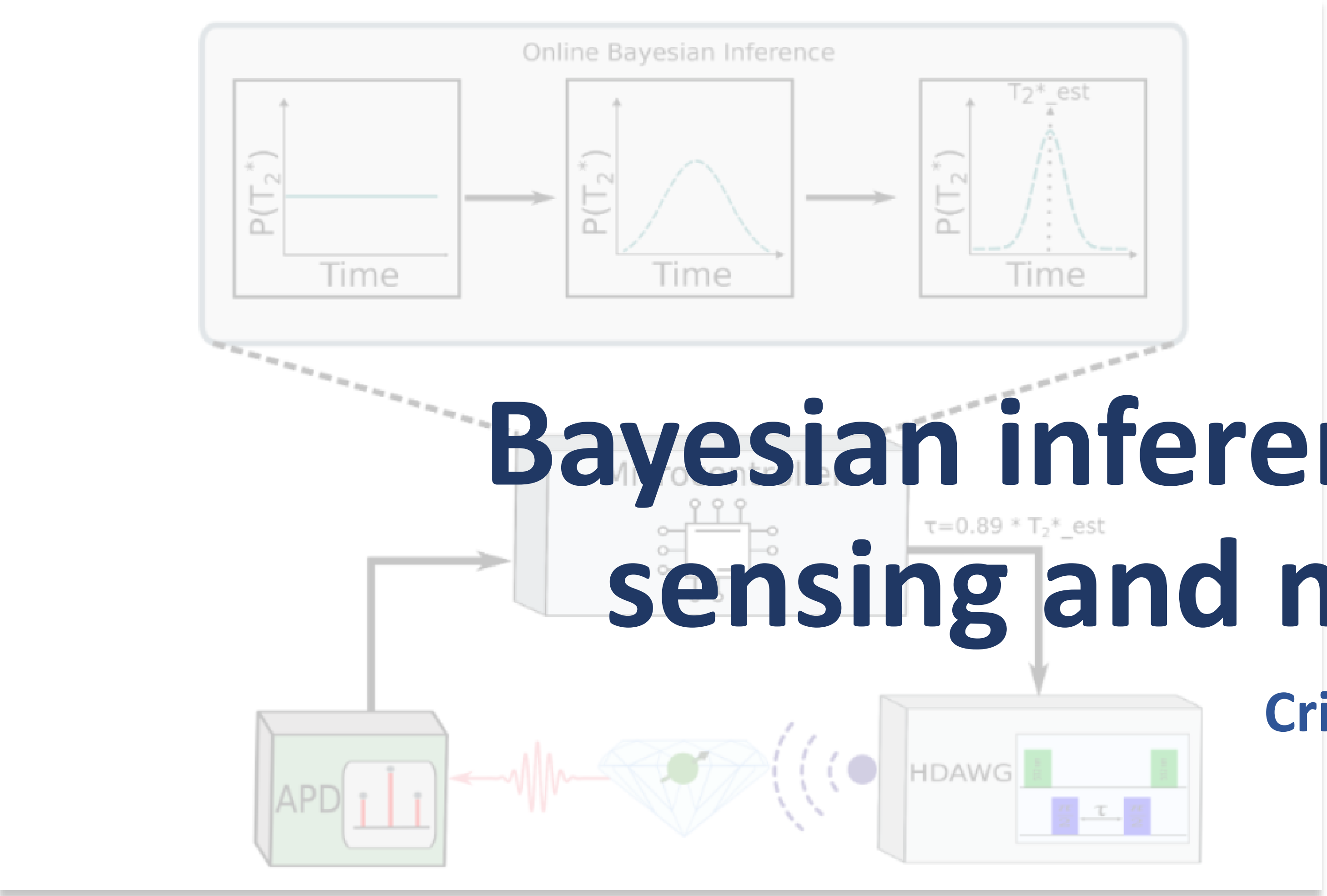


Bayesian inference for quantum sensing and model learning

Cristian Bonato, Quantum Photonics Lab,
Heriot-Watt University, Edinburgh (UK)
qpl.eps.hw.ac.uk

UKRI Engineering and Physical Sciences Research Council
LEVERHULME TRUST
WEIZMANN UK SUPPORTING THE WEIZMANN INSTITUTE OF SCIENCE
SCIENCE FOR THE BENEFIT OF HUMANITY
[dstl] The Science Inside
QUANTIC The UK Quantum Technology Hub in Quantum Enhanced Imaging

Collaboration

Cristian Bonato
(experimental physics)

Erik Gauger
(theoretical physics)

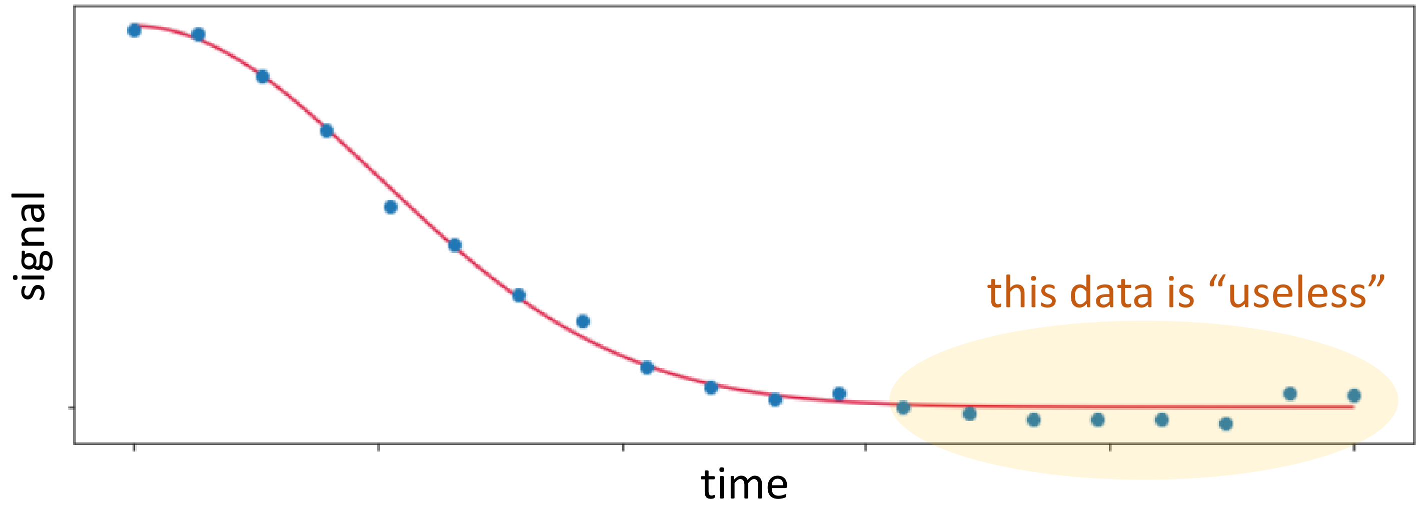
Yoann Altmann
(signal processing)

WEIZMANN UK
SUPPORTING THE WEIZMANN INSTITUTE OF SCIENCE
SCIENCE FOR THE BENEFIT OF HUMANITY

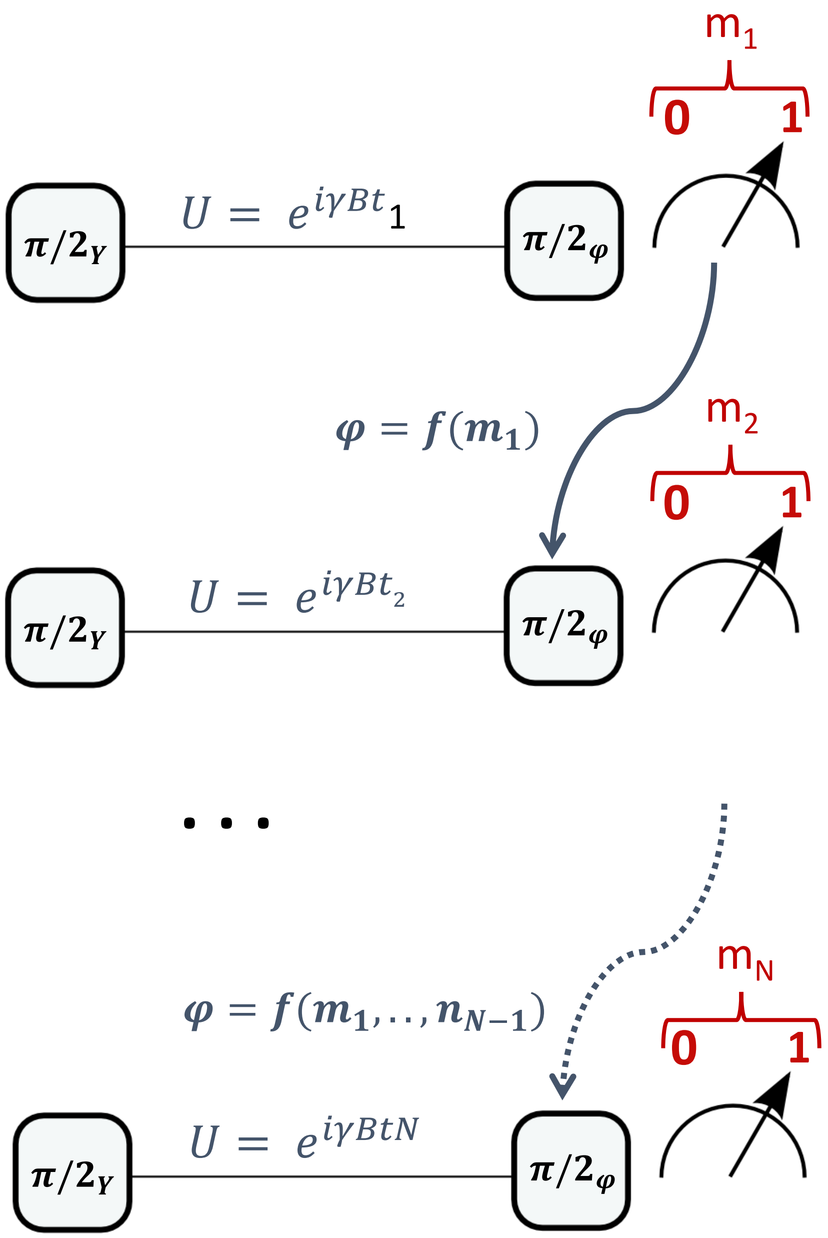
QUANTIC
The UK Quantum Technology Hub
in Quantum Enhanced Imaging

Standard way of taking data: sweep a parameter

For example, if you want to measure the loss of quantum coherence, you perform a sequence of Ramsey experiments, sweeping the delay time over a pre-determined range:



Adaptive quantum sensing experiments



Adaptive measurement: use information from earlier measurement outcomes to estimate the a quantity and optimise parameters for later measurements in real-time

Adaptive Bayesian experiment design

$$P\,(B|m) \propto \boldsymbol{P(m|B)}\; P(B)$$

(1) Easily include **ALL information (imperfections, prior info, etc)** available

Ideal likelihood:

$\mathrm{Pr}["0"] \sim 1 + \cos(\boldsymbol{\gamma B}t + \varphi)$

γ = 28 MHz/mT

photon collection efficiency

dark counts

decoherence

$$P(\mu = 0|f_B) = \frac{(1 + F_0 - F_1)}{2} + \frac{(F_0 + F_1 - 1)}{2} e^{-(t/T_2^*)^2} \cos[2\pi f_B t + \vartheta]$$

$$P(\mu = 1|f_B) = 1 - P(\mu = 0|f_B)$$

(2) integrate **online adaptation**:

Current $P(B)$ can be used to optimise settings for next measurement

Our strategy is to maximise Fisher information:

Fisher information:

$$\mathcal{I}(\theta) = -\operatorname{E}\left[\left.\frac{\partial^2}{\partial\theta^2}\log f(X;\theta)\right|\theta\right]$$

Cramer-Rao bound:

$$\operatorname{var}(\hat{\theta}) \geq \frac{1}{I(\theta)}$$

It’s an asymptotic bound, but it works well for simple cases of single-peaked distributions

Summarising:

(1) probability distribution $P_k(x)$,
(encodes your knowledge about *x*)

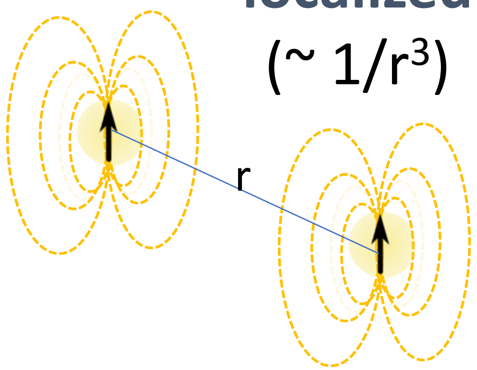
(2) select value of k to make the most change to P(x),
(we use Fisher information).

(3) perform your measurement with optimal settings,
getting outcome m.

