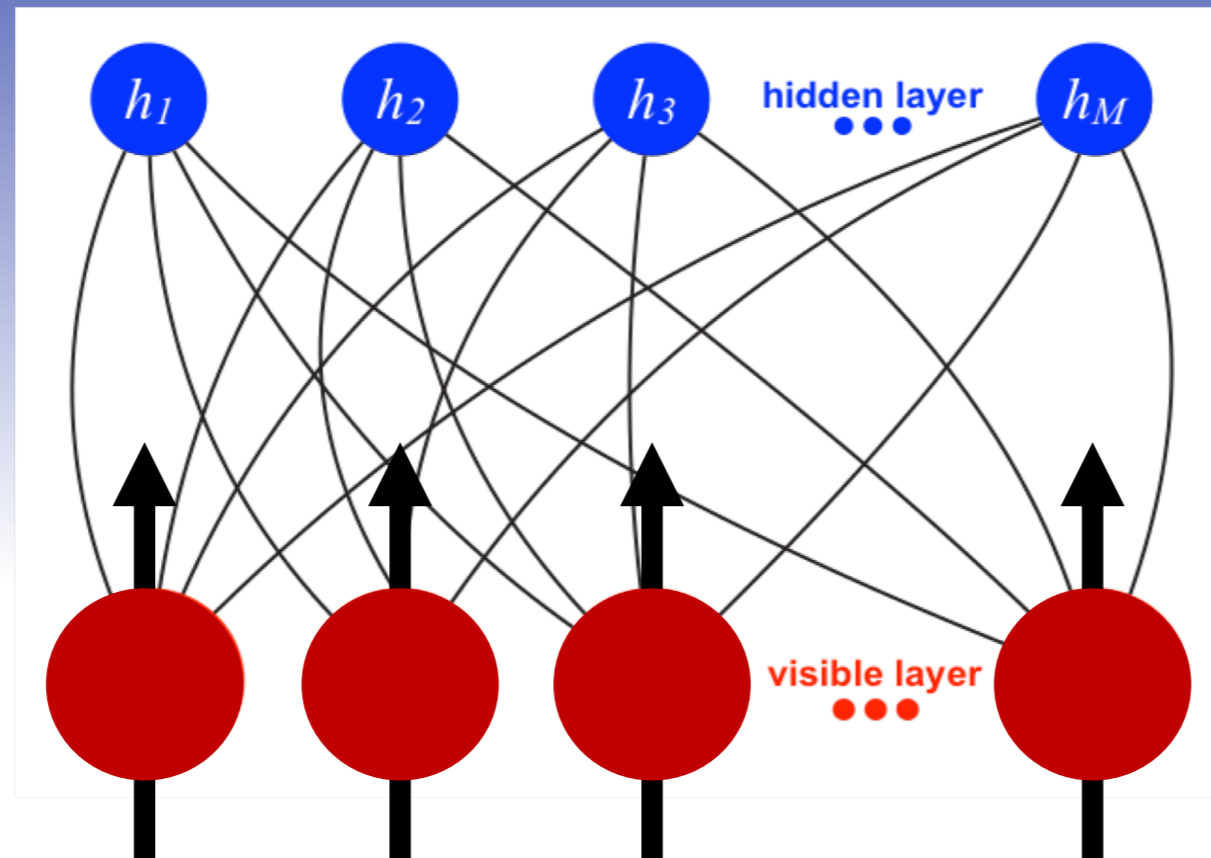
$i\partial_t \Psi(t) = \mathcal{H}(t)\Psi(t)$ impossible for large systems



Existing variational algorithms
(DMRG, MPS, PEPS...)
capture area-law entanglement

fail with large entanglement
(high-dimension, dynamics)

Machine Learning in Many-Body physics



Giuseppe Carleo, Matthias Troyer.

"Solving the Quantum Many-Body Problem with Artificial Neural Networks". Science 355, 602 (2017).

Hidden neurons represent correlations: $\alpha = \frac{\# \text{ hidden neurons}}{\# \text{ spins}}$

Quantum Entanglement in Neural Network States

Dong-Ling Deng,^{1,*} Xiaopeng Li,^{2,3,1} and S. Das Sarma¹

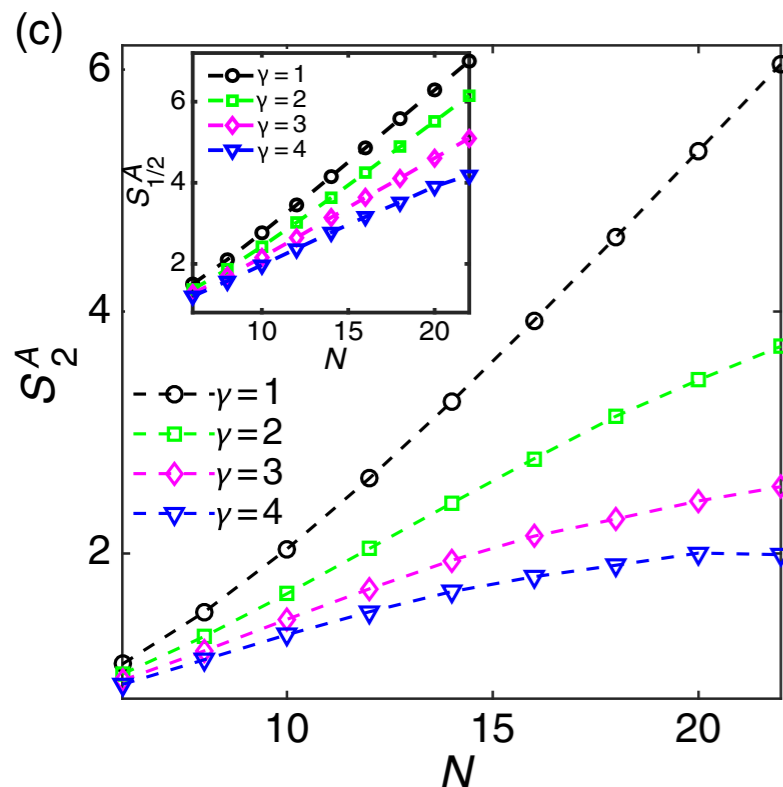
¹*Condensed Matter Theory Center and Joint Quantum Institute, Department of Physics, University of Maryland, College Park, Maryland 20742-4111, USA*

²*State Key Laboratory of Surface Physics, Institute of Nanoelectronics and Quantum Computing, and Department of Physics, Fudan University, Shanghai 200433, China*

³*Collaborative Innovation Center of Advanced Microstructures, Nanjing 210093, China*

(Received 25 January 2017; revised manuscript received 18 March 2017; published 11 May 2017)

dimensions and bipartition geometry. For long-range RBM states, we show by using an exact construction that such states could exhibit volume-law entanglement, implying a notable capability of RBM in representing quantum states with massive entanglement. Strikingly, the neural-network representation for

