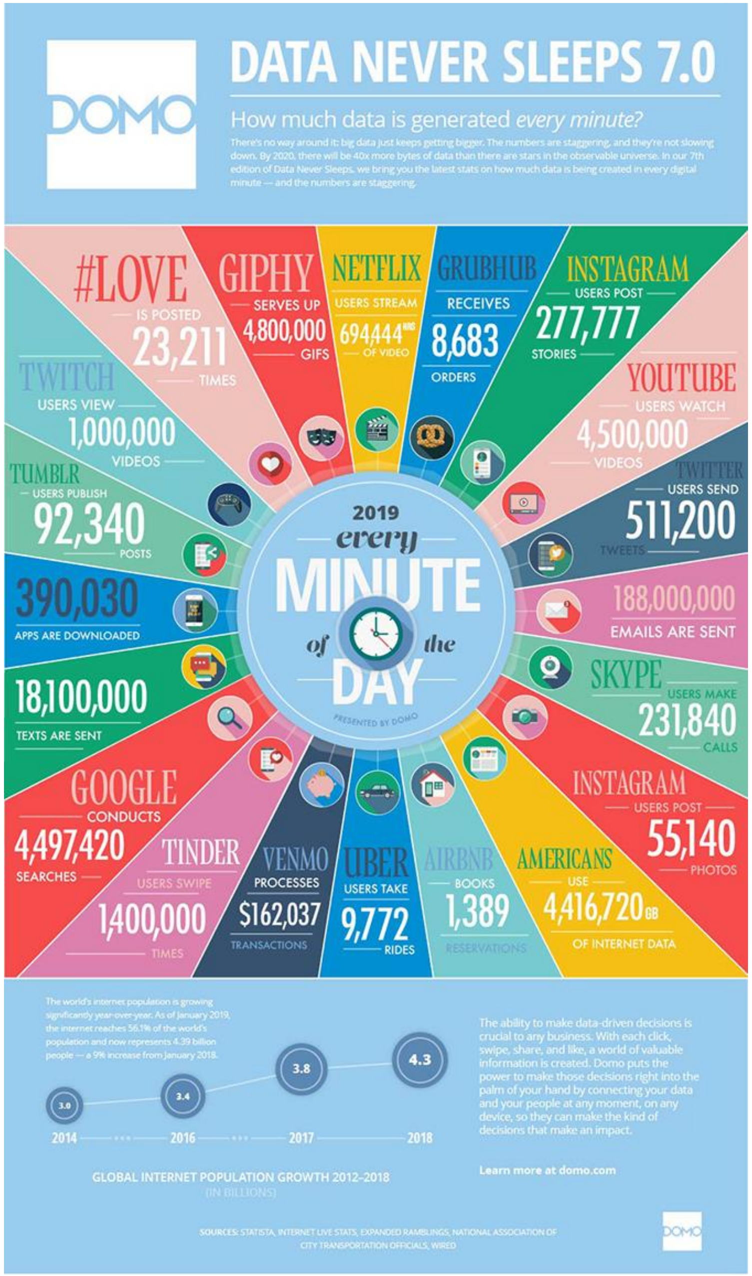
Condor Cloud:
Accelerating material discovery

Research IT Club

Daniel Reta
daniel.reta@manchester.ac.uk



2018: **2.5 quintillion** bytes per day.

2020: 40x more bytes of data than stars in the universe

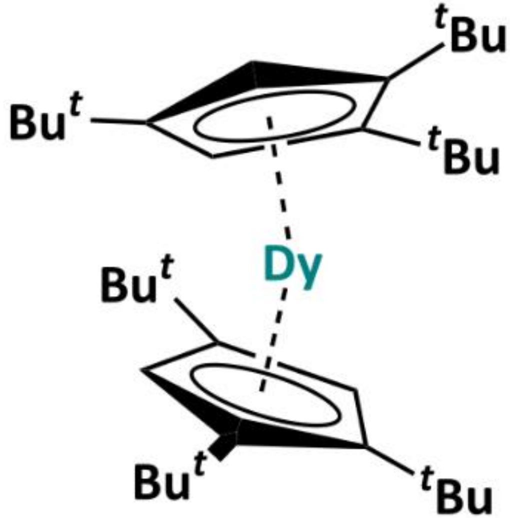
Cost:

2029: Data centres will triple power consumption.[1]

2040: 14 % of global green house emissions.[2]

[1]: Page 12 in "2015 international technology roadmap for semiconductors itrs"
[2]: The Guardian, Dec. 2017. "tsunami of data could consume fifth global electricity by 2025".