(4) update Pk(x) using Bayes rule, for outcome m

Our application: spin-based quantum sensors

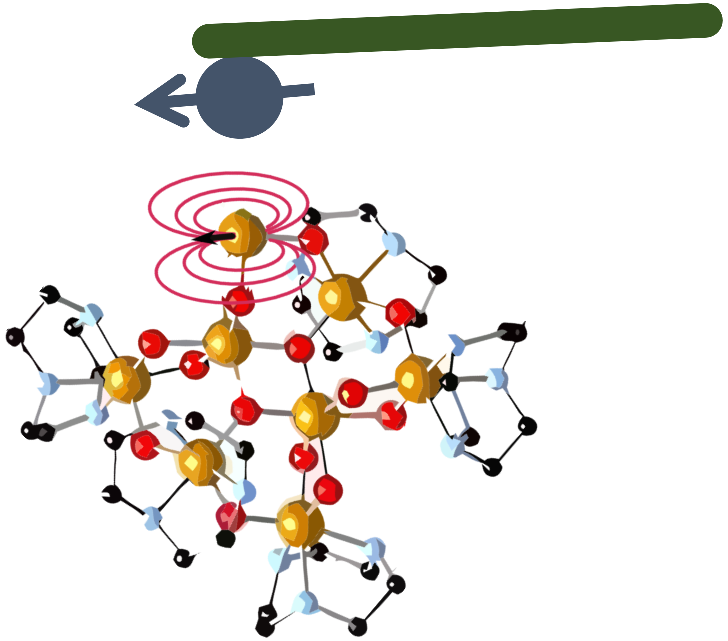
magnetic interaction is
'localized'
(~ $1/r^3$)



spins need to be **VERY** close to interact

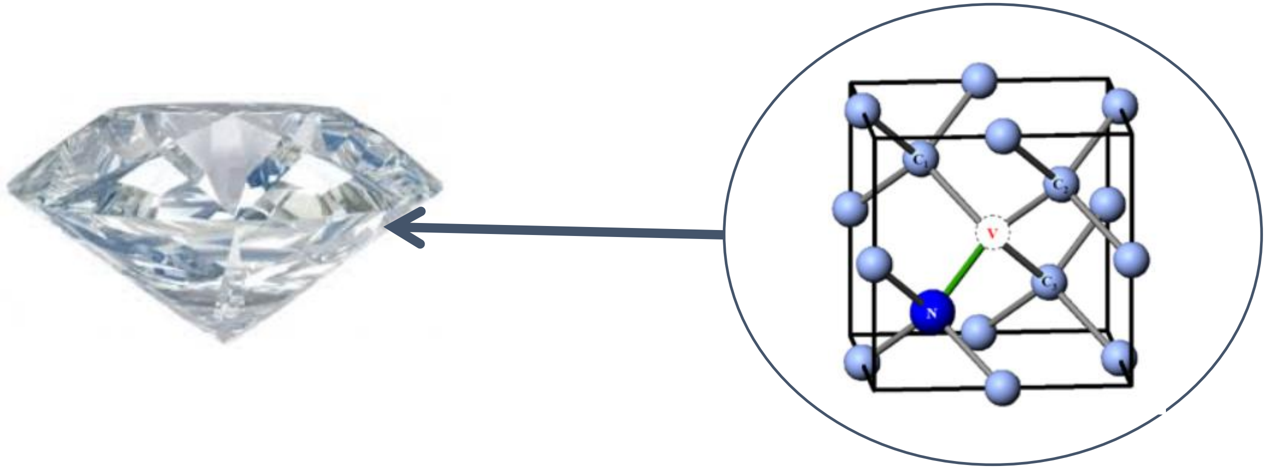
For (arbitrary) interaction strength of 100 kHz:

- e-spin/e-spin, r = 15 nm
- e-spin/^{13}C nuclear spin, r = 1.2 nm

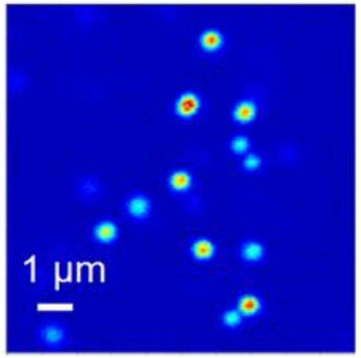


Since spins only interact when they are close by, one can achieve nanoscale spatial resolution!

Our system: nitrogen-vacancy (NV) centre in diamond

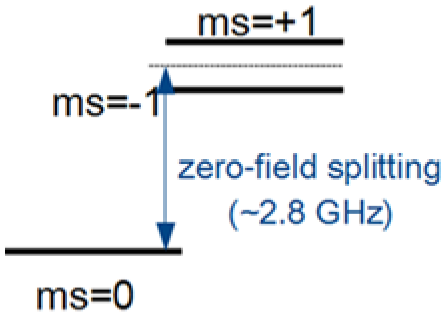


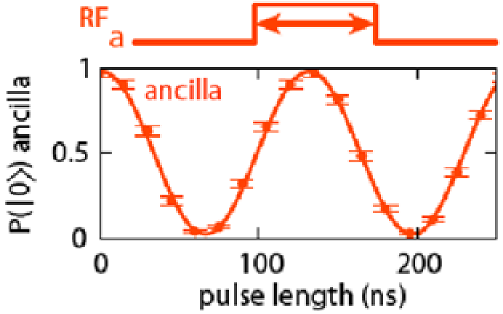
optically-active



spin state can be read-out by a change in photoluminescence

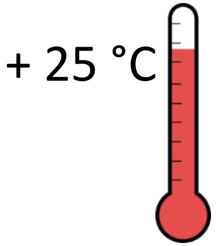
paramagnetic ground state (S=1)





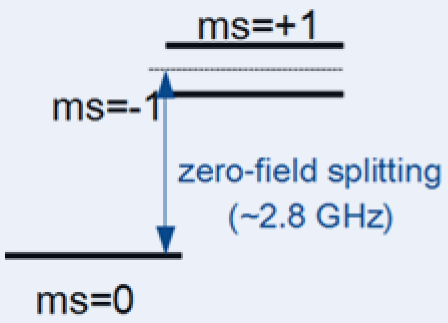
spin manipulation by microwave pulses

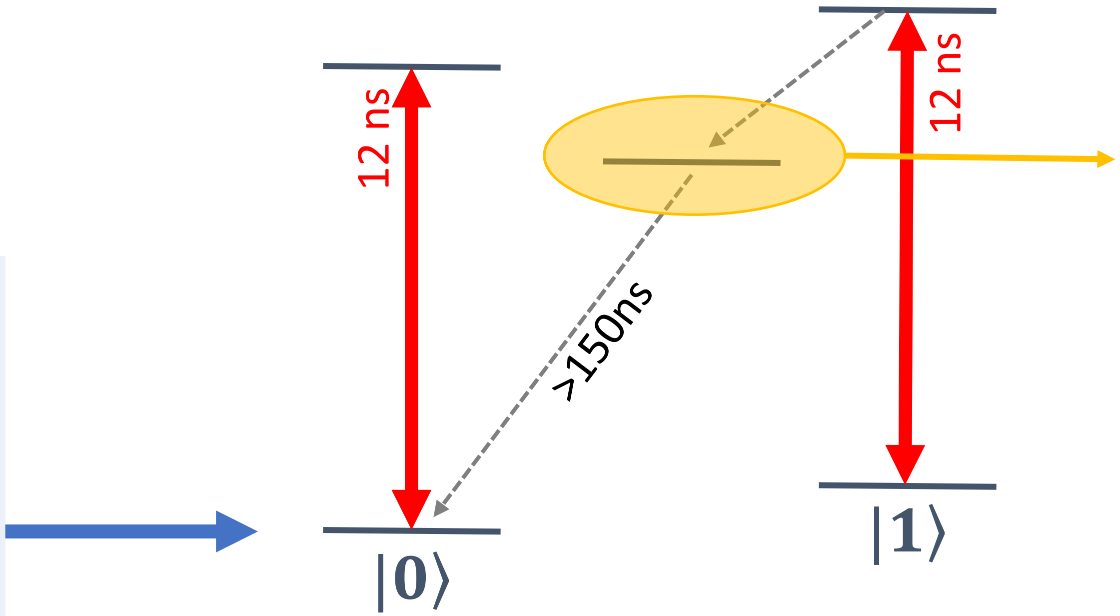
What’s unique about the NV centre in diamond?



Electron spin can be polarised and readout at room temperature:

ground state:
S=1 electronic spin





metastable state (lifetime >150ns)
with strong spin-orbit coupling

Quantum sensing modalities

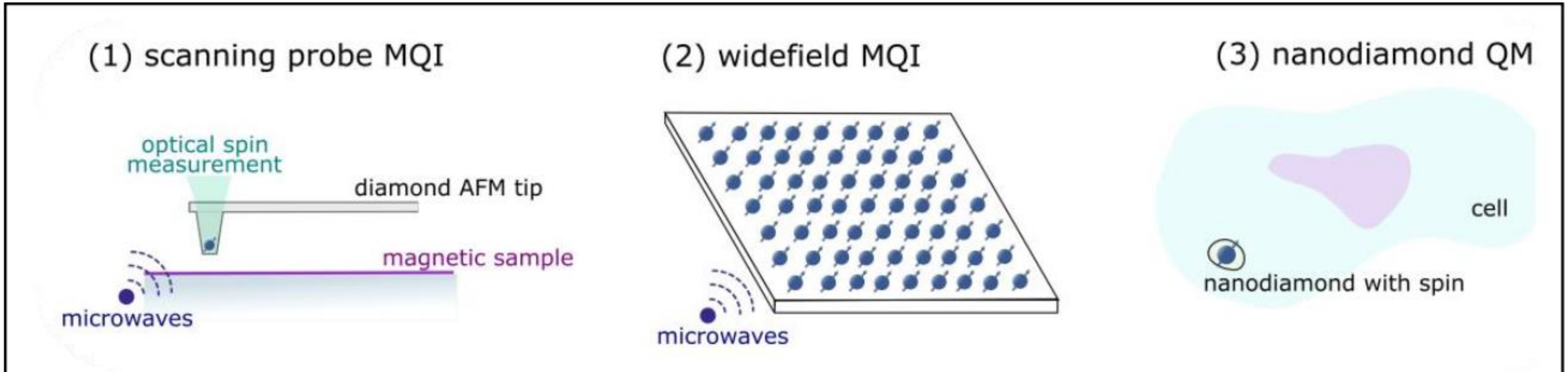
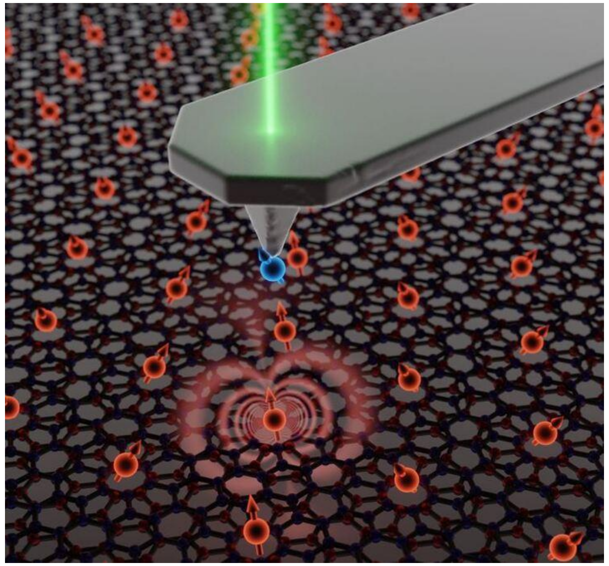


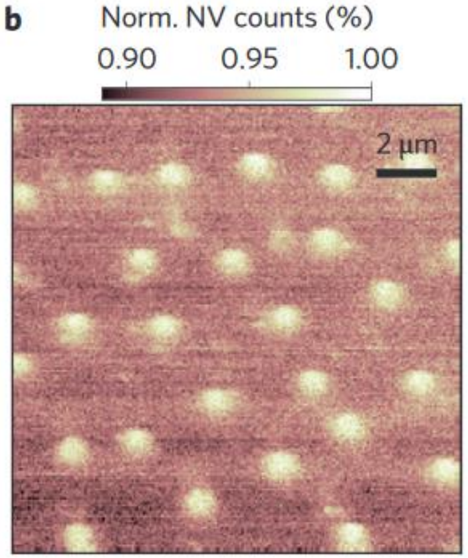
Fig 1: *different configurations for quantum magnetic imaging (QMI).*



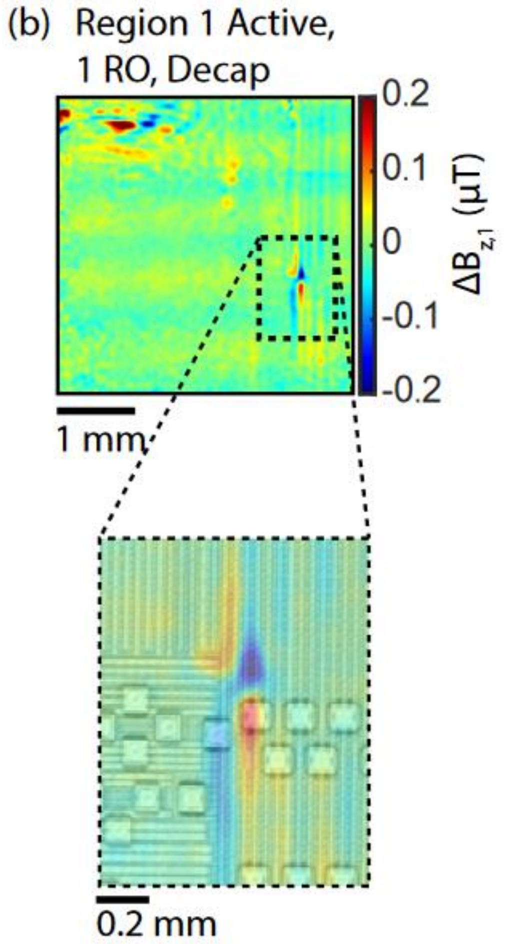
Heriot-Watt’s ESRC Quantum Magnetometry facility,
DSTL project “Quantum magnetometry of complex 2D materials”

Nanoscale magnetic fields

Single spins are already being used as sensors in different fields:



Imaging vortices in superconductors
Nature Nanotech 11 (2016)



Imaging of currents in an electronic chip
Phys. Rev. Applied 14, 014097 (2020)