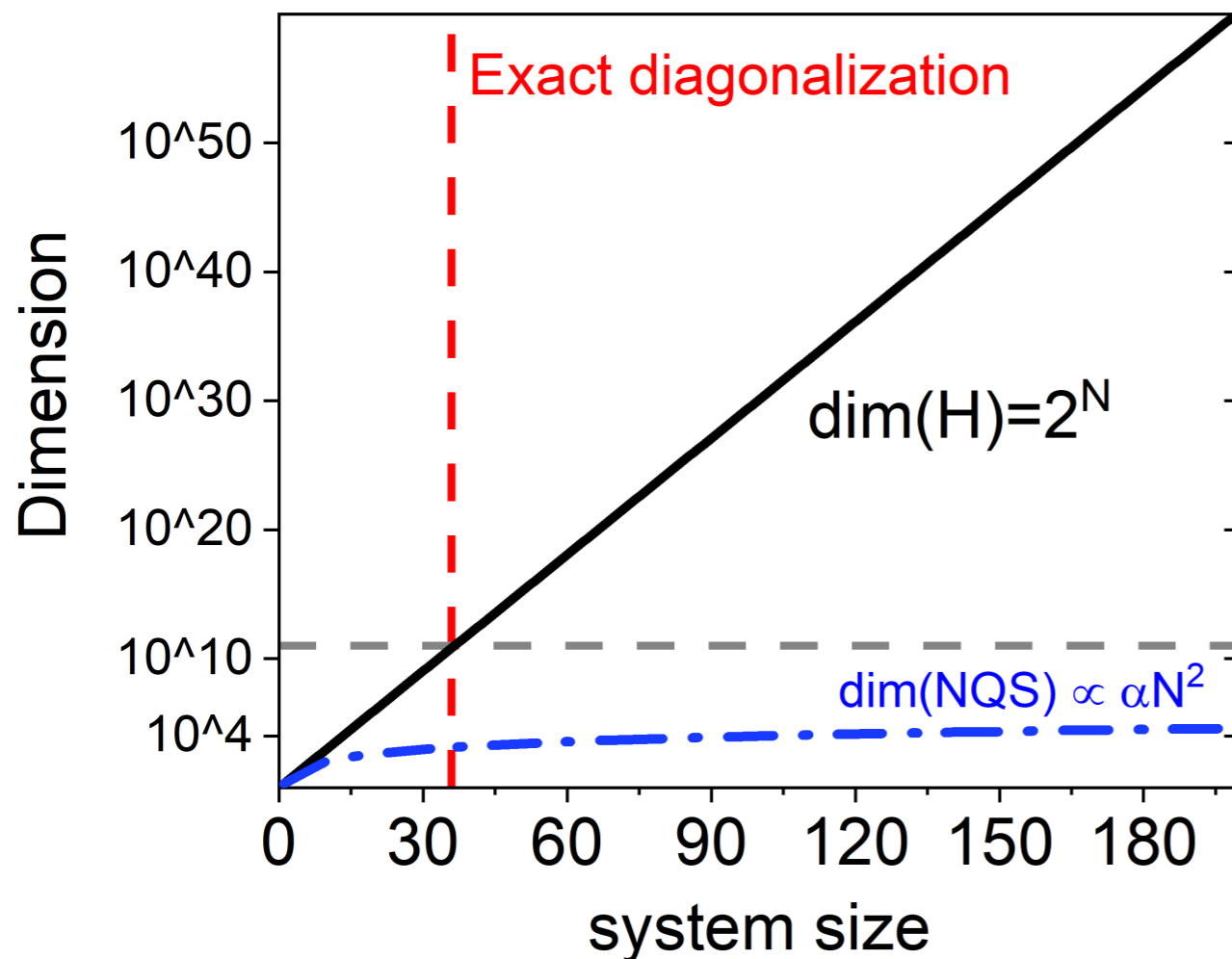


- Reduction from 2^N to αN parameters.
- Wavefunction based: no memory limitation on accessible simulation time
- Simulate dynamics of non-local correlations in systems relevant for magnetic materials

Neural Quantum States

Probability of Neural Network as variational Ansatz for wave function

$$\Psi_{\mathcal{W}}(\sigma) = P_{ANN}(\sigma) = \sum_{\{h_i\}} e^{\sum_j a_j \sigma_j^z + \sum_i b_i h_i + \sum_{ij} w_{ij} \sigma_i^z h_j}$$



Solve by optimizing network parameters $\mathcal{W} = \{a_i, b_j, w_{ij}\}$ using Monte Carlo methods

Ground state

$$\|\hat{H}\Psi - E\Psi\|_{\mathcal{W}}$$

Dynamics

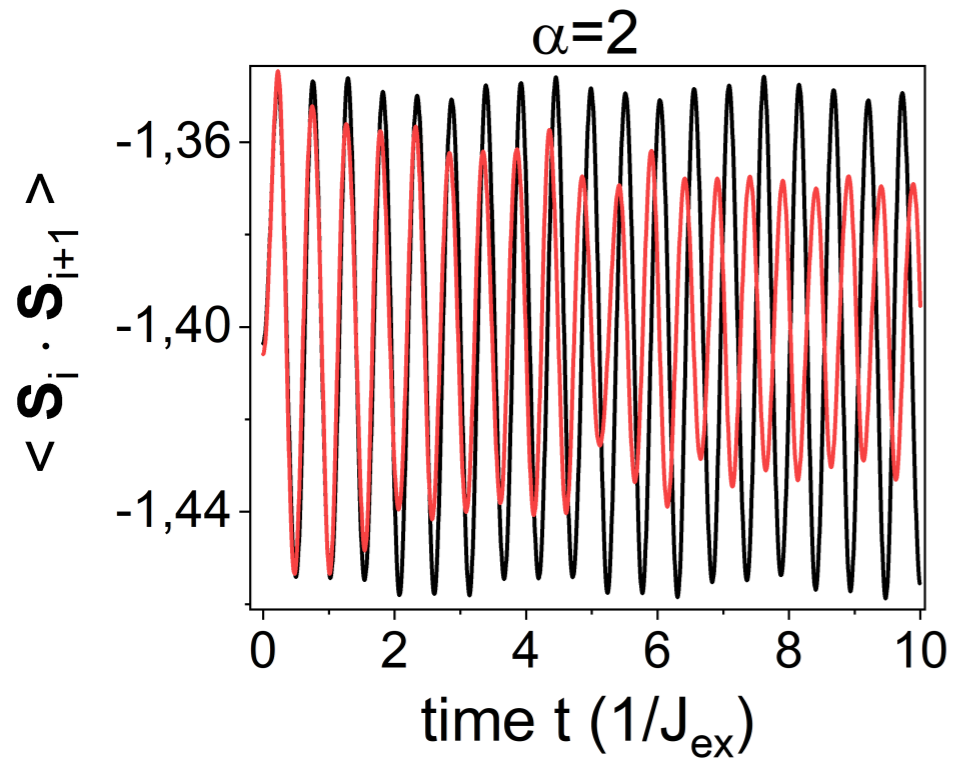
$$\|i\partial_t \Psi(t) - \mathcal{H}(t)\Psi(t)\|_{\mathcal{W}}$$

ODE for network parameters

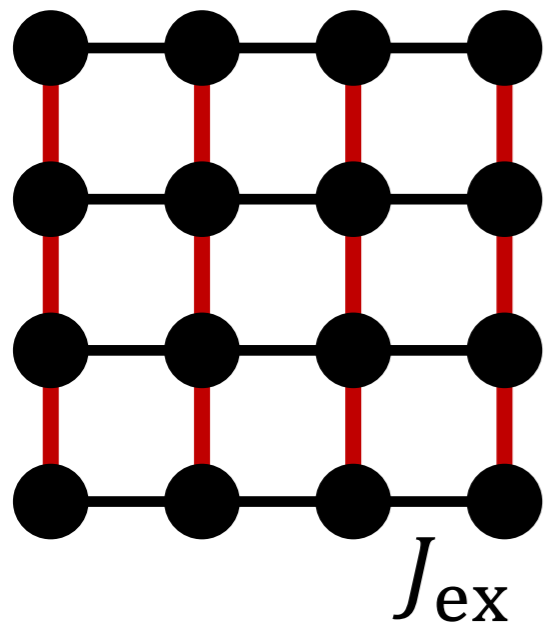
$$i S_{kk'}(t) \dot{\mathcal{W}}_{k'}(t) = \mathcal{F}(\mathcal{W}(t))$$

Here: apply NQS to study dynamics of magnon pairs

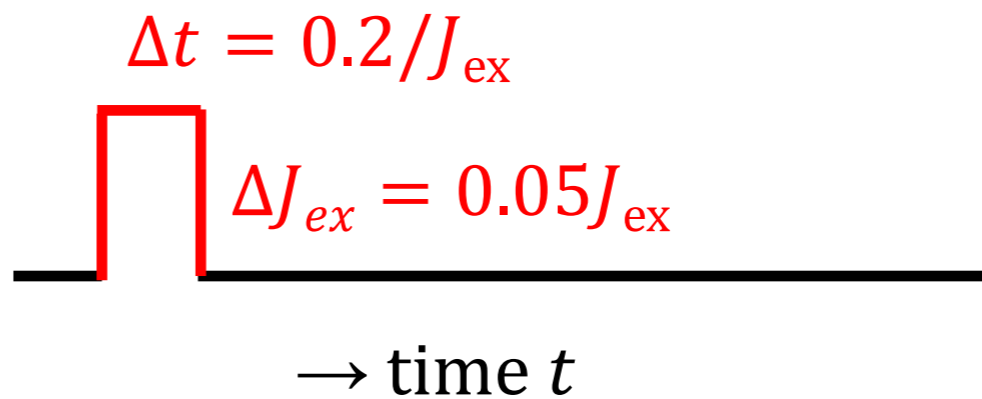
NQS vs ED (4x4)