Molecules as candidates for high-density data storage



SMM:
Single Molecule Magnet

Nature, **2017**, *548*, 439–442

Nanosized:
Retain information at the molecular level.

Solution processable:
Deposition over surfaces.

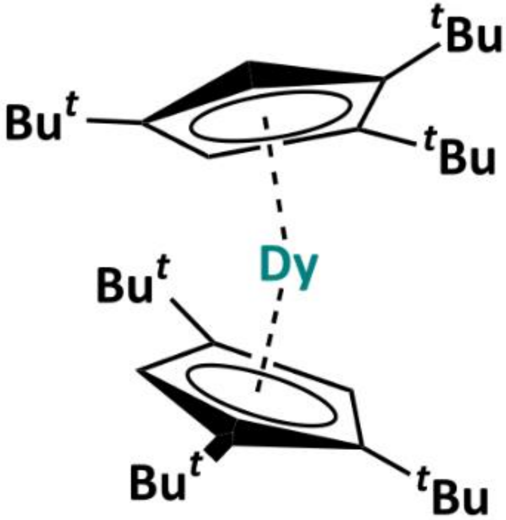
Reproducible:
Arrays of identical bits.

Delicate:
Air and temperature sensitive.

Not entirely molecular:
Properties change with surroundings.

Operational:
Information is lost @ RT.

Molecules as candidates for high-density data storage

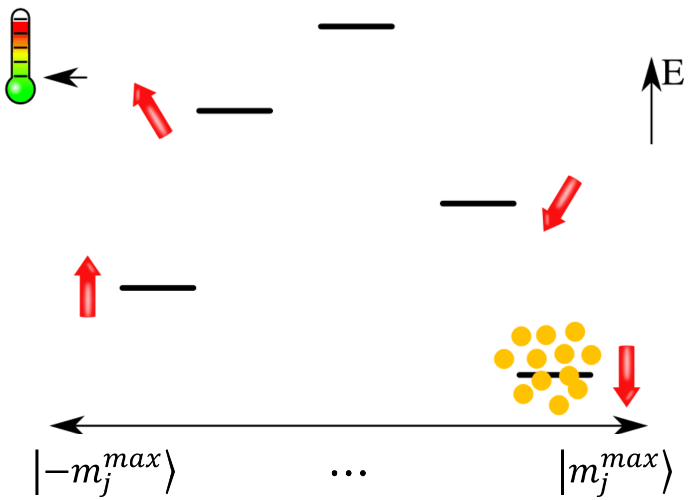


SMM:
Single Molecule Magnet

Nature, **2017**, *548*, 439–442

Our aim is to propose design strategies that improve performance.

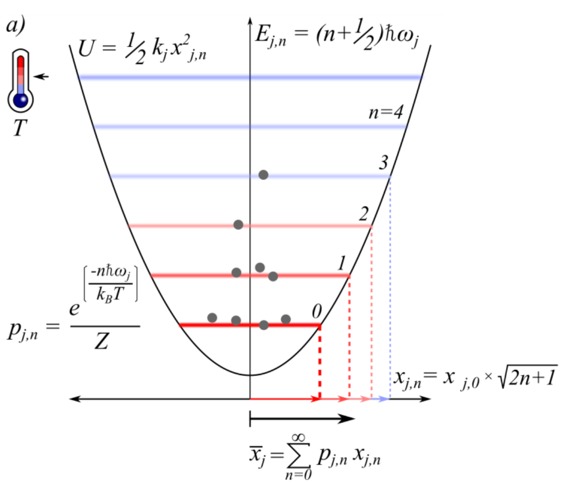
Where Amazon Web Services (AWS) Condor becomes handy



Master matrix:

$$\frac{d}{dt}p_i(t) = \sum_{f \neq i}[\gamma_{if}p_f(t) - \gamma_{fi}p_i(t)]$$

To calculate γ_{if} one has to repeat many nearly identical calculations.



Molecule with N atoms:

3N-6 modes.
4 points each.

N = 90 → 1056 jobs

~47k jobs,
2 weeks limit

Why AWS Condor vs other architectures

Required program:	• Molcas Runs in 1 core (no need for MPIs). ~1 GB memory per job. 4-6 hours per job. Installed in AWS – spin up an instance with program image for each job.
Desired performance:	• As large of a throughput as possible.