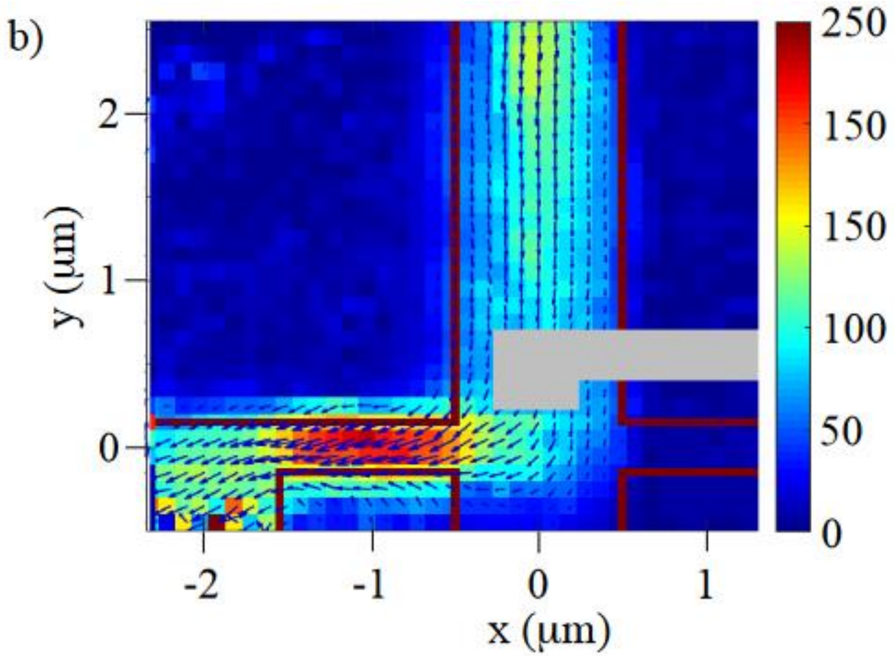
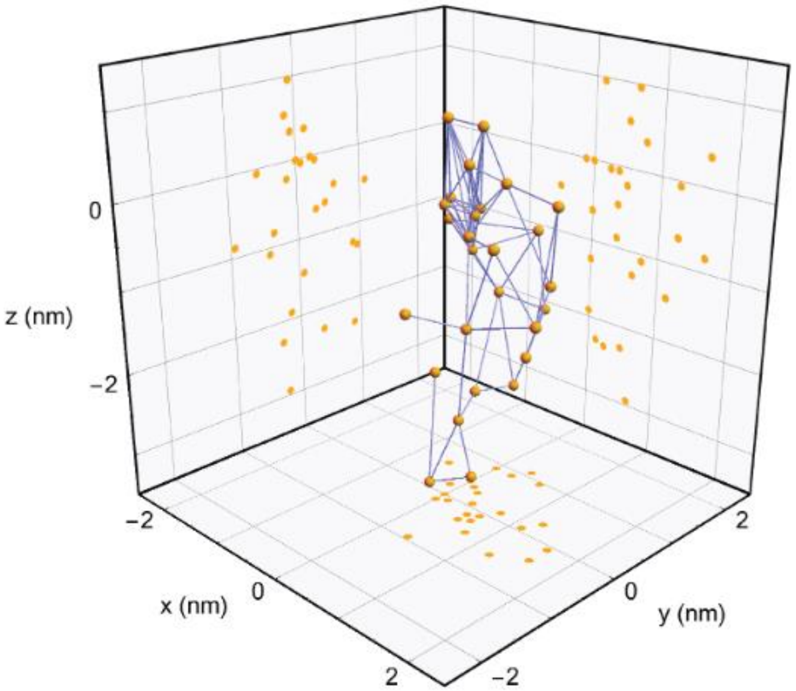


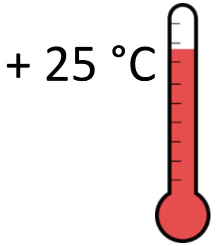
Image "viscous" flow of Dirac electron fluid in graphene
Nature 583, 537 (2020)



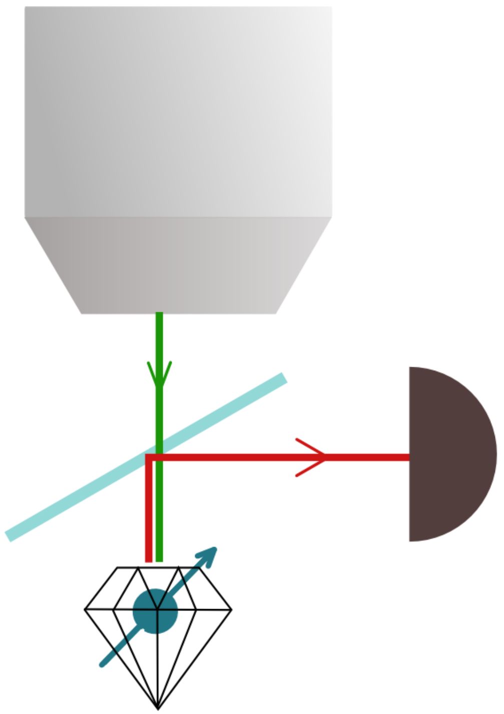
Detection of 27 individual ^{13}C nuclei in diamond
Nature 576, 411 (2019)

…just few examples from the tens of papers published every year

NV centres in diamond: room temperature readout

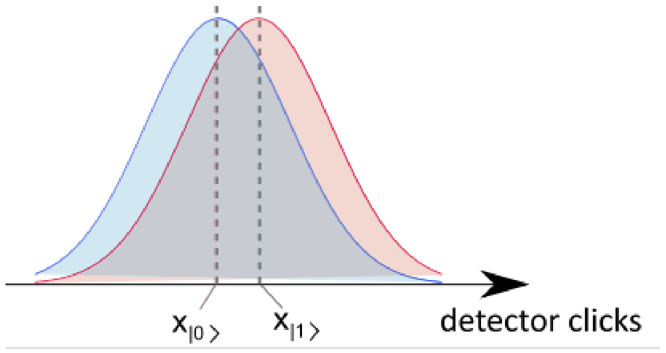


Recipe:
- shine green laser
- collect photoluminescence



more luminescence (~0.03 photons/shot) → **spin in $|0\rangle$**

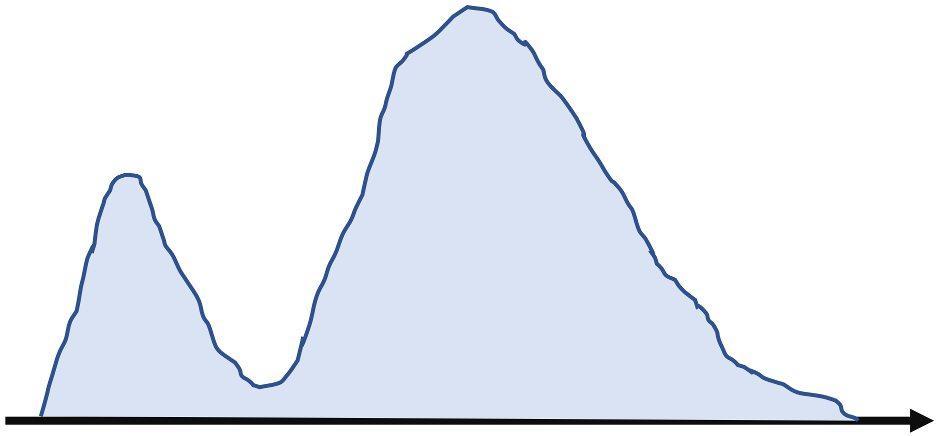
less luminescence (~0.02 photons/shot) → **spin in $|1\rangle$**



A single shot does not give us the spin info, we need to **repeat R (e.g. R=10,000) times**

Bayesian framework:
we update using "we detected r photons in R trials"

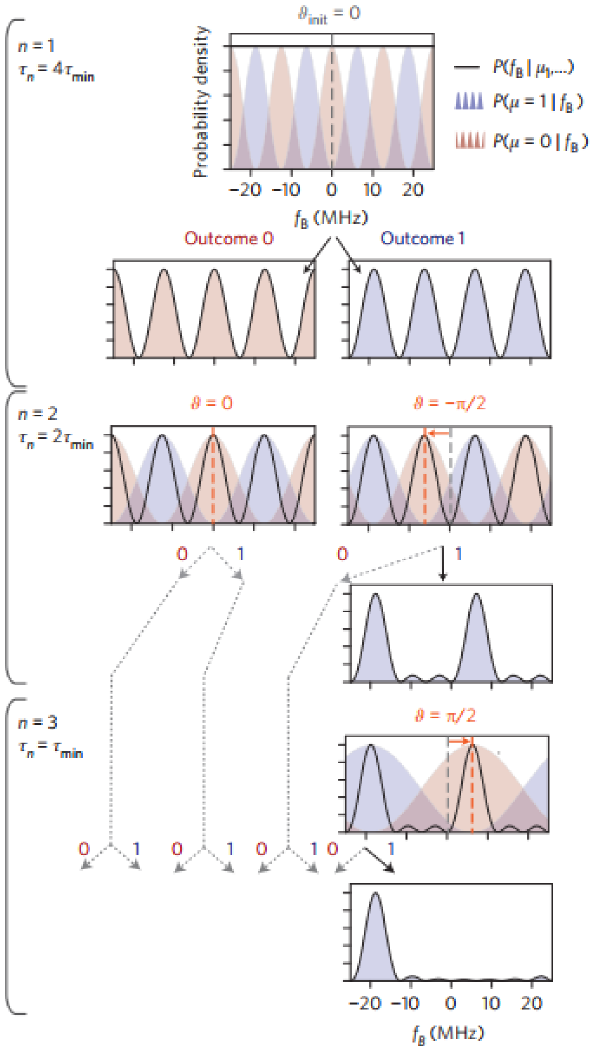
Implementation: how do you store/process the probability distribution?



discretise probability distribution $\{x_i\}$
(more obvious way: uniform discretization)

store $\{x_i\}$ in memory (not a big deal)

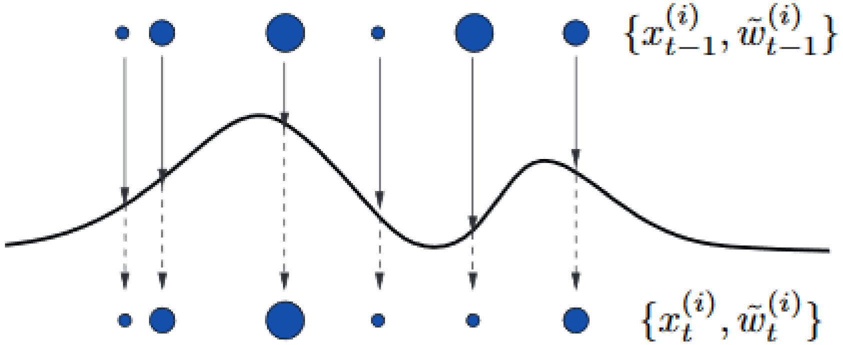
after each measurement, update all $\{x_i\}$
(complexity O(N))



when you start, you know nothing so you need a broad range

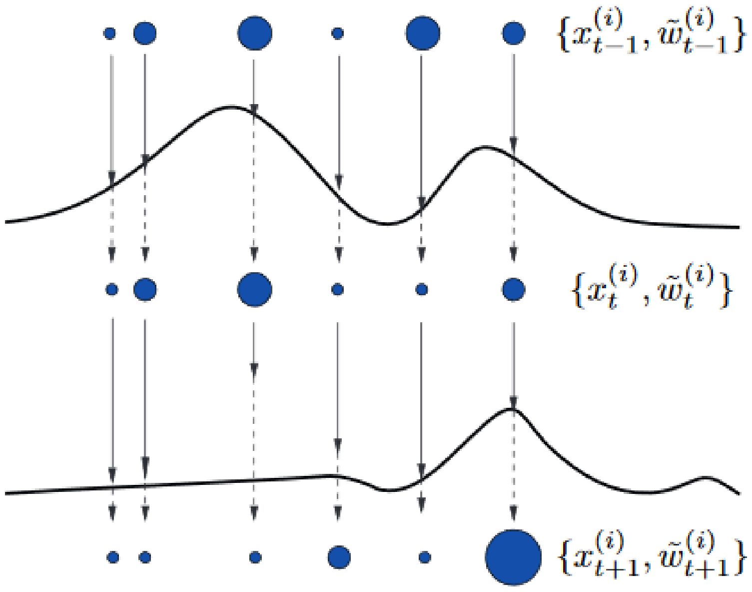
… as the measurement progresses, there are low-probability regions which are useless, but still occupy resources

Particle filtering (or sequential MonteCarlo)

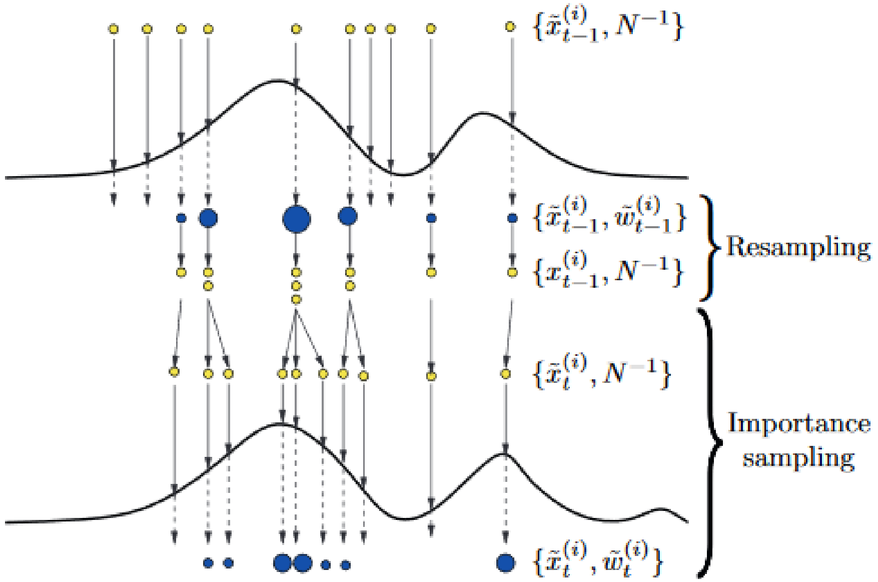


Adapted from (Doucet *et al.* , 2001)

Bayesian update:



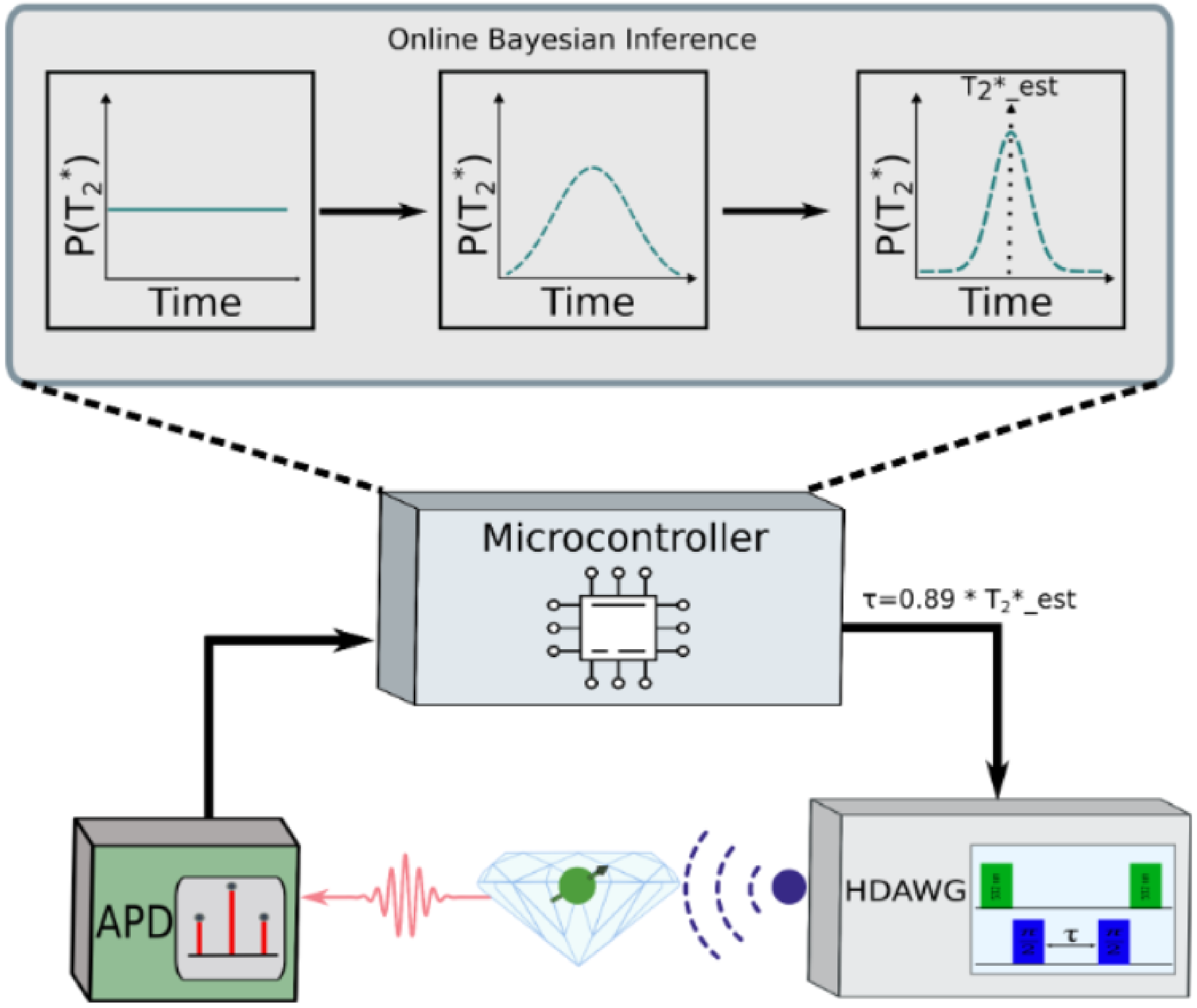
Re-sampling:



Used for quantum Hamiltonian learning by:

R Santagati et al, "*Magnetic-Field Learning Using a Single Electronic Spin in Diamond with One-Photon Readout at Room Temperature*", Phys Rev X (2019)

Experiment idea



Muhammad

Ben

Christiaan

Real-time Micro-controller: Adwin Pro II

AWG: Zurich Instruments HDAWG4

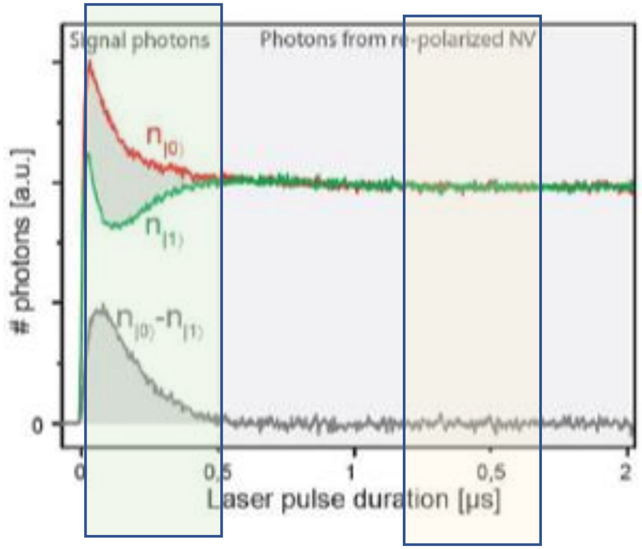
Real-time feedback loop duration: 50us

MJ Arshad et a, arxiv:2210.06103 (2022)

Electronics Detailed Schematic

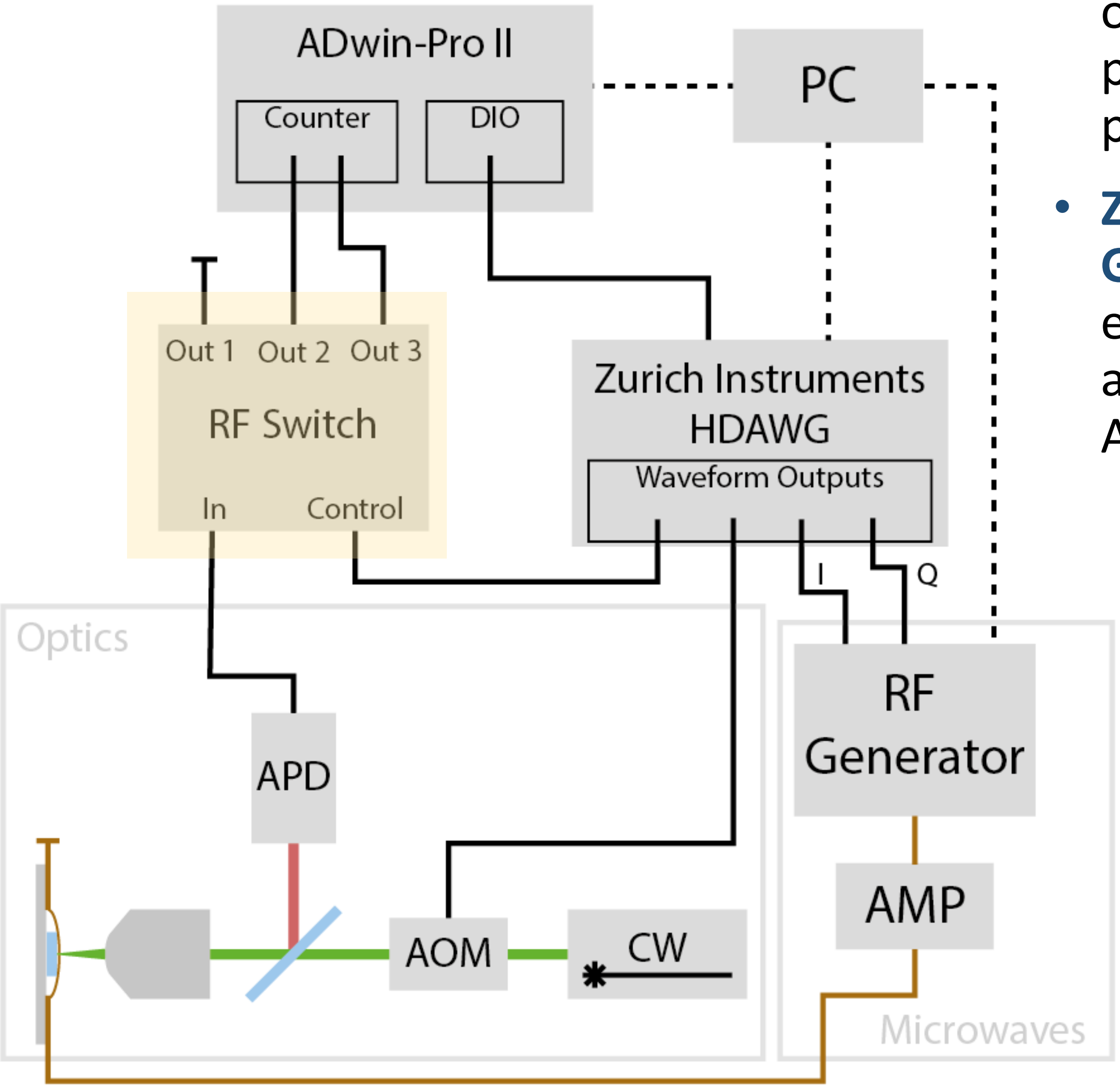
RF switch

Filters out signal and normalization in NV time-resolved PL:



Here we have a difference in photon count based on the spin state

Here we have no spin-related difference (we can use this to detect system drifts – "normalisation")



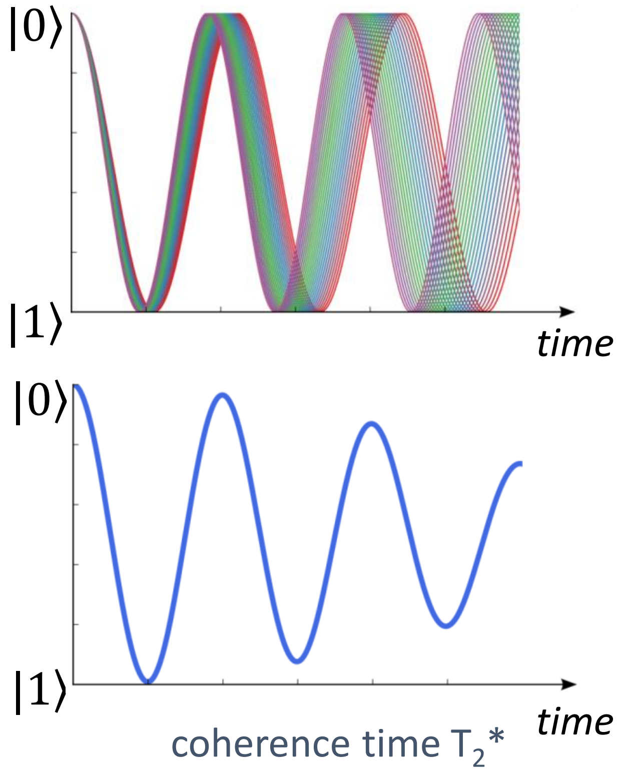
- Adwin (**microcontroller**) initiates experiment, reads out photon counts and provides optimised parameters
- **ZI Arbitrary Waveform Generator** controls the experimental apparatus and routes count signals to ADwin counters

Feedback loop: 50-100 microseconds

Example: measuring loss of quantum coherence

Decoherence:
quantum systems lose their "*quantumness*" by interacting with the environment.

Example: fluctuations in magnetic field induce fluctuations in spin precession frequency



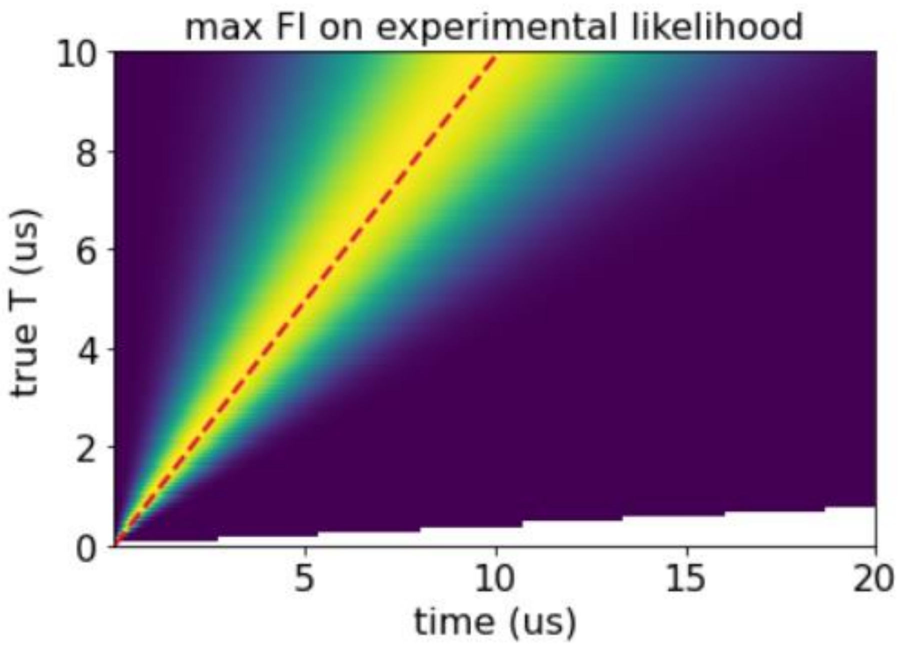
The loss of quantum coherence can be generally described as:

$$p(t) \propto \frac{1}{2}\left(1 - e^{-\chi(t)}\right)$$

$$\chi(t) \propto \left(\frac{t}{T_\chi}\right)^\beta$$

How do you adaptively choose best settings?