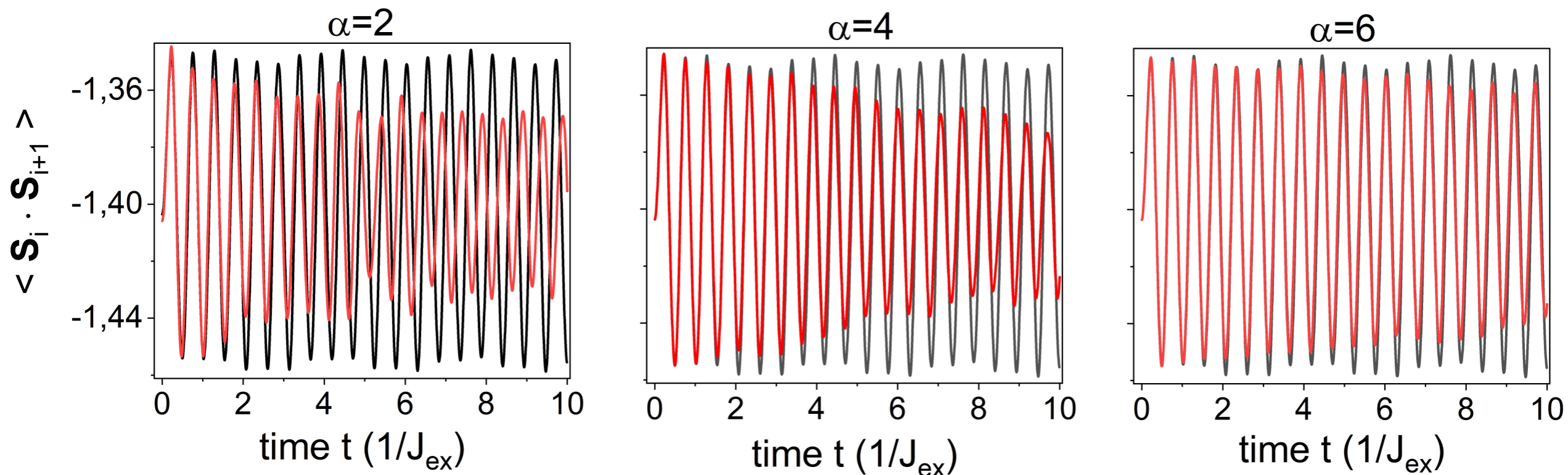
— ED
— NQS



$J_{ex} + \Delta J_{ex}(t)$

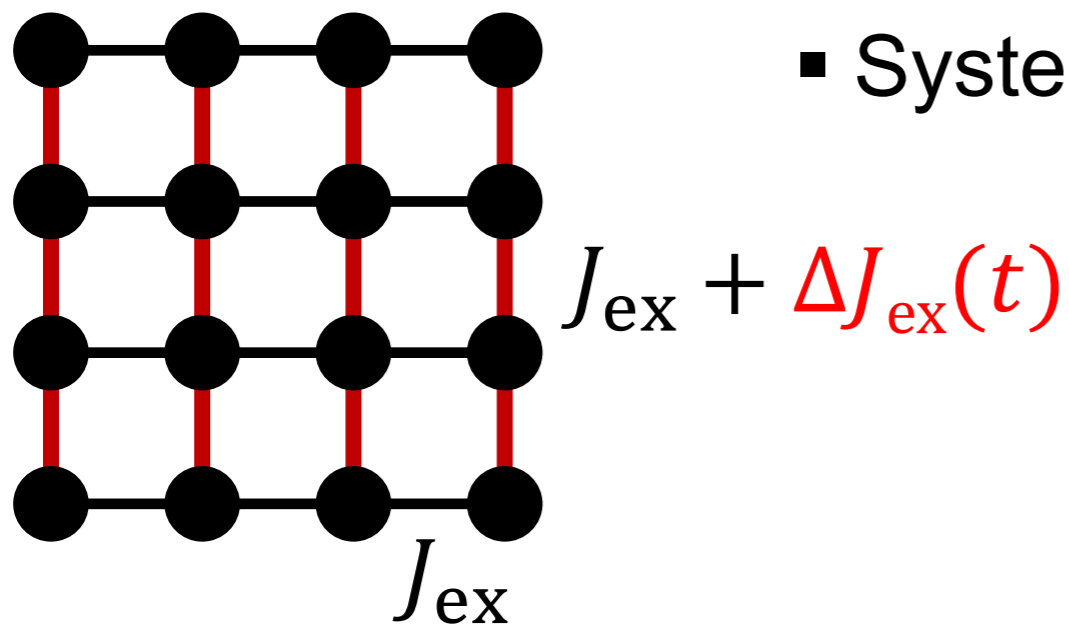


NQS vs ED (4x4)



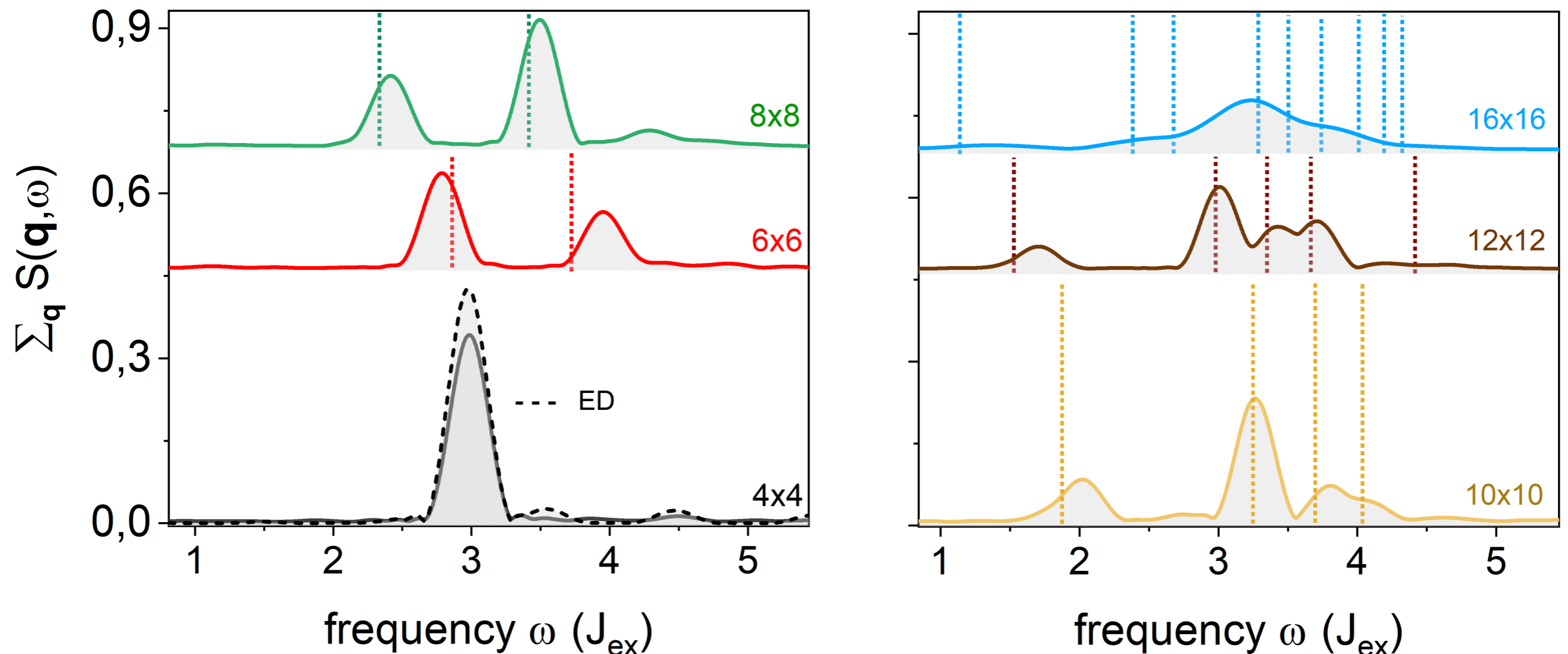
— ED
— NQS

- Pronounced damping for small α
- Systematic improvement with increasing α



Comparison interacting magnon theory

$$S(\vec{q}, \omega) = \int dt e^{i\omega t} S(\vec{q}, t)$$



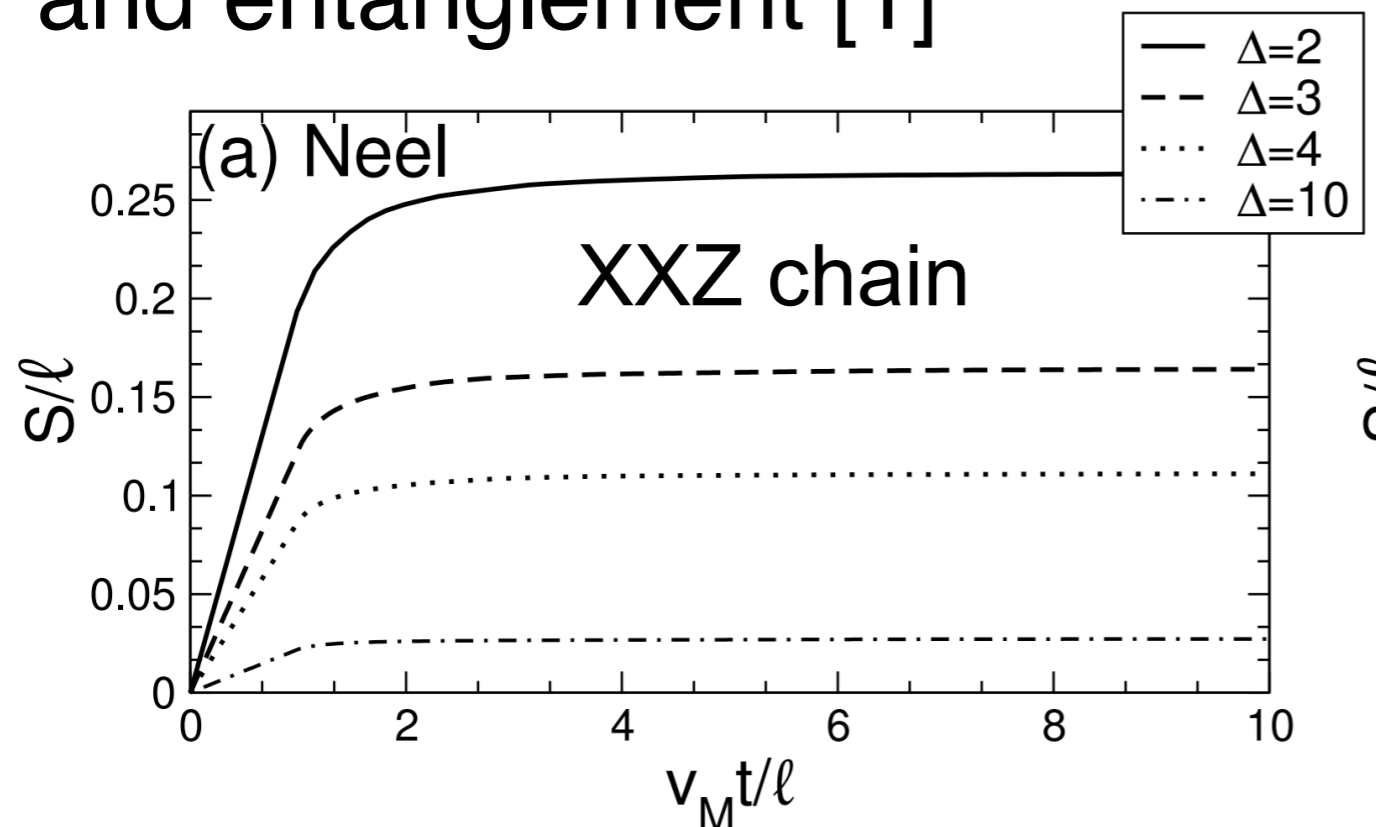
G. Fabiani, J.H. Mentink SciPost Phys. **7**, 004 (2019)
<https://github.com/ultrafast-code/ULTRAFast>