AWS Condor specifics

- **Spot** *vs* on-demand requests for computing nodes:

  Using spare capacity is 80% cheaper – no priority over on-demand users.

- Instance type: **r5.large**

  1 computing node, containing 2 cores (halves the requirements).
  8 GB each core.

- **Price** (https://aws.amazon.com/ec2/spot/pricing/)**:**

  $0.021/hour (*vs* $0.126/hour on-demand)

- Changes to standard Condor submission script:

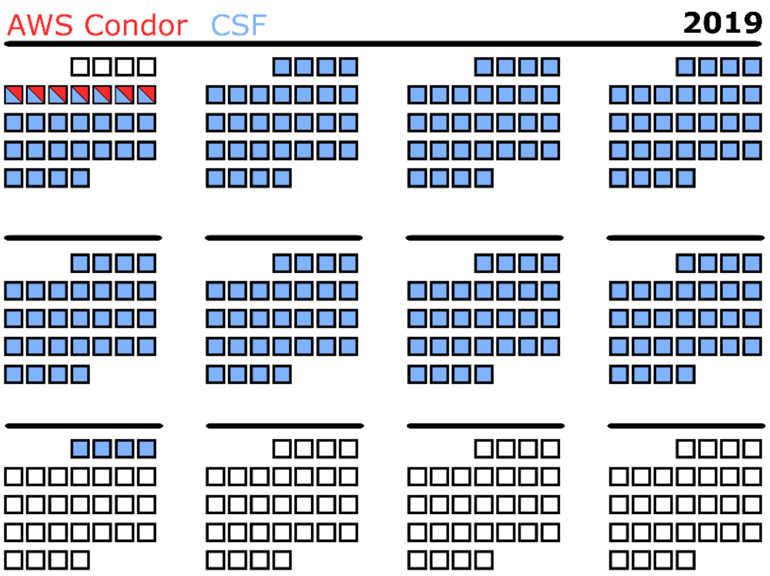
  +MayUseAWS=True

- Internal script that boosts throughput by monitoring queued jobs – submission in batches.

Results & conclusions

- ~ 1500 jobs per day, using 1600 cores at a time:

  32x throughput increase *vs* CSF.



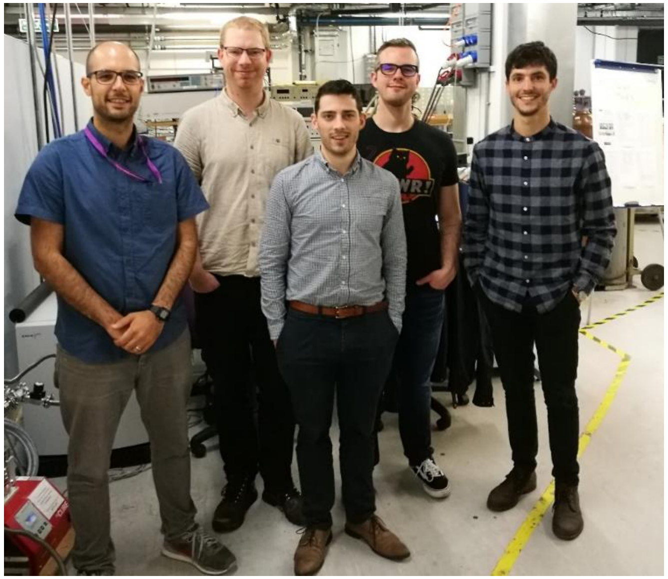
- Price:

  On r5.large, 5 hrs job = $0.1
  → 47k jobs = 23.5k nodes = $2.35k

- Proof of concept**:**

  If you have a computational problem that is highly parallel, AWS Condor is an ideal solution - all necessary tools and expertise are now tested, reliable and user-friendly, so go talk to the Research IT.

Acknowledgements



Dr. N. F. Chilton   Dr. D. Mills

Dr. C. Goodwin   Dr. F. Ortu

Chilton group

The people who made this work:

Daniel Corbett

Dr. Chris Paul

Simon Hood

its-ri-team@manchester.ac.uk

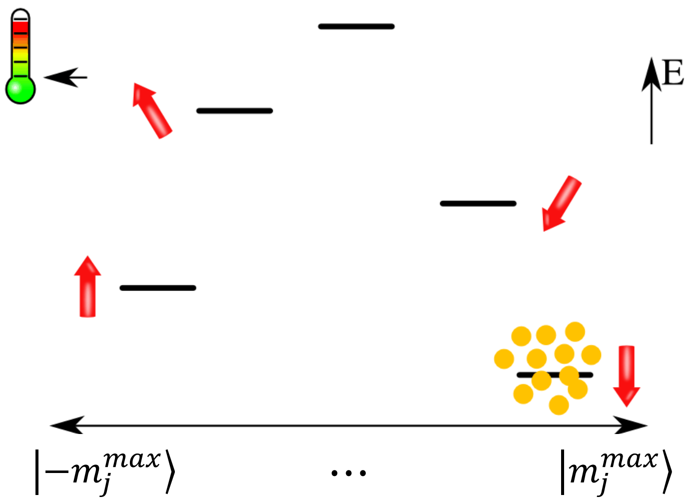
MANCHESTER 1824

The University of Manchester

EPSRC

Engineering and Physical Sciences Research Council

Our approach to ab initio spin dynamics



Master matrix:

$$\frac{d}{dt}p_i(t) = \sum_{f \neq i} [\gamma_{if} p_f(t) - \gamma_{fi} p_i(t)]$$

Markov-process. Spin dynamics independent of molecular dynamics.

Construct matrix with transition rates γ_{if} as elements: Set of LDE.

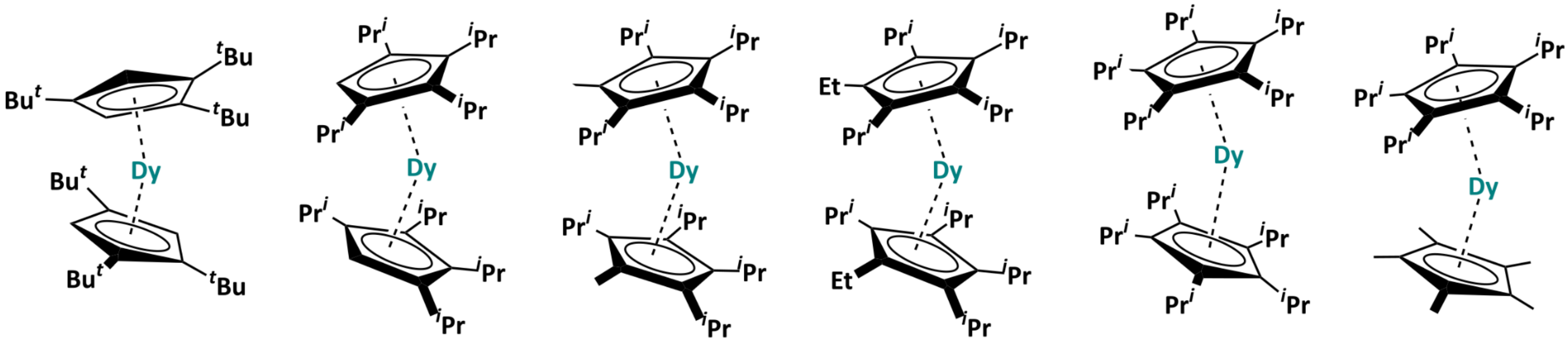
Diagonal elements $\gamma_{ii} = \sum_{i \neq f} -\gamma_{if}$ ensures conservation of population.

Eigenvalues are $-{}^{1}/_{\tau_k}$. One is zero, representing equilibrium.

C. A. P. Goodwin, F. Ortu, D. Reta, N. F. Chilton, D. P. Mills, *Nature*, **2017**, *548*, 439

Systems studied

Nature, **2017**, *548*, 439 | *Chem. Sci.* **2019**, *9*, 8492 | *Science* **2018,** *362*, 6421