Our approach: simple "analytical" near-optimal max(FI)
(formula needs to be simple for adaptive choice to be fast so that **computations do not slow sensing down**)



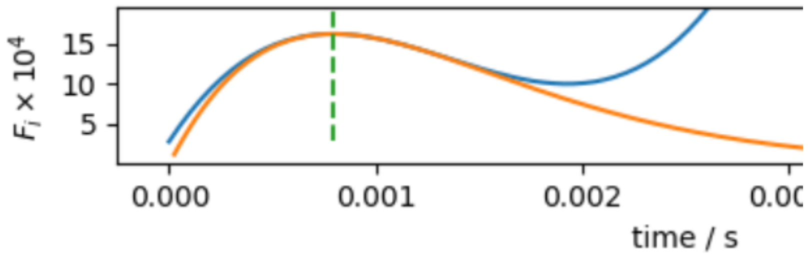
Fisher Information for $T1$:

$$F_i(t, T_1) = \frac{t^2}{T_1^4\left(e^{\frac{2t}{T_1}} - 1\right)}$$

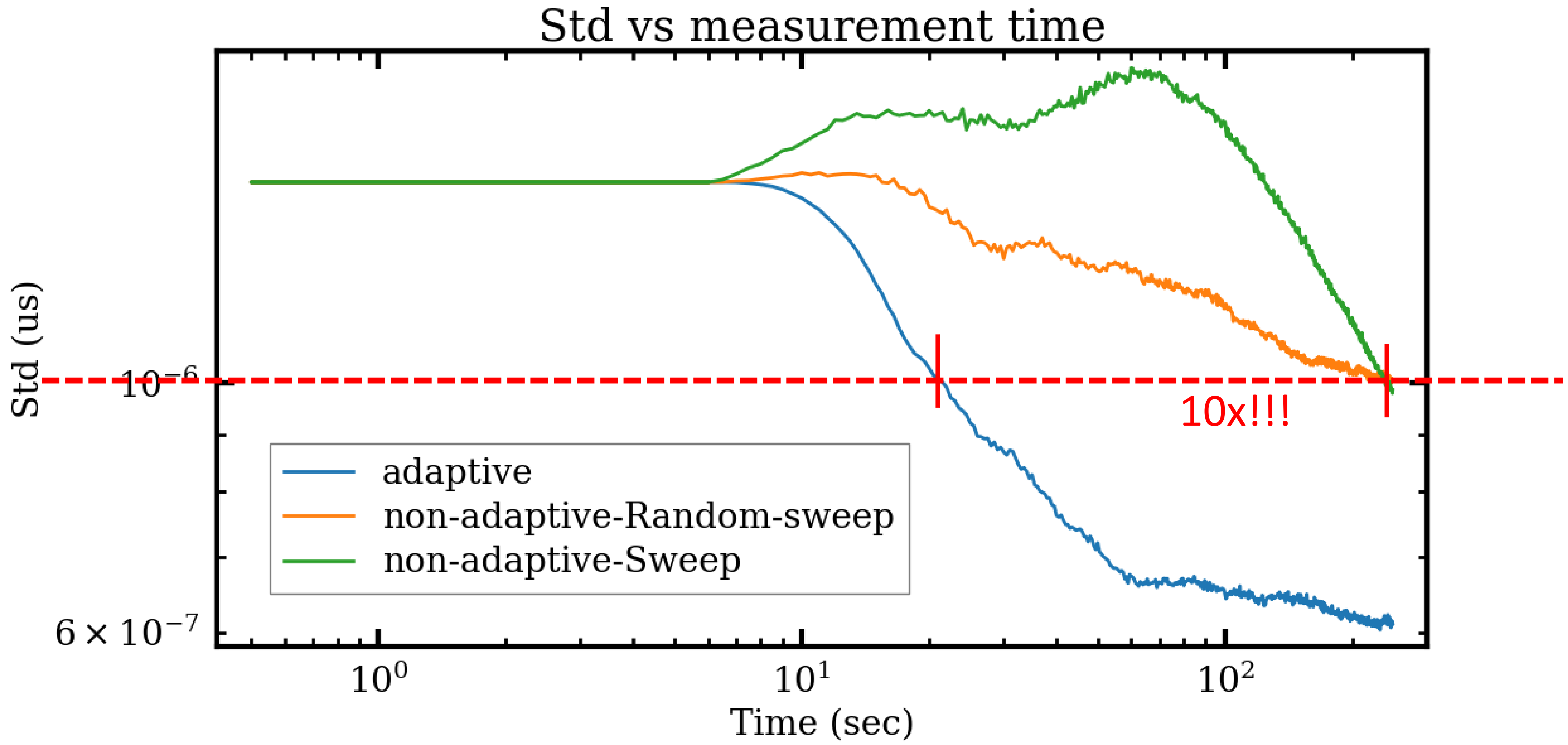
No analytical solution, Taylor expansion:

$$F_i(t, T_1) \approx \frac{0.028T_1^3 + 0.390T_1^2 t - 0.347T_1 t^2 + 0.085t^3}{T_1^5}$$

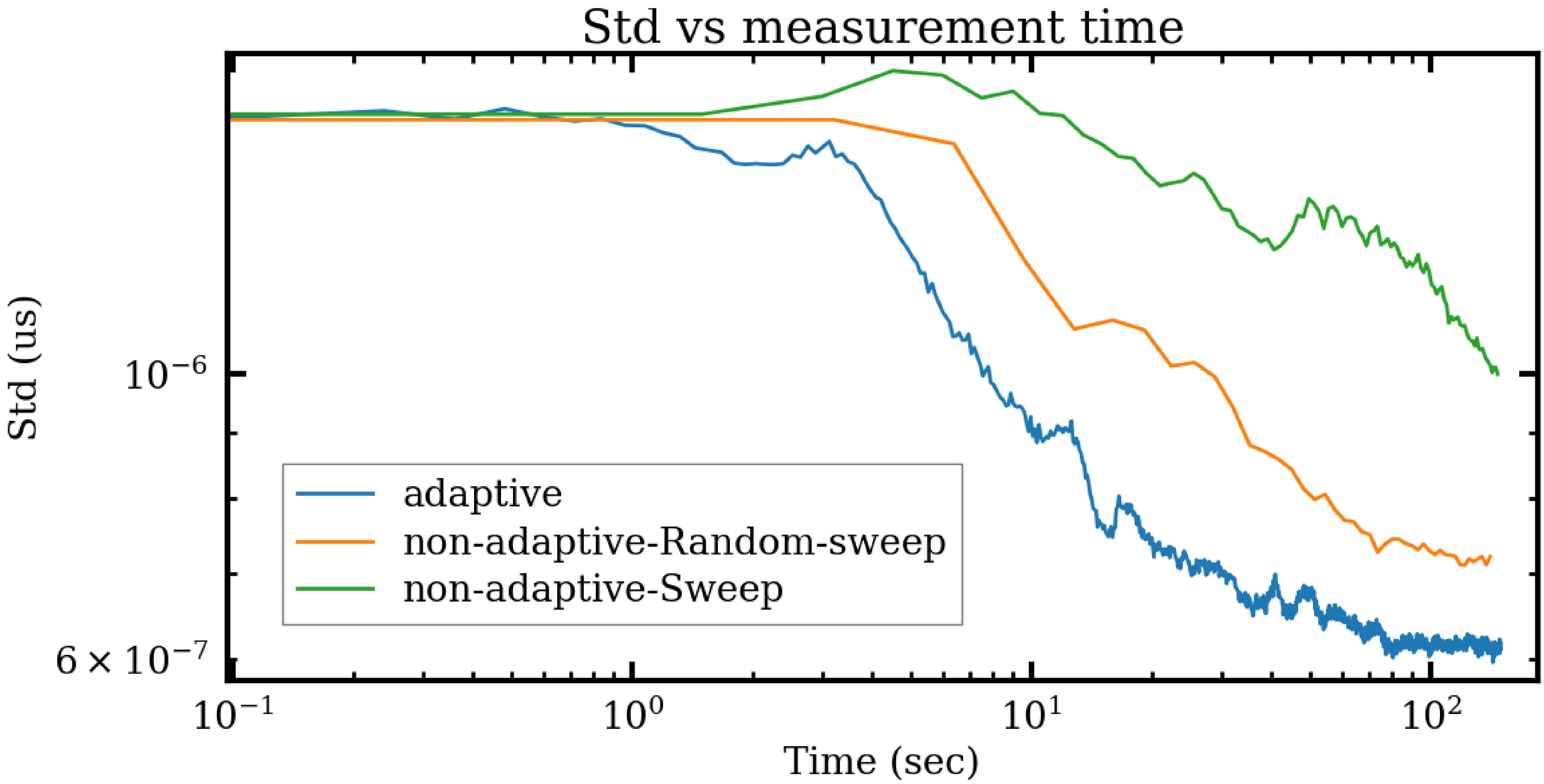
approximately: t_opt ~ 0.8*T1_est



Experimental T2* estimation (averaged-readout with $R=10^6$ reps)



Experimental T2* estimation (averaged-readout with R=10^5 reps)



What should we optimise when sensing time is not constant?

A longer measurement that yields the same sensitivity as a shorter one should be penalised!

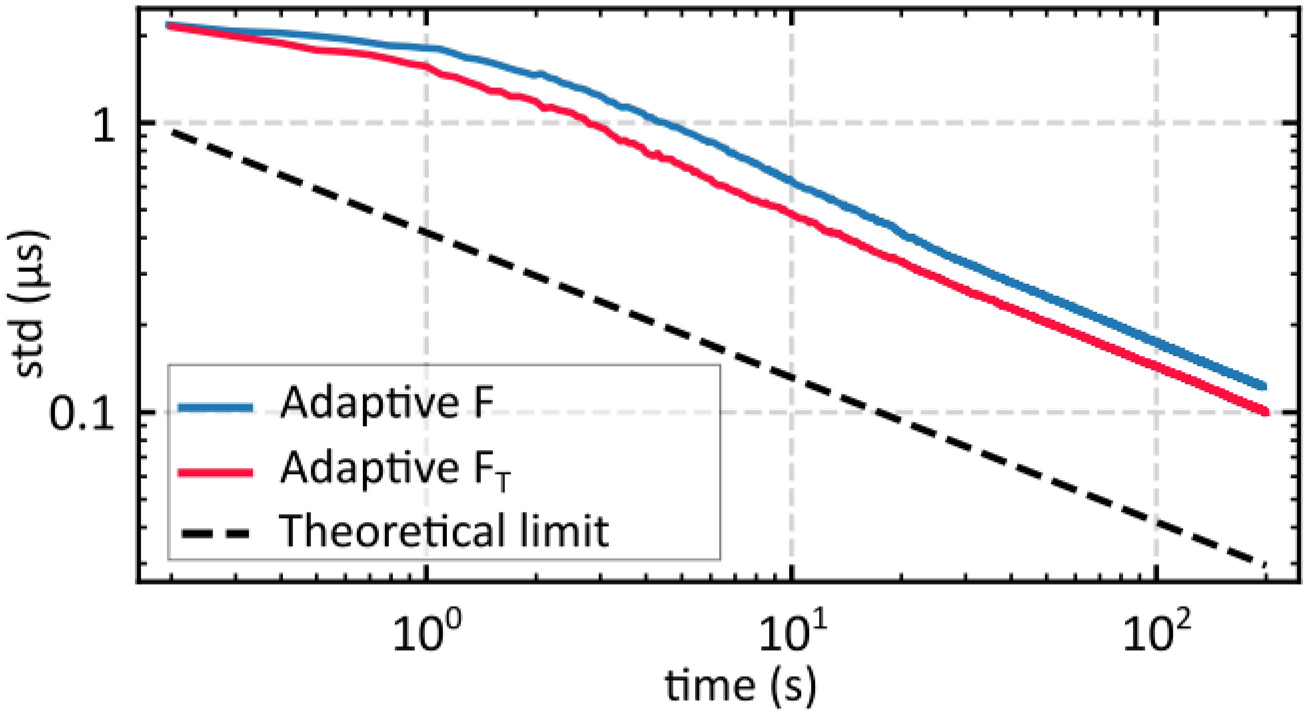
Instead of looking at variance, we can look at sensitivity, commonly defined as Var*T

From the Cramer-Rao bound:

$$\mathrm{Var} * T > \frac{1 * T}{FI}$$

Instead of maximising FI, let's maximise:

$$\mathbf{F_T = F/T}$$



Multi-parameter estimation

The loss of quantum coherence can be generally described as:

$$p(t) \propto \frac{1}{2}\left(1 - e^{-\chi(t)}\right)$$

$$\chi(t) \propto \left(\frac{t}{T_\chi}\right)^{\beta}$$

decay exponent provides information about the statistics of the noise acting on the spin sensors

Can we estimate β and T_χ simultaneously?

Multi-parameter estimation

Problem: the determinant of the Fisher information is zero!

Why are β and T_χ correlated?

They are not. But with just one sensing time, they become correlated (one equation with two unknowns)

Solution: use two sensing times!

Multi-parameter estimation

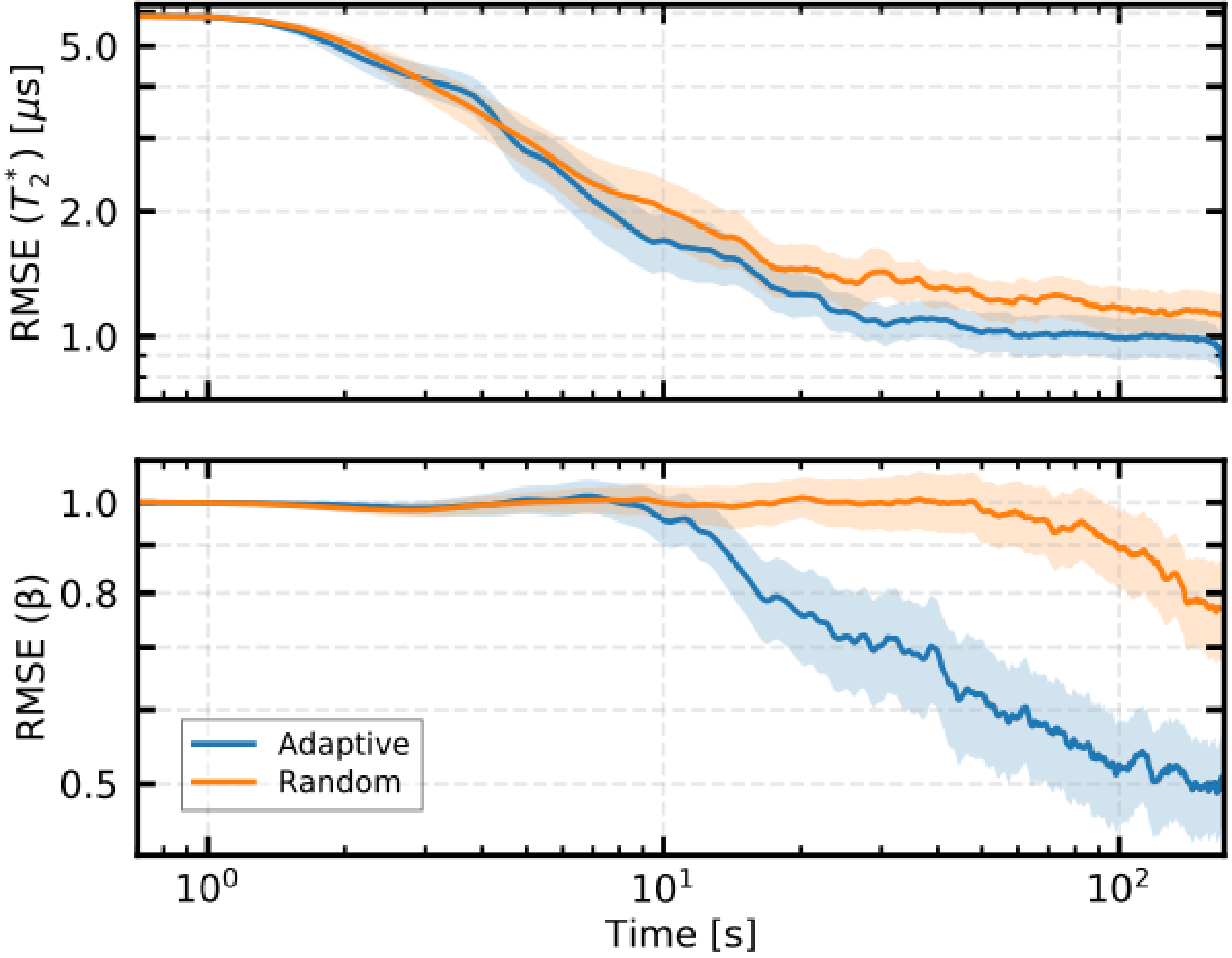
Determinant of the Fisher information matrix:

$$\det \hat{F}_{\mathcal{B}} = \frac{N^2 \left(\frac{\tau_0}{T_\chi}\right)^{2N} \left(\frac{\tau_1}{T_\chi}\right)^{2N} \left(\log^2\left(\frac{\tau_0}{T_\chi}\right) - 2\log\left(\frac{\tau_0}{T_\chi}\right)\log\left(\frac{\tau_1}{T_\chi}\right) + \log^2\left(\frac{\tau_1}{T_\chi}\right)\right)}{T_\chi^2 \left(-\exp\left[2\left(\frac{\tau_0}{T_\chi}\right)^N\right] - \exp\left[2\left(\frac{\tau_1}{T_\chi}\right)^N\right] + \exp\left[2\left(\frac{\tau_0}{T_\chi}\right)^N + 2\left(\frac{\tau_1}{T_\chi}\right)^N\right] + 1\right)}$$