RPA results from Lorenzana and Sawatsky, PRB **52**, 8576 (1995)

Entanglement and spin correlations

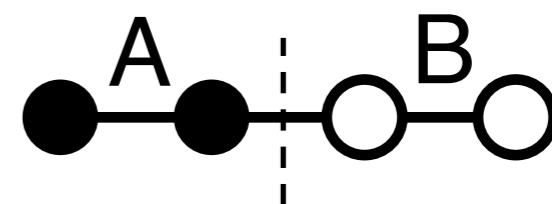
How to measure entanglement?

1D systems: direct link between spreading of correlations and entanglement [1]



Renyi entropy ($\gamma=2$)

$$S_\gamma(\rho_A) = \frac{1}{1-\gamma} \log(\text{Tr } \rho_A^\gamma)$$



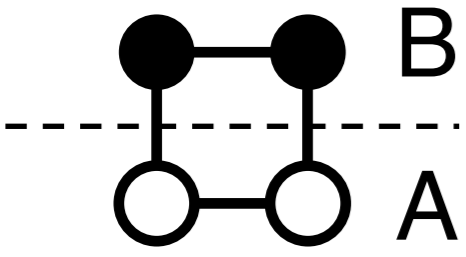
A and B not entangled:
 $S_A=0$ (product state)

[1] V. Alba, P. Calabrese, SciPost Phys. 4, 017 (2018)

Entanglement and spin correlations

How to measure entanglement?

2D systems: 2x2 system



Analytical result after $\Delta J_{\text{ex}}(t)$

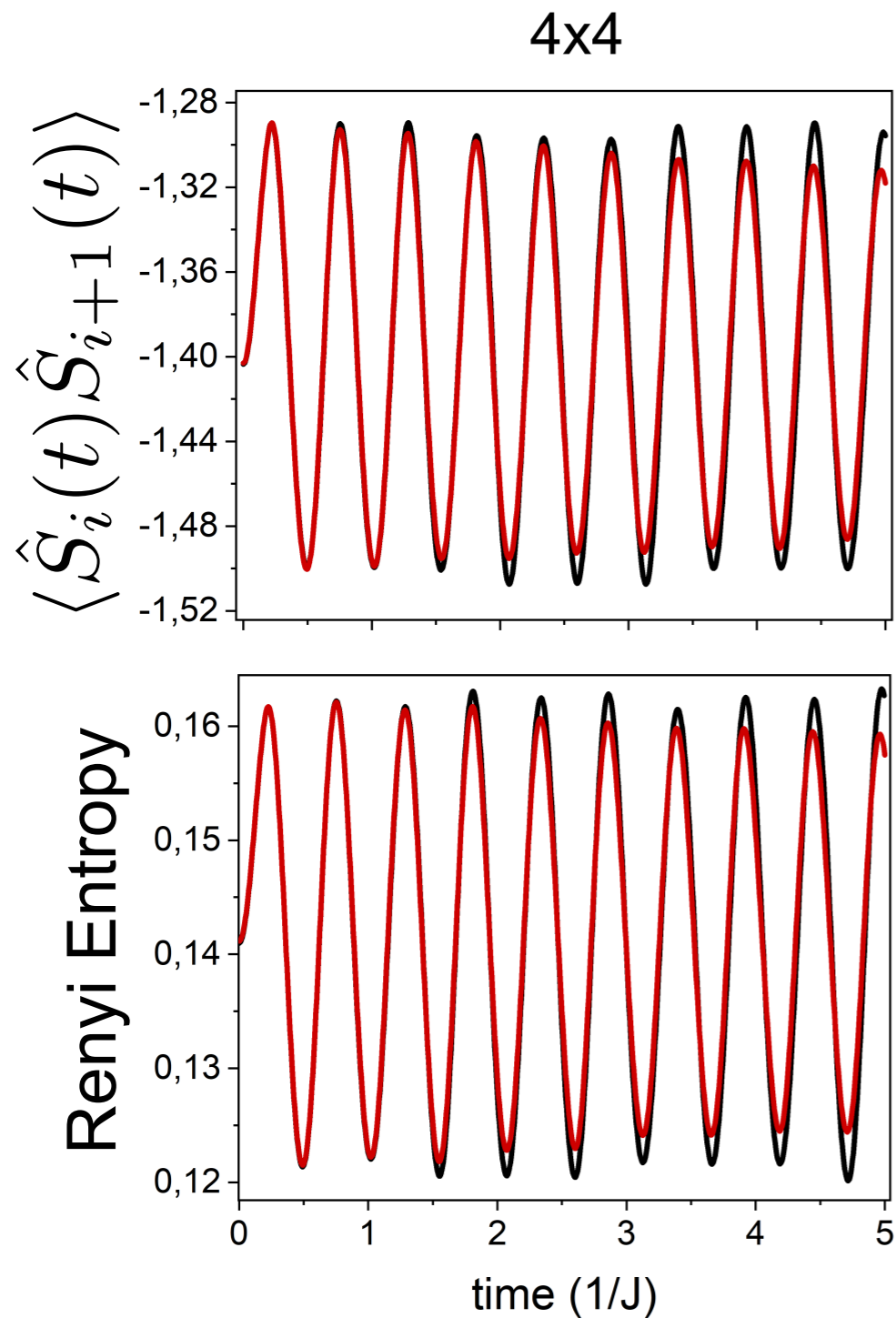
$$S_A(t) = -\log \left(\text{const} + 12 \langle S_1^z S_2^z \rangle^2 \right)$$

General result:

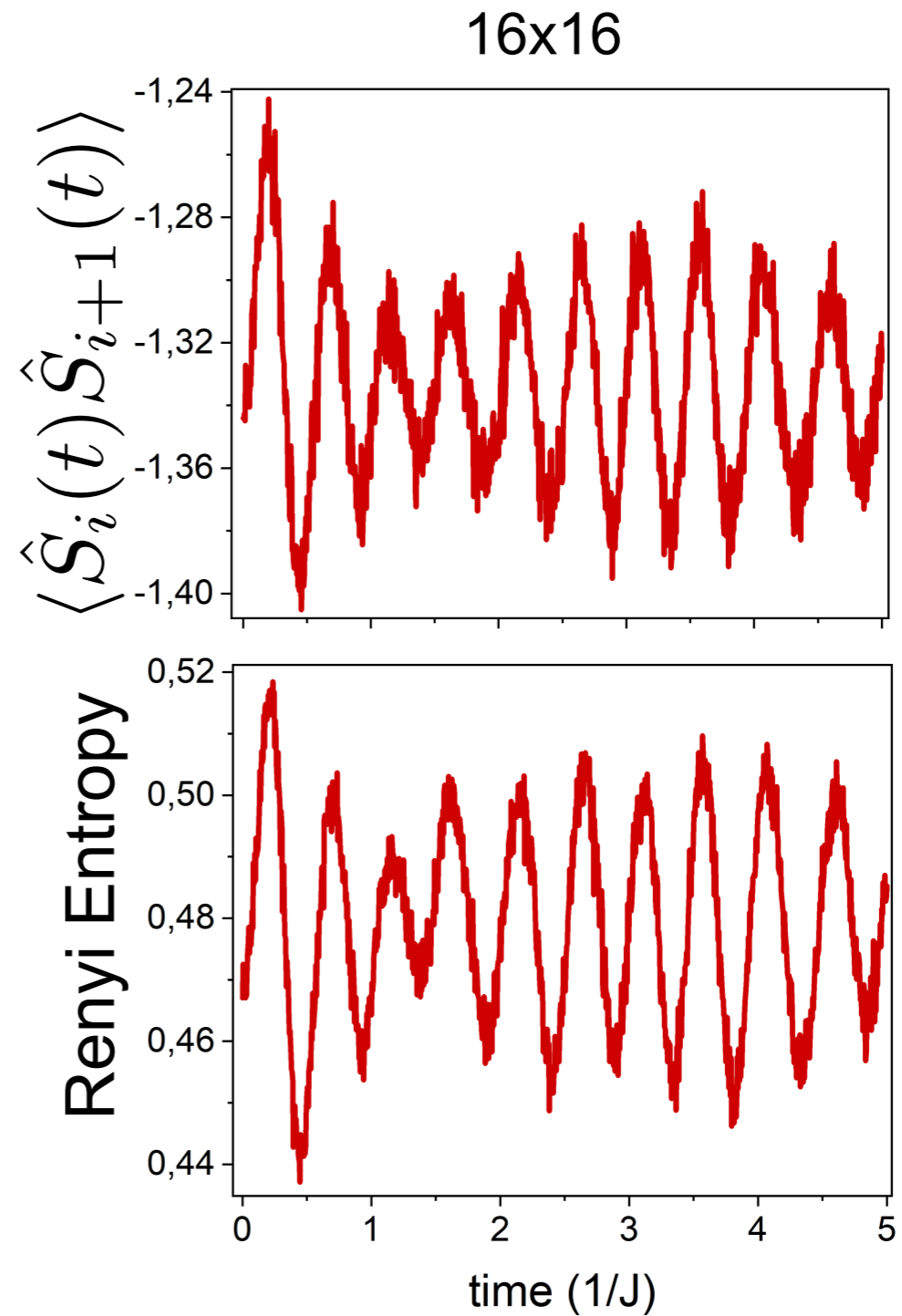
$$S_A = N_A \log(2) - \log \left(1 + \sum_i \sum_{\alpha} \langle S_i^{\alpha} \rangle^2 + \sum_{ij \in A} \sum_{\alpha\beta} \langle S_i^{\alpha} S_j^{\beta} \rangle^2 + \sum_{ijk} \sum_{\alpha\beta\gamma} \langle S_i^{\alpha} S_j^{\beta} S_k^{\gamma} \rangle^2 + \dots + N_A - \text{point correlators} \right)$$

Suggests direct link between entanglement dynamics and nearest-neighbor spin correlations!

Entanglement and spin correlations

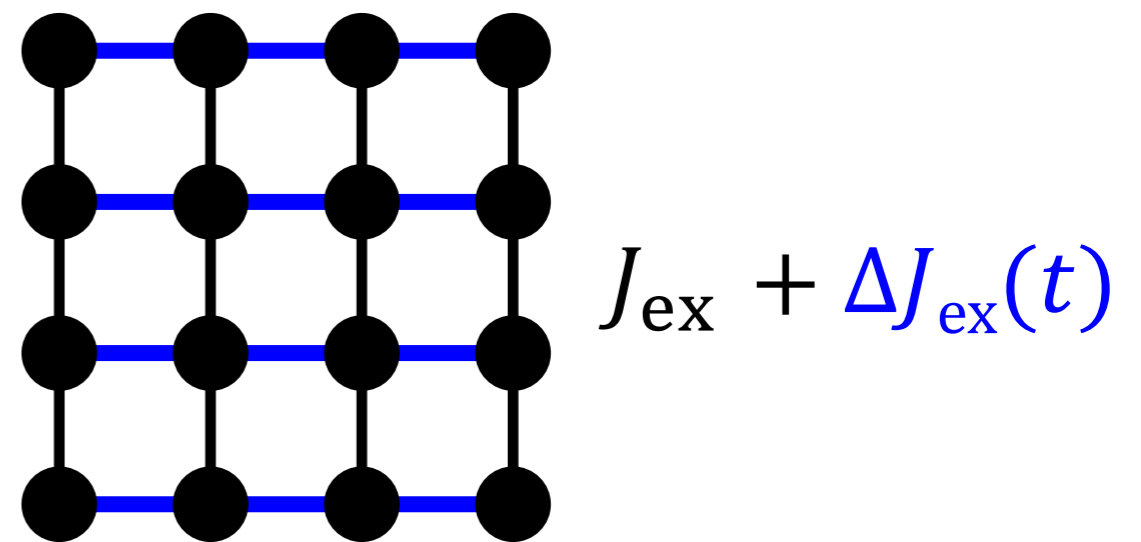
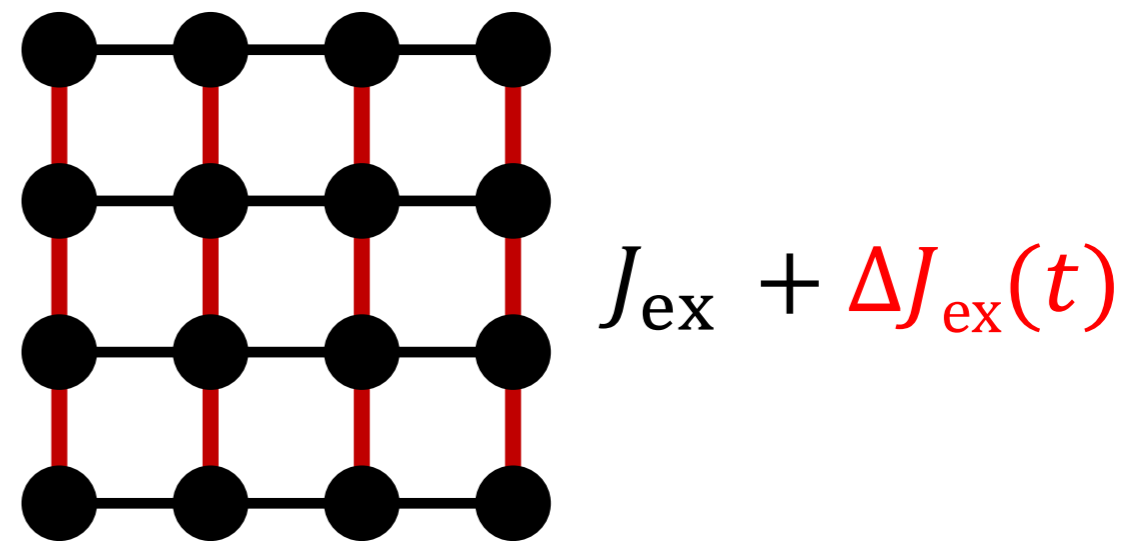
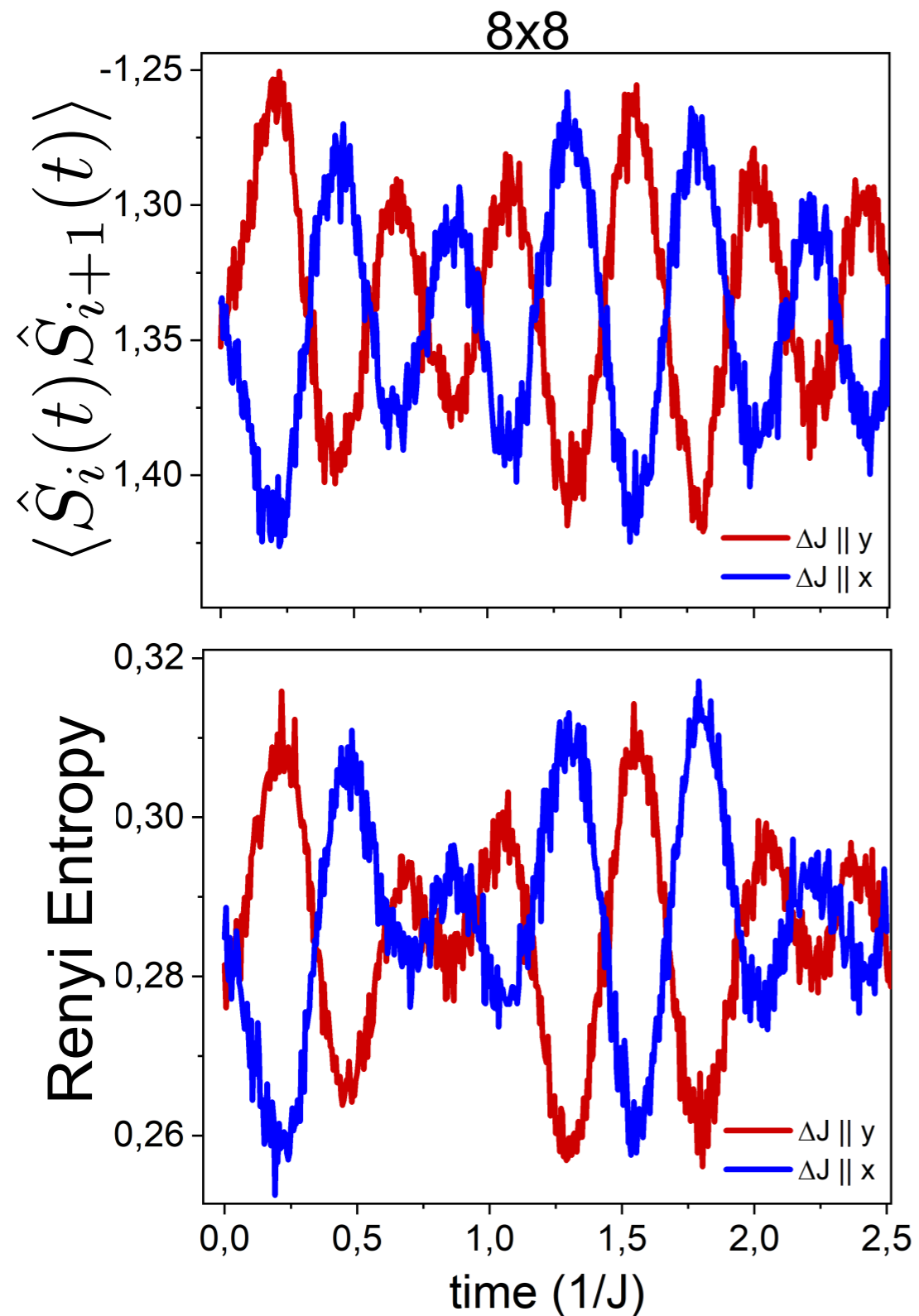