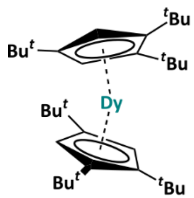
T_B (K)						
	56	17	62	59	56	67

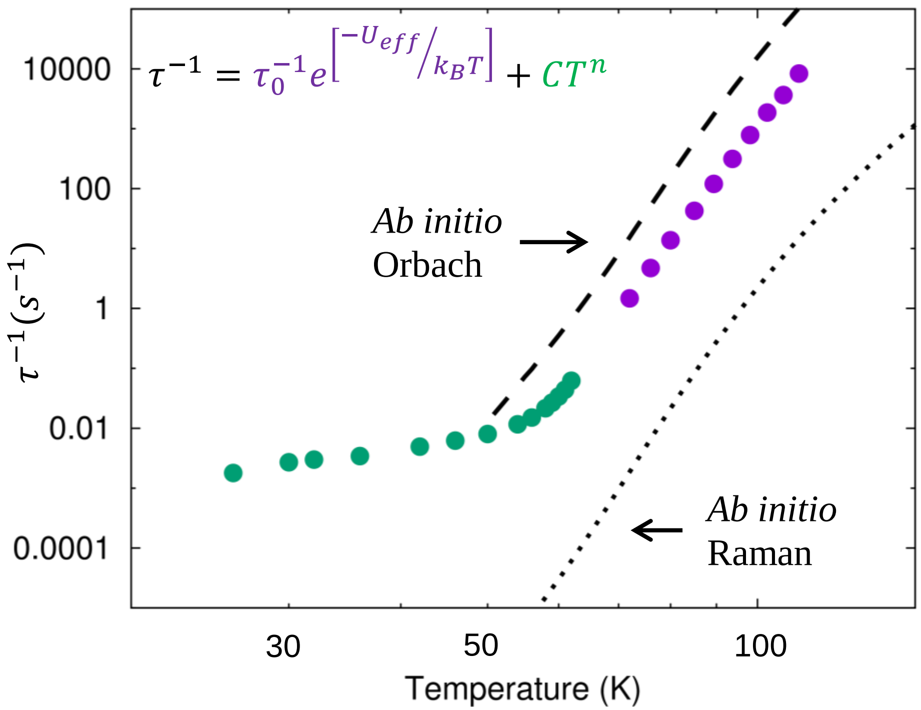
Geometry optimisation: Gaussian 09, PBE, GD3,
Dy → Y (Iso=162.5) Stuttgart ECP, cc-pVTZ Cp C, cc-pVDZ remaining,
mode energies corrected against IR.

Electronic structure: MOLCAS, CAS(9,7)SCF-SO-RASSI, Sextets, S.A. 21 roots,
Dy ANO-RCC-VTZP, Cp C ANO-RCC-VDZP, ANO-RCC-VDZ remaining,
$(3N_{atoms} - 6) \cdot N_{distortions}$ CASSCF calculations.