Simple approximation for its maximum:

$$\tau_{1_{opt}} = \begin{cases} 0.313\tau_0 + 1.04\hat{T}_\chi, & \text{if} \quad \tau_0 < 0.83\hat{T}_\chi \\ 0.7\tau_0, & \text{if} \quad 0.83\hat{T}_\chi < \tau_0 < 0.96\hat{T}_\chi \\ 0.109\tau_0 + 0.55\hat{T}_\chi, & \text{if} \quad 0.96\hat{T}_\chi < \tau_0 \end{cases}$$

Multi-parameter estimation



MJ Arshad et a, arxiv:2210.06103 (2022)

What do we need this for?

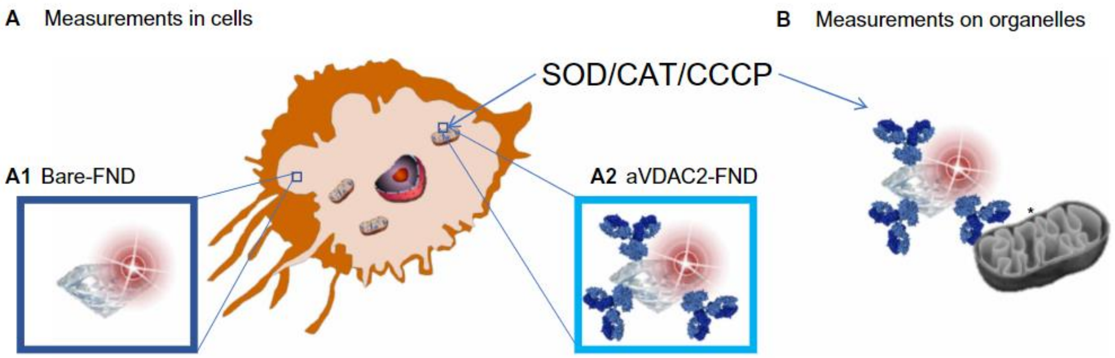
Decoherence of the central spin can give us information about the local environment

SCIENCE ADVANCES | RESEARCH ARTICLE

APPLIED PHYSICS

Quantum monitoring of cellular metabolic activities in single mitochondria

L. Nie[1†], A. C. Nusantara[1†], V. G. Damle[1], R. Sharmin[1], E. P. P. Evans[1], S. R. Hemelaar[1], K. J. van der Laan[1], R. Li[1], F. P. Perona Martinez[1], T. Vedelaar[1], M. Chipaux[2*], R. Schirhagl[1*]



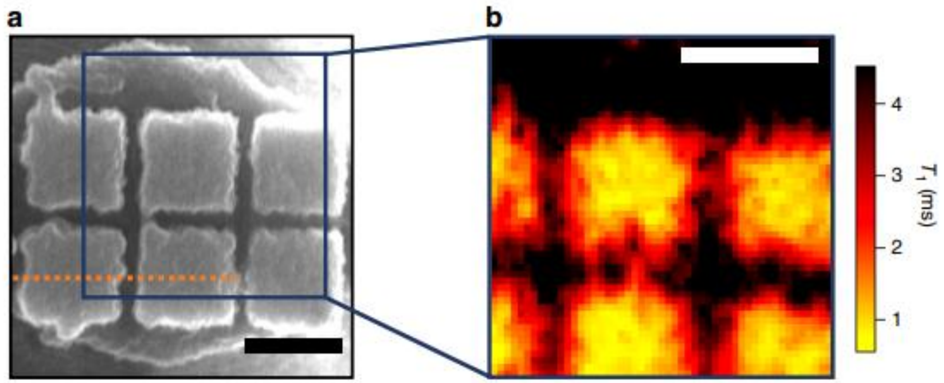
nature COMMUNICATIONS

ARTICLE

DOI: 10.1038/s41467-018-04798-1 OPEN

Nanoscale electrical conductivity imaging using a nitrogen-vacancy center in diamond

Amila Ariyaratne[1], Dolev Bluvstein[1], Bryan A. Myers[1] & Ania C. Bleszynski Jayich[1]



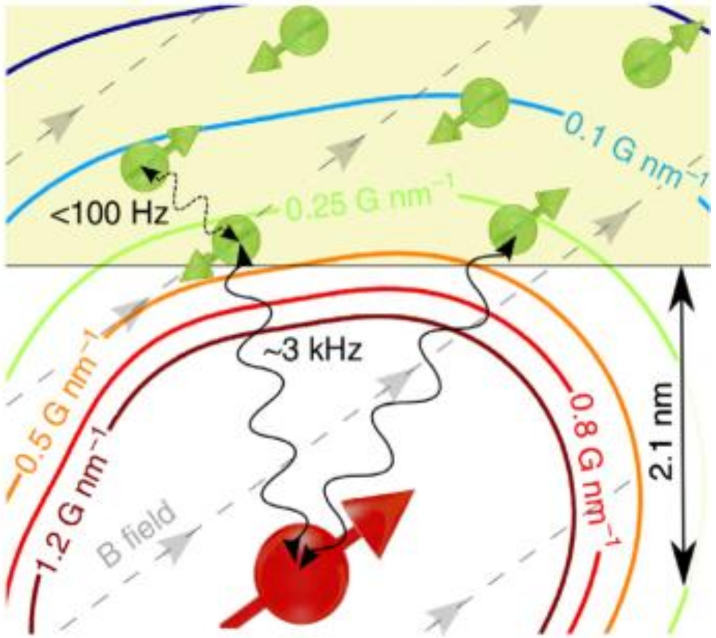
What’s the longer-term vision for this?

Our lab's goal: adaptive automated nanoscale magnetic resonance

Detecting nuclear spins is important…

Current limits:
- volume 40 um^3
- number spins: 10^{13} $Hz^{1/2}$

Solution: go NANO!

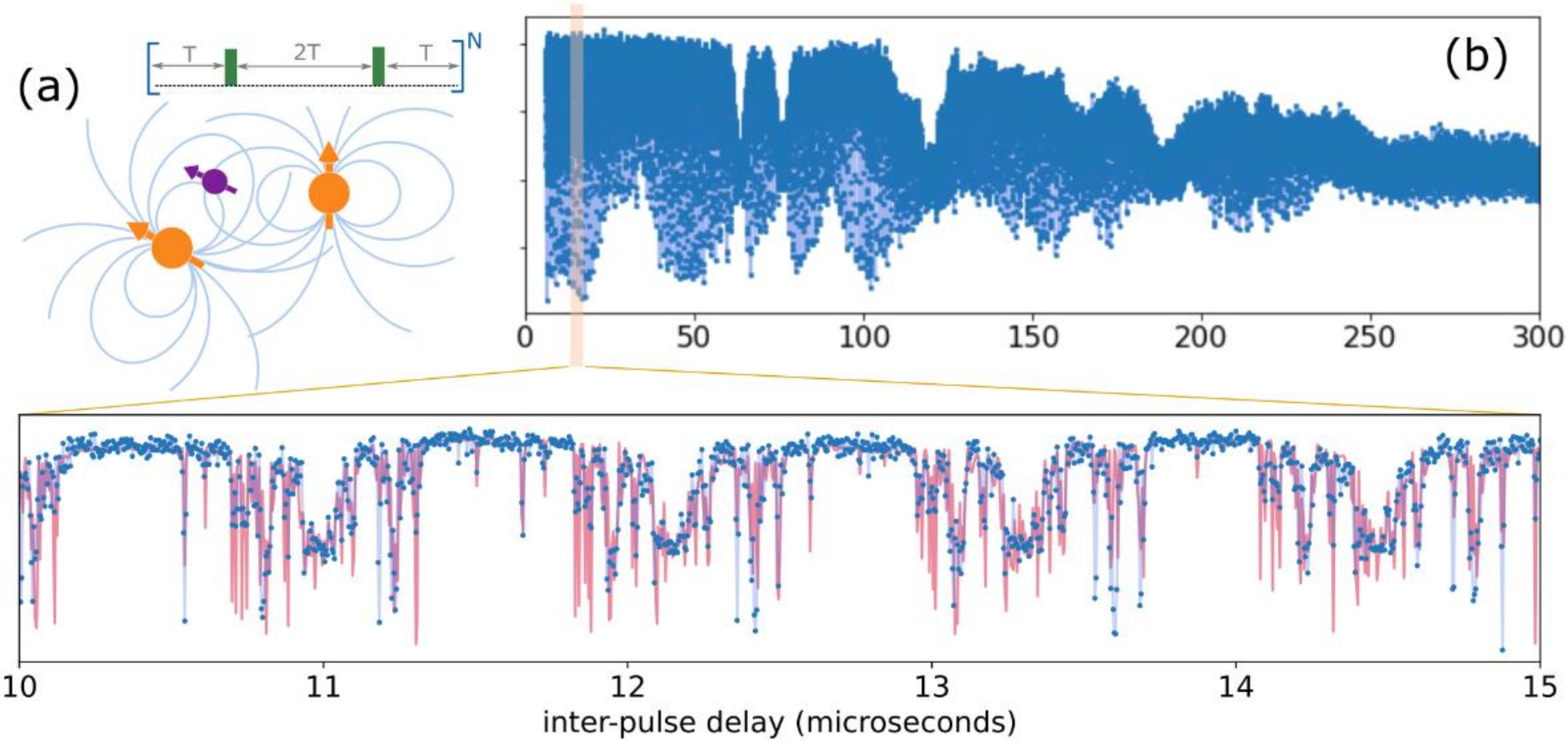


Use a single spin as **nearby** quantum sensor and detect nuclear spins by their dipolar coupling (statistical polarisation!)

See work from Taminiau (Delft), Degen (EH), Wrachtrup (Stuttgart)

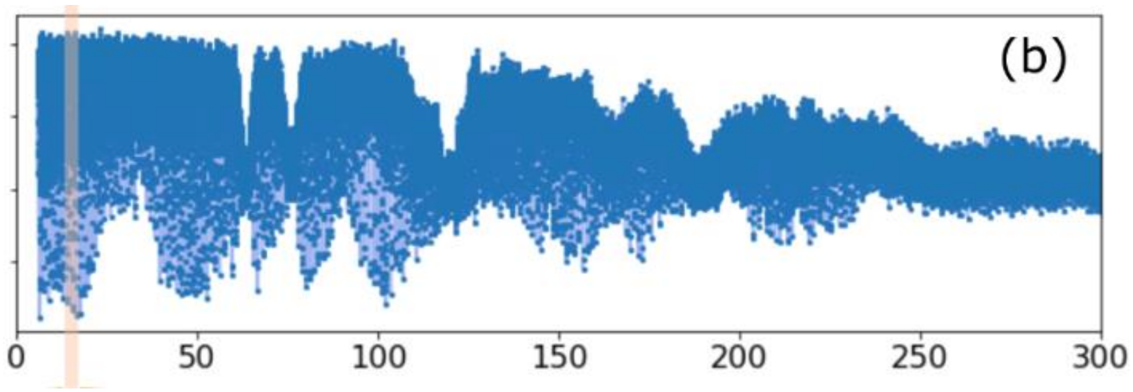
Our lab's goal: adaptive automated nanoscale magnetic resonance

The signal from many individual nuclear spins is complex, data acquisition is time-consuming:



(data from Taminiau group)

Our lab's goal: adaptive automated nanoscale magnetic resonance



(1) Can we adaptively optimise data taking for each point?

(2) Can we adaptive take only the points that give more information?

(3) How do we automatically fit the data and then link hyperfine values to position (need DFT prior)?

Automating physics:
Learning models of quantum systems from data

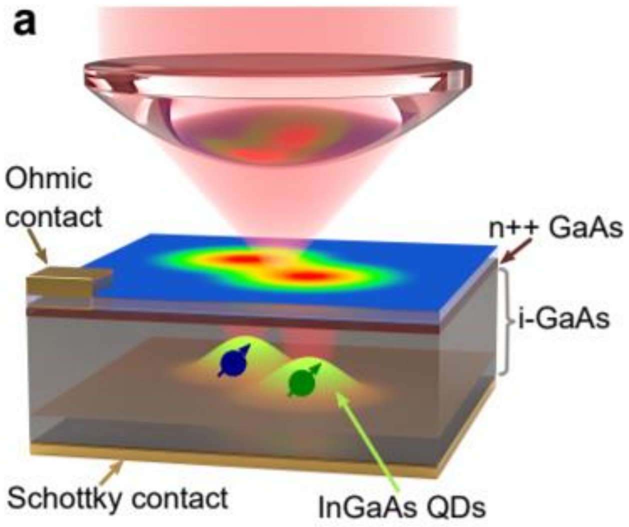
Stewart Wallace

Erik Gauger
(open quantum
system theorist)

Yoann Altmann

Can we learn the model for two "cooperatively-emitting" quantum emitters?