- Exact Diagonalization
- NQS wavefunction

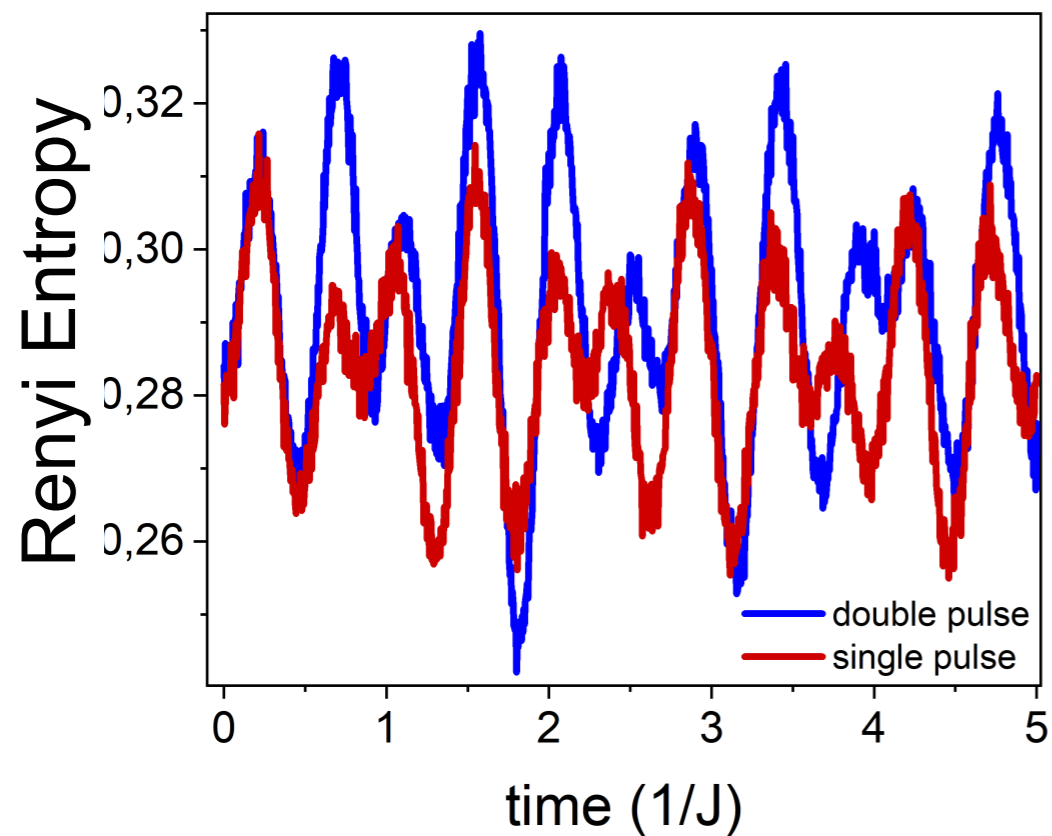
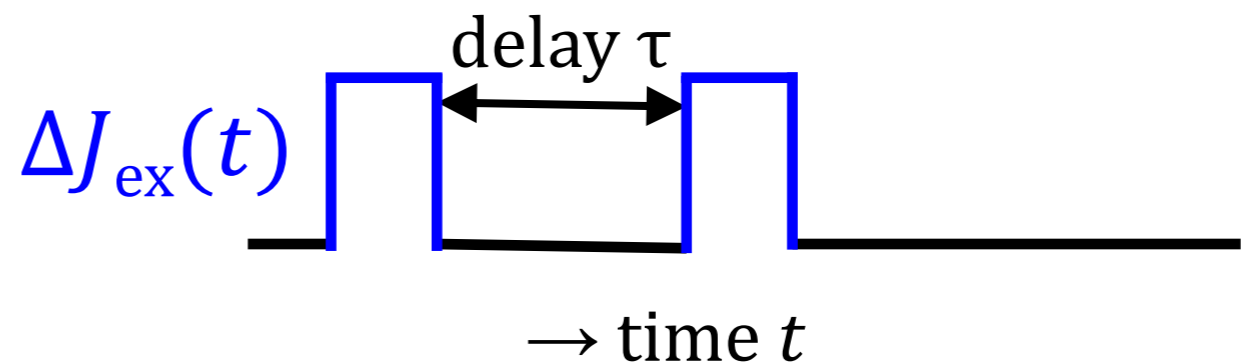
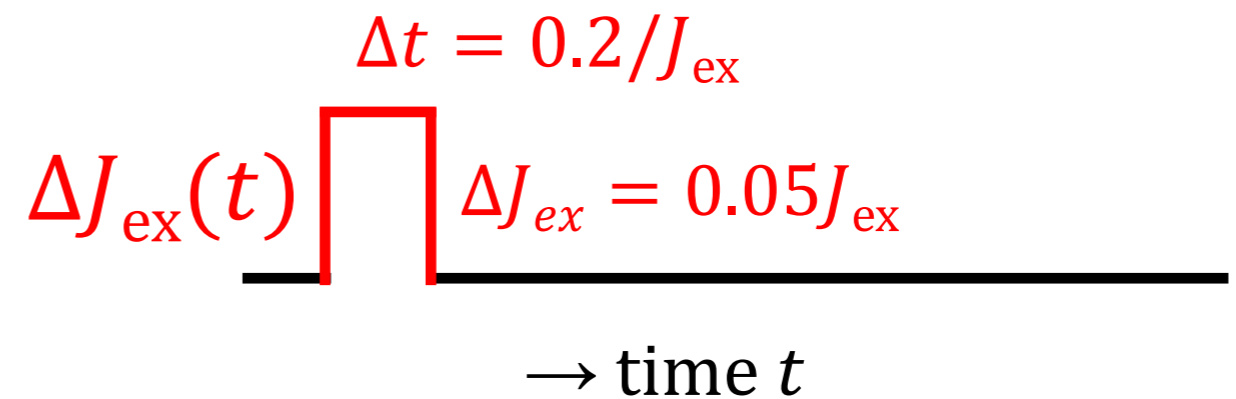
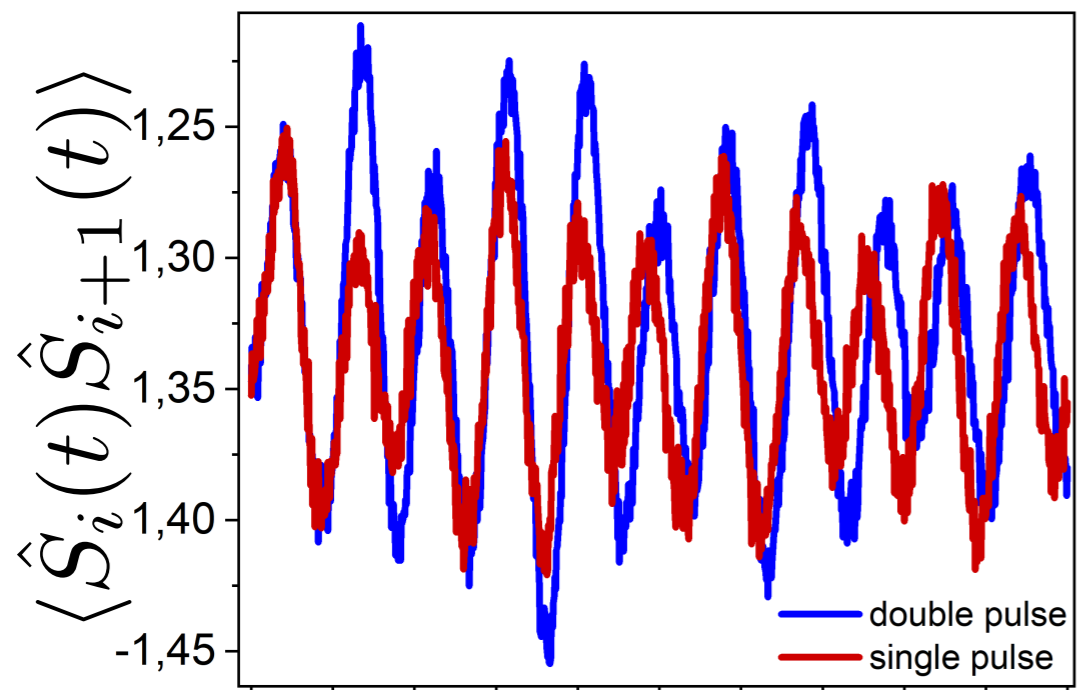