C. A. P. Goodwin, F. Ortu, D. Reta, N. F. Chilton, D. P. Mills, *Nature*, **2017**, *548*, 439

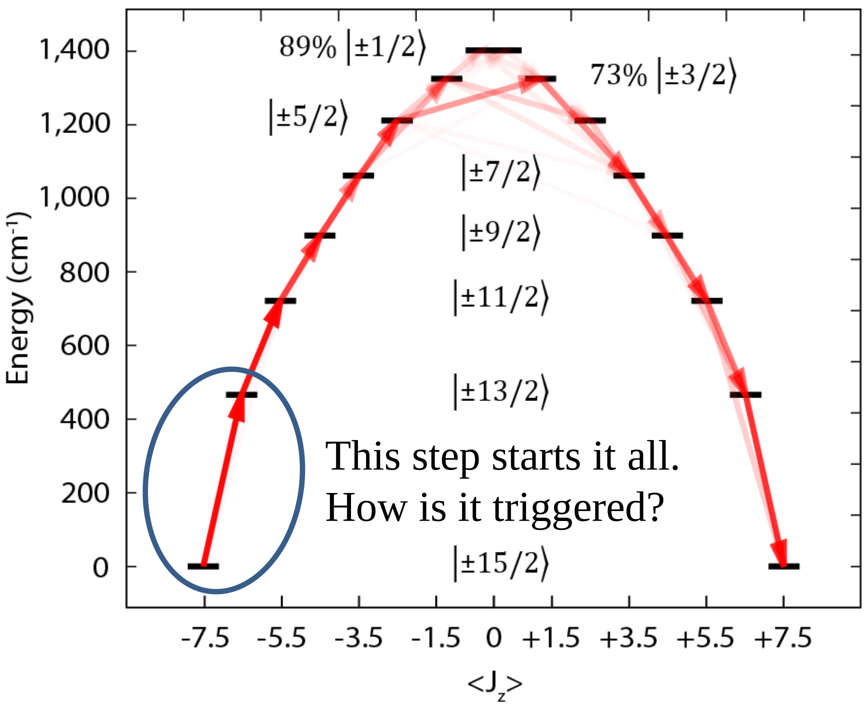
Ab initio spin dynamics: Results





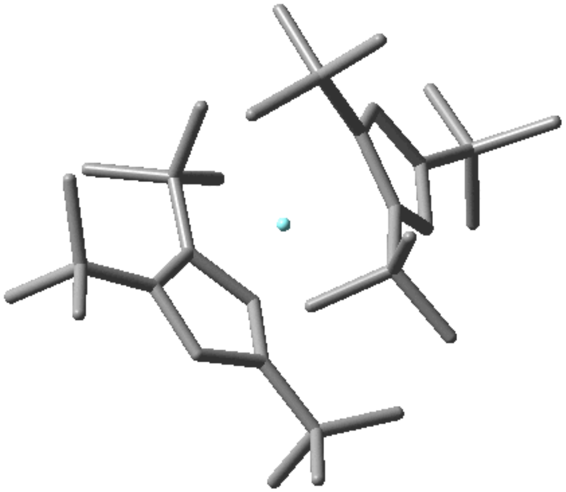
Transitions increasing populations

$$p_T(0) = p_{|-15/2\rangle}(0) = 1$$



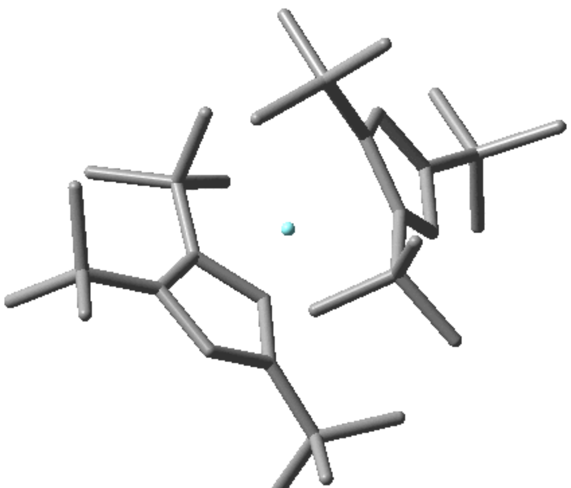
C. A. P. Goodwin, F. Ortu, D. Reta, N. F. Chilton, D. P. Mills, *Nature*, **2017**, *548*, 439

Ab initio spin dynamics: Results



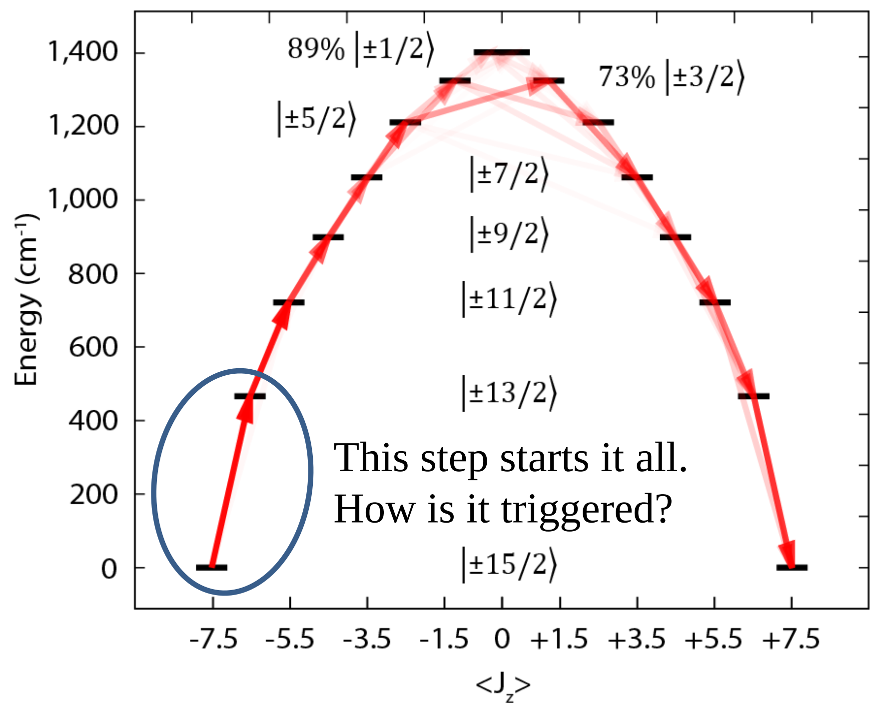
$$460 < E_{mode}(cm^{-1}) < 470$$

$$\Delta E_{states} = 461\ (cm^{-1})$$



Transitions increasing populations

$$p_T(0) = p_{|-15/2\rangle}(0) = 1$$



C. A. P. Goodwin, F. Ortu, D. Reta, N. F. Chilton, D. P. Mills, *Nature*, **2017**, *548*, 439