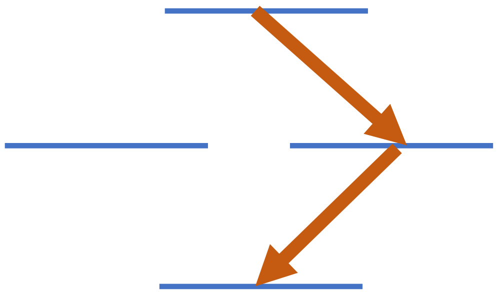
Two quantum dots brought into resonance:



Data from Gerardot's group at HWU
Zhe Xian Koong, Science Adv (2022)

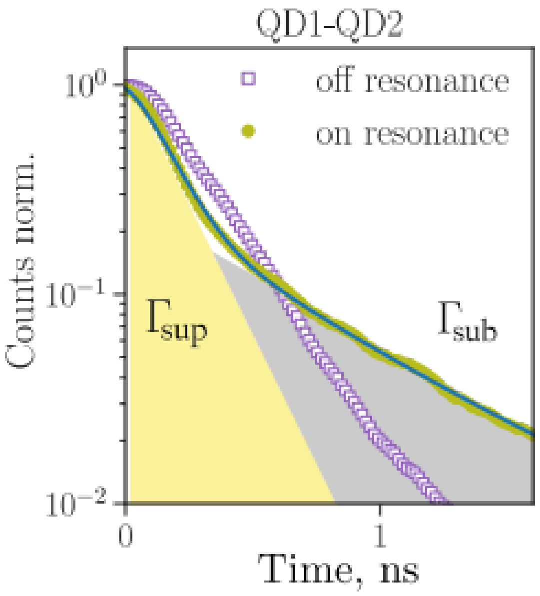
Independent emitters:

Super-radiance:

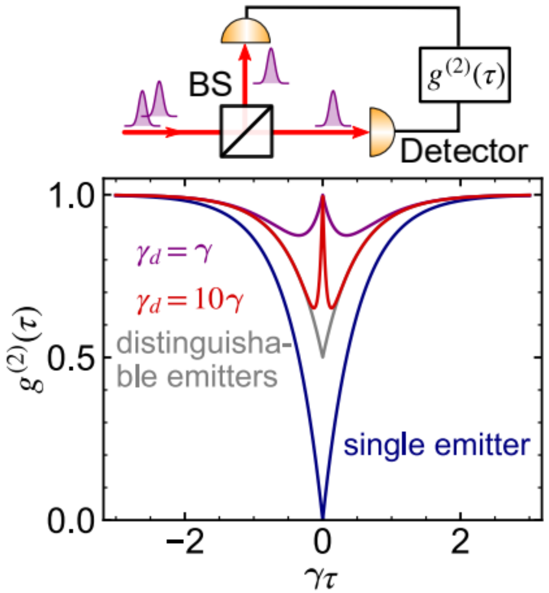


Signatures

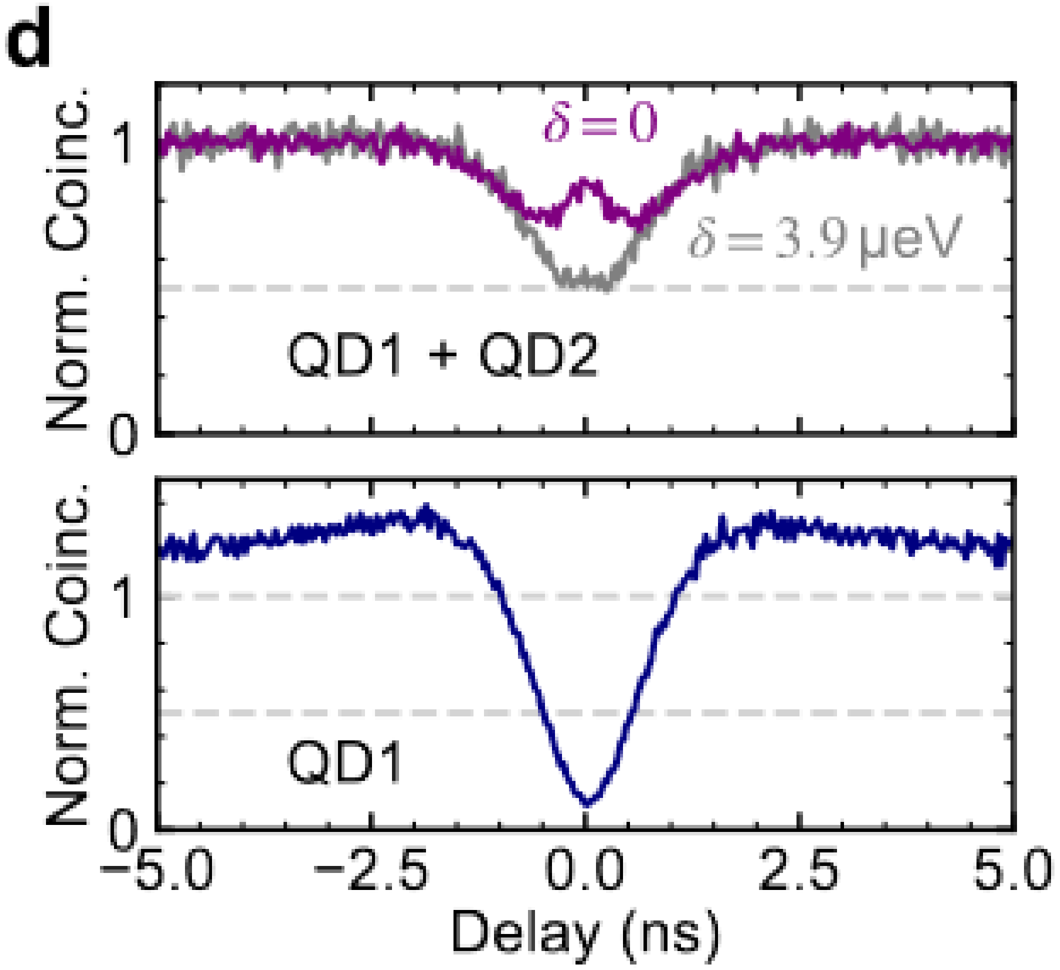
Lifetime:



Correlation (g2):



Can we learn the model for two "cooperatively-emitting" quantum emitters?



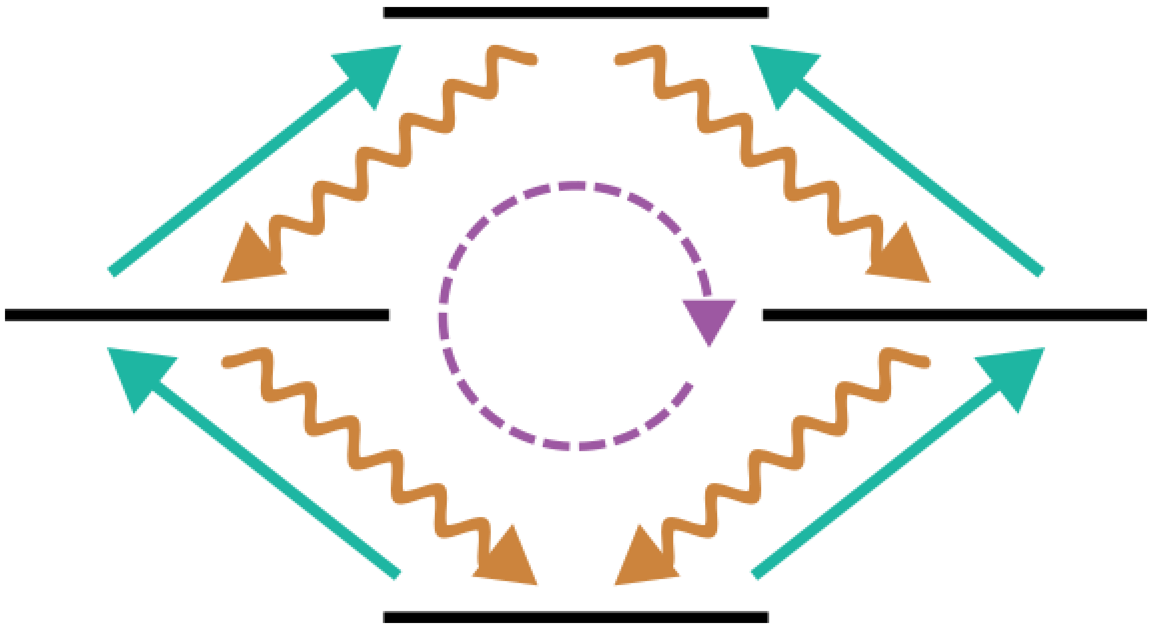
<u>Experimental data:</u>

- lifetime does not appear to be changed
- g2 shows a small peak

What is going on?

Data from Gerardot's group at HWU
Zhe Xian Koong, Science Adv (2022)

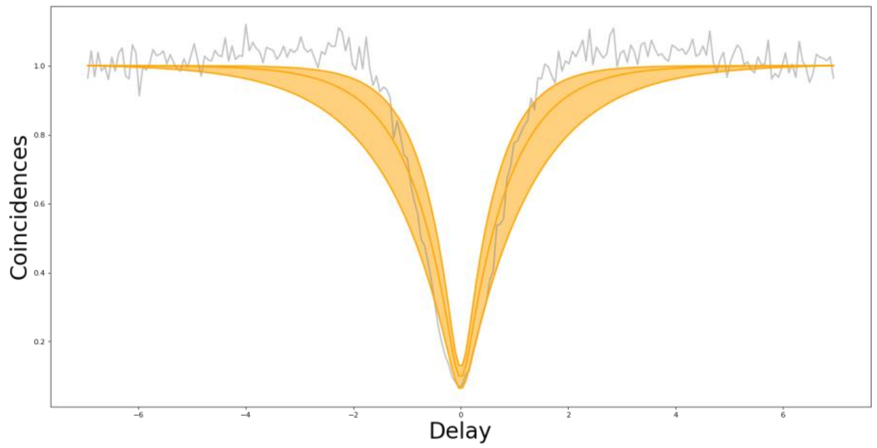
Again, using Bayesian inference
(just a more complex algorithm known as **Markov-chain MonteCarlo**)