- SWAP operators* on NQS wavefunctions
- *Hastings et al., PRL 104 157201 (2010)

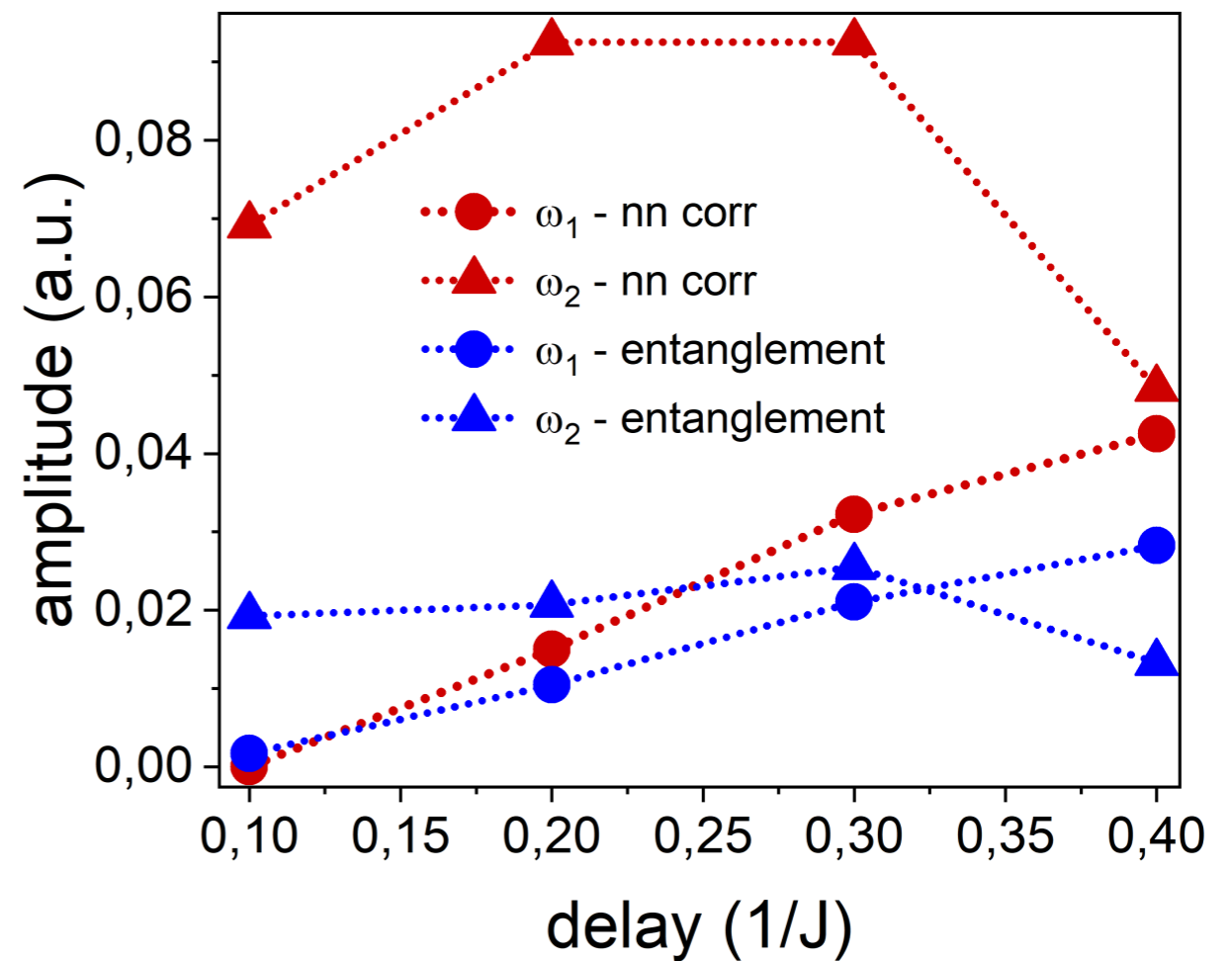
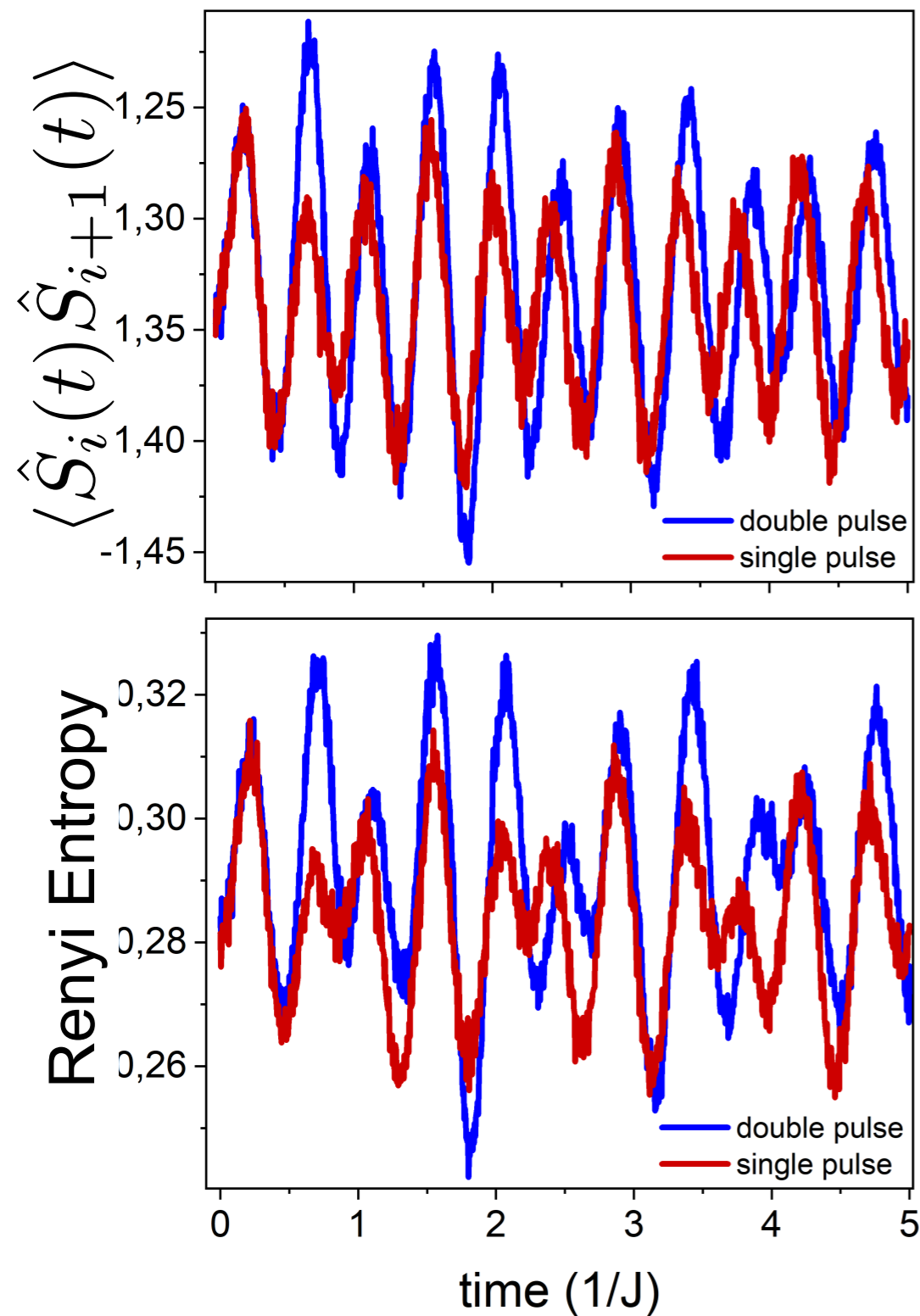
Control of the phase