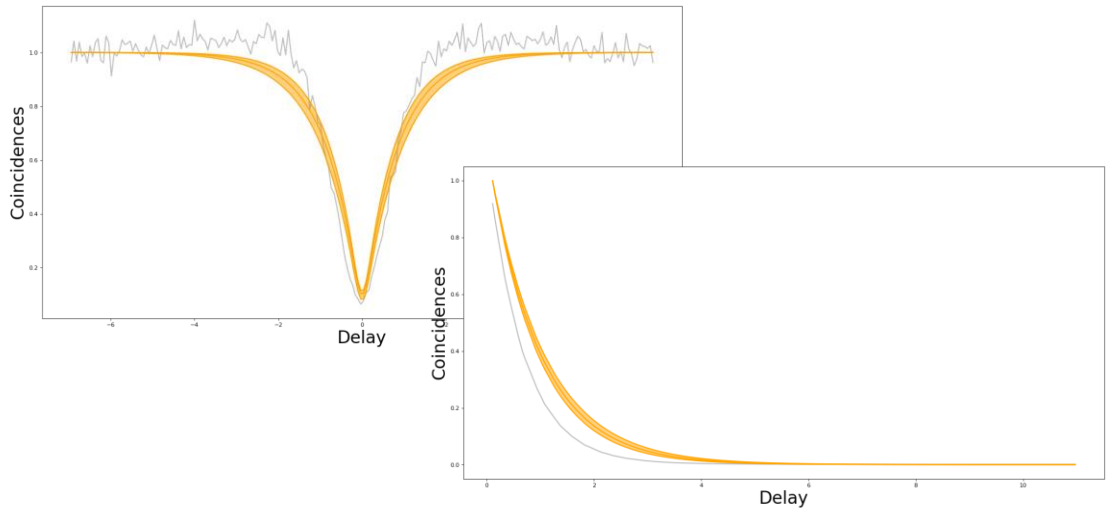


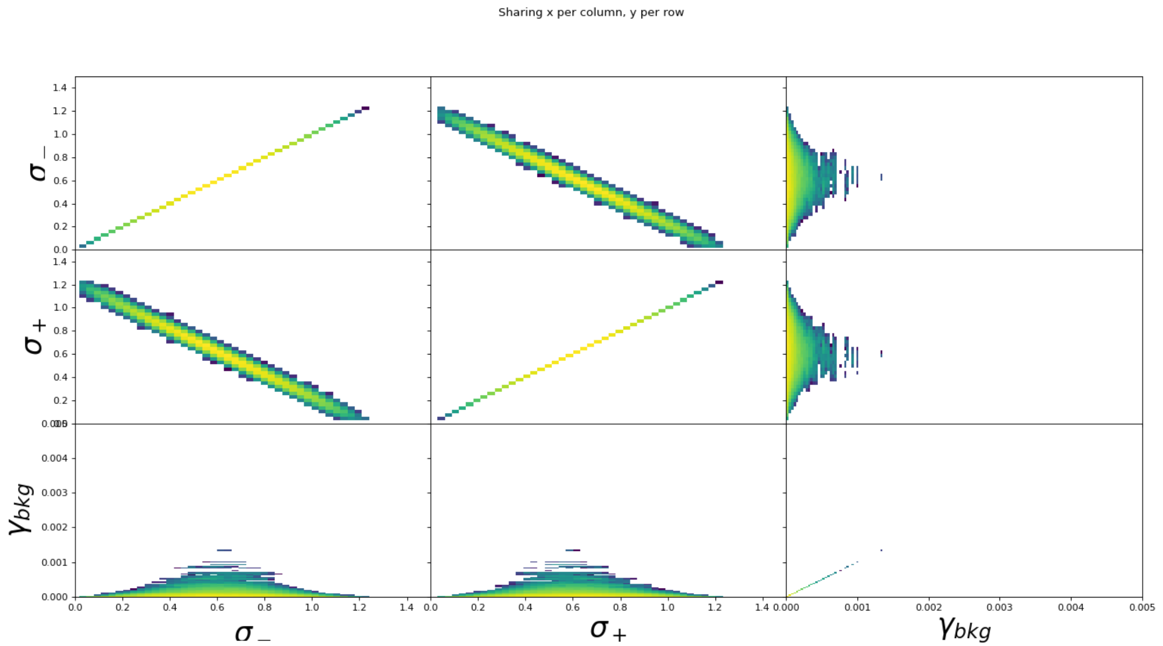
10 parameters = 9 rates + background

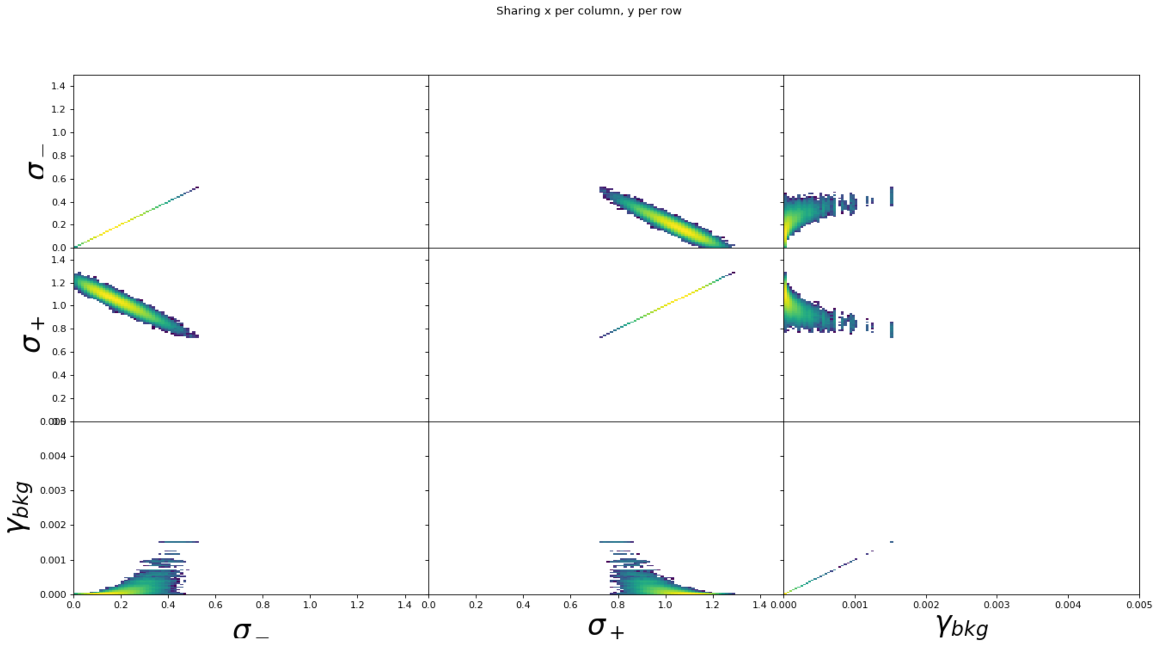
- method to sample a probability distribution
- a Markov-chain is used to walk across the parameters space
- every proposed move in parameters space can be accepted or rejected depending on how well it explains the data (Metropolis-Hastings algorithm)

Single emitter estimation (3 parameters)

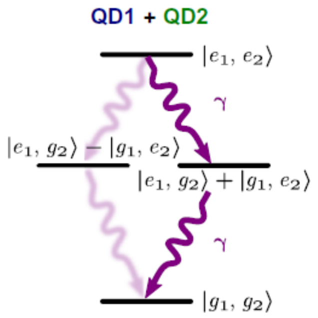


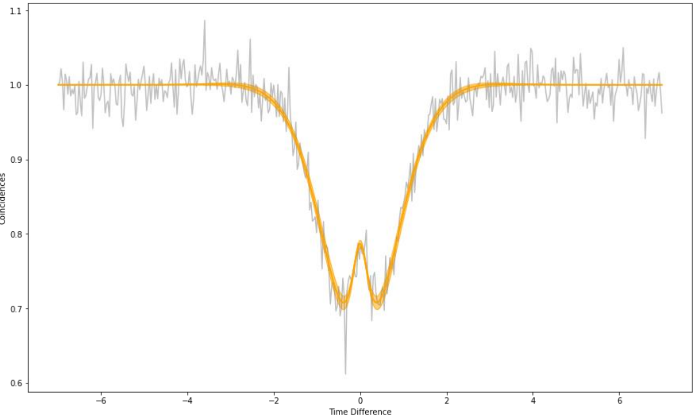


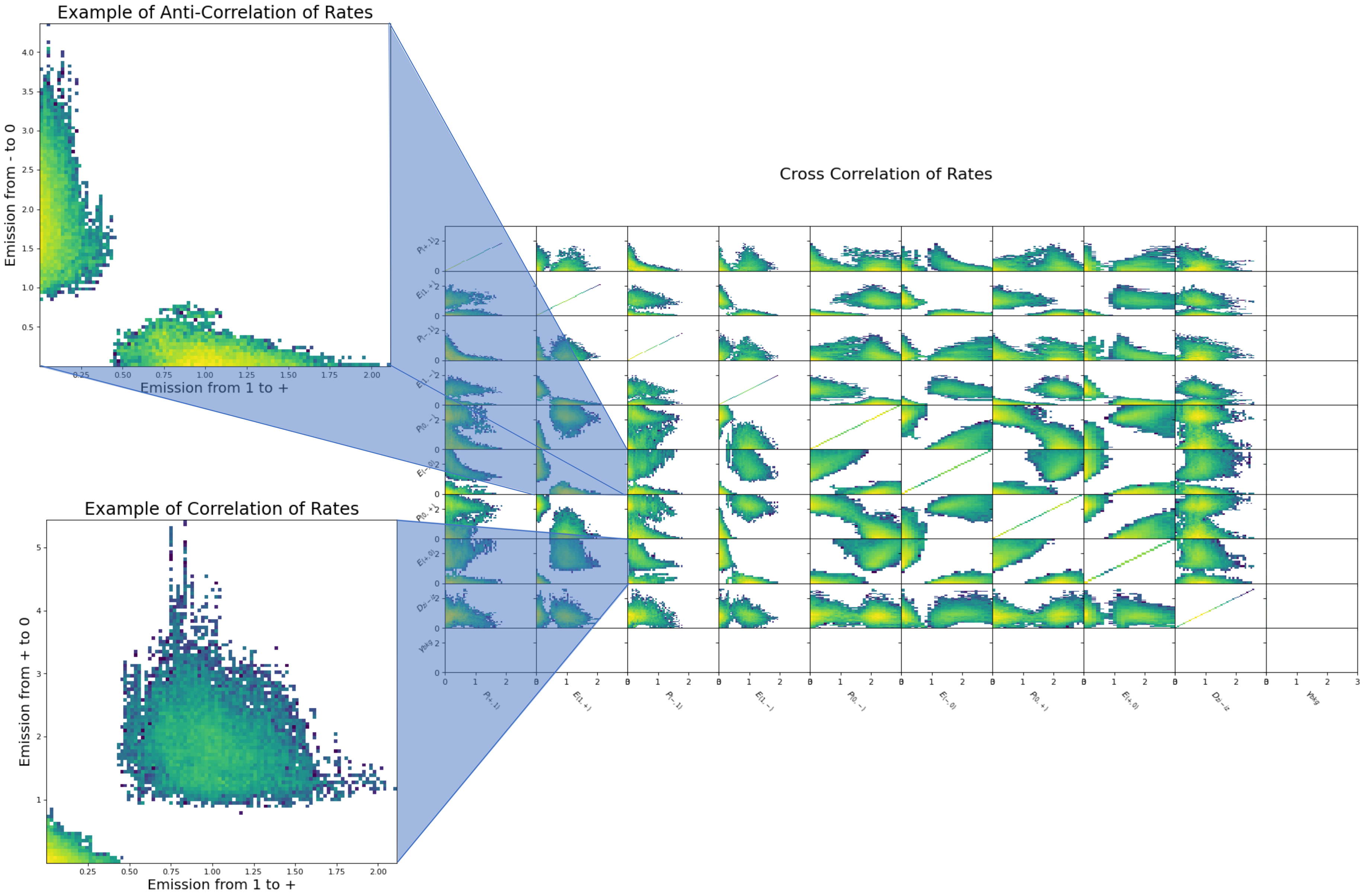




Two emitters estimation (10 parameters)







Can we learn a Lindblad master equation without making assumptions?

We are learning a differential equation of the form:

$$\dot{\rho} = -i[\hat{H}, \rho] + \sum_{v} (\hat{L}_v \rho \hat{L}_v^\dagger - \frac{1}{2}[\hat{L}_v^\dagger \hat{L}_v, \rho])$$

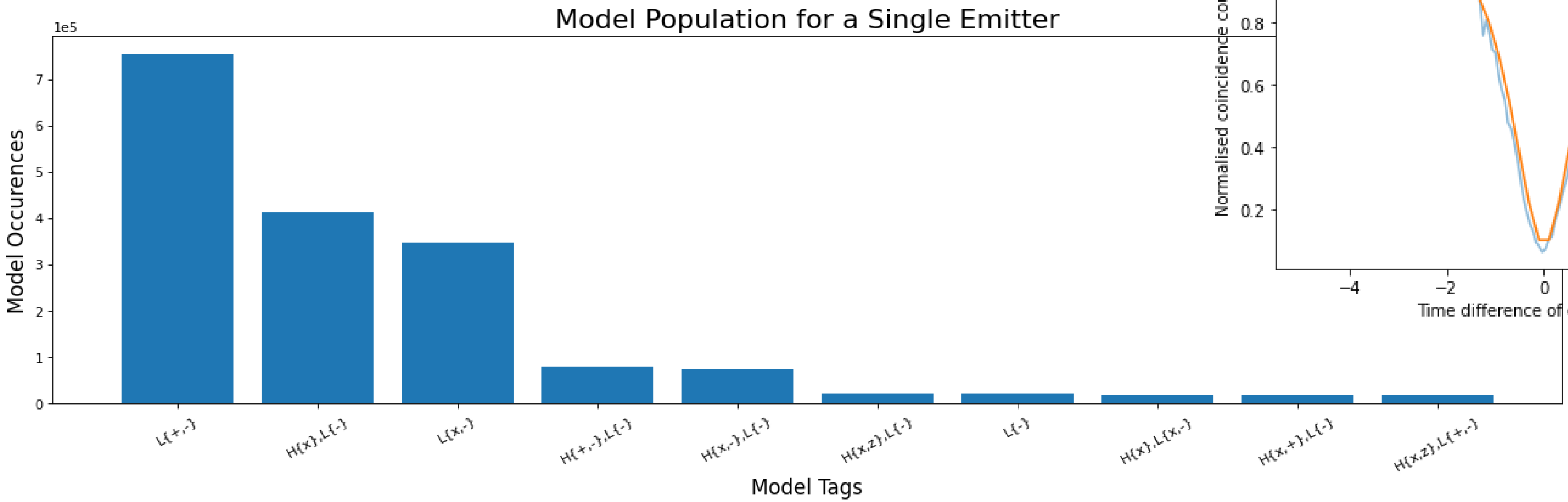
where the number of terms is unknown.

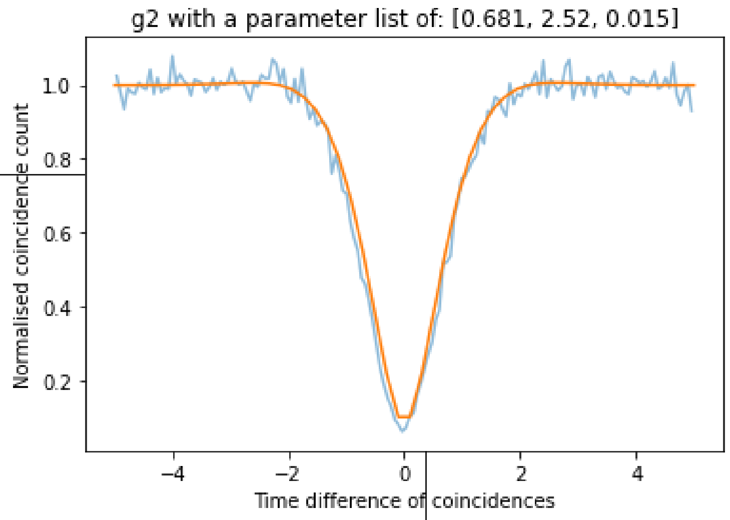
our approach:

reversible-jump MCMC

Can we learn a Lindblad master equation without making assumptions?

First results (showing goodness of fit for different models):





… work in progress …

Can we learn a Lindblad master equation without making assumptions?

General problem:
The size of quantum systems scales **exponentially** with size!

N quantum levels -----> matrix $2^N \times 2^N$

e.g. already 256 (complex) parameters with 4 levels
This problem can only be solved in a scalable way with a quantum computer!

To conclude:

(1) Self-optimising experiments
Considerable speed-up for long measurements (nano-MRI for NVs)

(2) “Machine learning” can help physics
We can use most sophisticated tools to help us do physics
(e.g. are there any alternative explanations for our data?)

(3) **Capitalising on new instruments**
“Programmable” AWGs (Zurich Instruments, Quantum Machines), or directly FPGAs (for the brave)

(4) **Other platforms?**
This is obviously not restricted to NVs

nature reviews physics
https://doi.org/10.1038/s42254-022-00552-1

Review article
Check for updates

Learning quantum systems

Valentin Gebhart[1,2], Raffaele Santagati[3], Antonio Andrea Gentile[4], Erik M. Gauger[5], David Craig[6], Natalia Ares[7], Leonardo Banchi[8,9], Florian Marquardt[10,11], Luca Pezzè[1,2] & Cristian Bonato[5]

Thanks to the team and the sponsors

qpl.eps.hw.ac.uk
c.bonato@hw.ac.uk

Muhammad Junaid Arshad (PDRA)

Ben Haylock (PDRA)

Pasquale Cilibrizzi (PDRA)

Stewart Wallace (PhD)

Issam Belgacem (PhD)

Nick Werren (PDRA)

Malte Kroj (intern)

Quantum (Theory)

Erik Gauger

Hannah Scott (PhD)

Applied statistics

Yoann Altmann

UKRI Engineering and Physical Sciences Research Council
LEVERHULME TRUST
WEIZMANN UK — SUPPORTING THE WEIZMANN INSTITUTE OF SCIENCE — SCIENCE FOR THE BENEFIT OF HUMANITY
[dstl] The Science Inside
QUANTIC — The UK Quantum Technology Hub in Quantum Enhanced Imaging