Control of the amplitude



Control of the amplitude



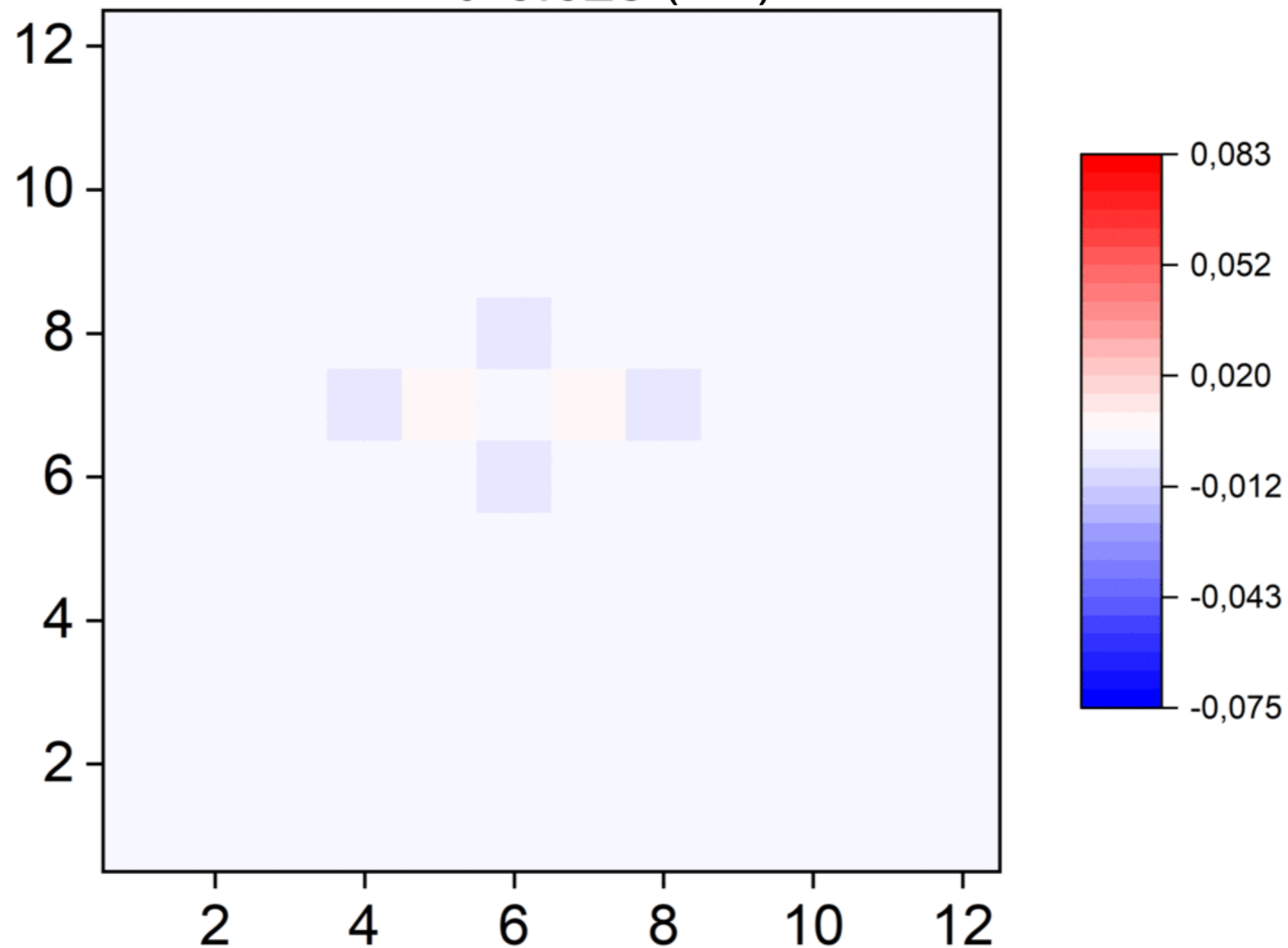
- ***Reinforces that dynamics is purely quantum***

Spreading of correlations

$$\langle S_i(t)S_j(t) \rangle - \langle S_i(0)S_j(0) \rangle$$

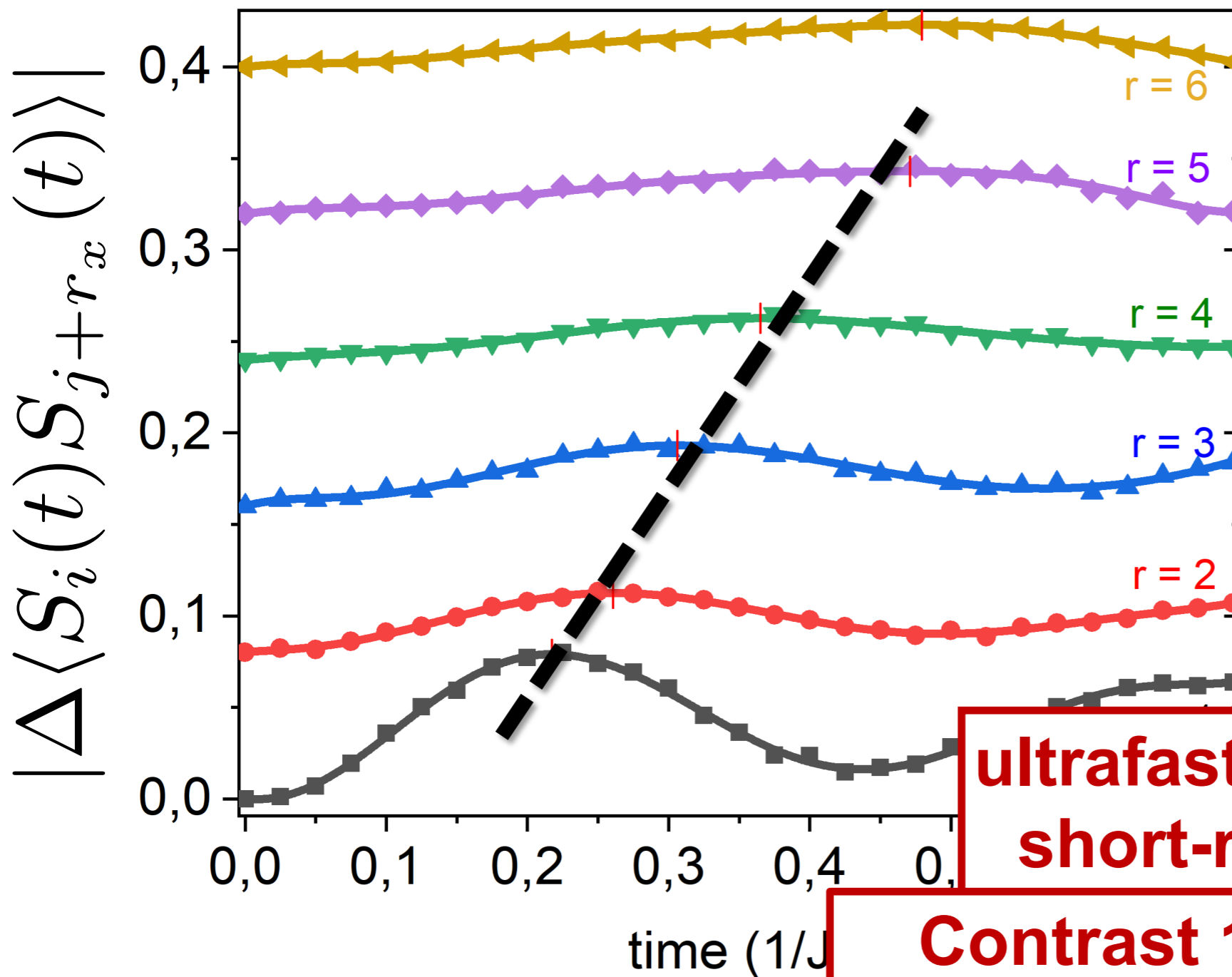
t=0.025 (1/J)

12x12



Similar anisotropic spreading known in Bose-Hubbard systems
Carleo et al., Physical Review A 89, 031602(R) (2014)

Velocity of spreading



Estimated speed
 $v = (4.2 \pm 0.8)Ja/\hbar$

Consistent with
 $\mathbf{k} = \mathbf{0}$ velocity
single magnons

$2v_g = 3.3Ja/\hbar$
 $\rightarrow 10 - 100 \text{ km/s}$

**ultrafast spreading even for
short-range correlations!**

**Contrast 1D: spreading not
related to entanglement growth**

Conclusions

- ✓ Control of J_{ex} induces >10 THz quantum dynamics in solid state
- ✓ Neural quantum states highly efficient for studying this dynamics in the 2D AFM Heisenberg model
- ✓ Entanglement manifest in nearest-neighbor spin correlations
- ✓ Coherent control of phase and amplitude of entanglement
- ✓ Ultrafast spreading of spin correlations in real space
- *Great potential for the fastest, smallest and most energy-efficient computing!*

