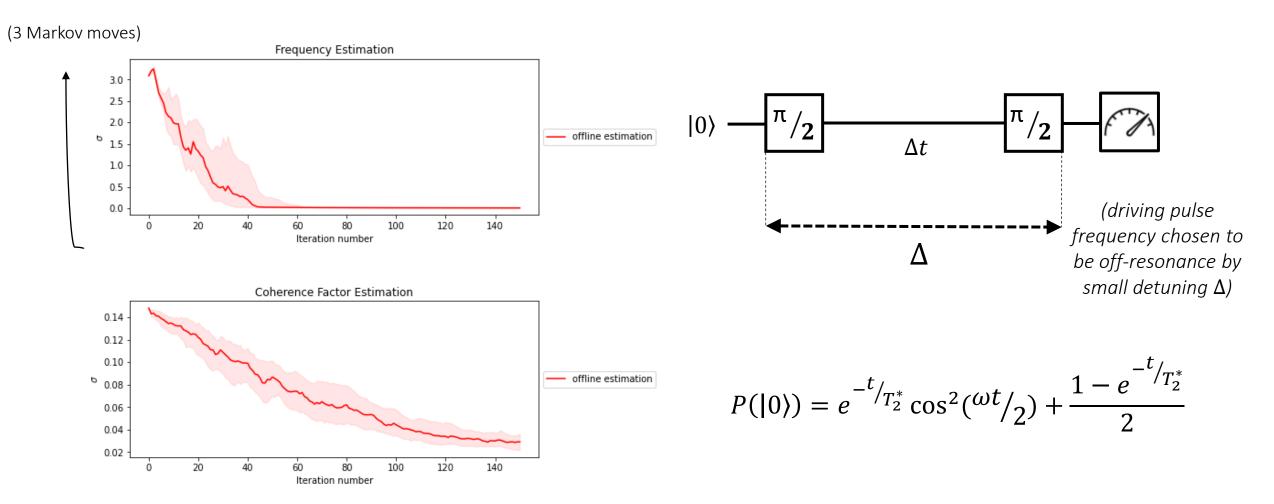
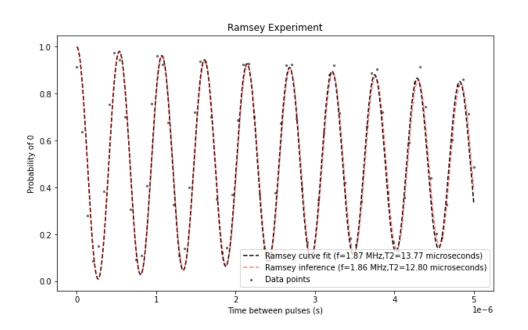
Ramsey experiment on IBMQ device 'Guadalupe'

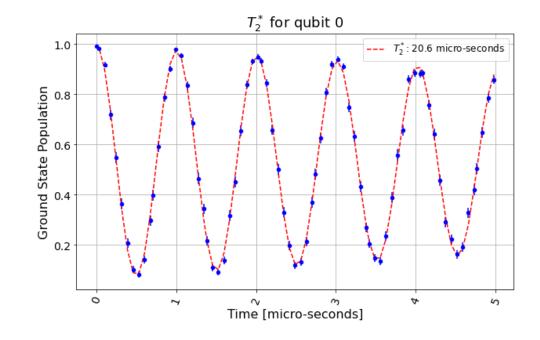


(All graphs use sequential importance resampling SMC with **15² particles** for 2d inference or **50 particles** for 1d, 1 datum added per step, gaussian **random walk Metropolis** for resampling, and a single Markov move per particle per resampler call unless otherwise stated, and take the median results over 100 runs uniformly distributed by 10 datasets collected from IBMQ's backends using OpenPulse.)

Ramsey experiment on IBMQ device 'Guadalupe'

- inference vs. curve fitting





Curve fit results, **512 shots** (Scipy):

$$\Delta f = 1.866 \pm 0.002 \, MHz$$
 $T_2^* = 14 \pm 3 \, \mu s$

Inference results, 2 shots (so 150 steps):

$$\Delta f = 1.863 \pm 0.006 \, MHz$$
 $T_2^* = 15 \pm 6 \, \mu s$

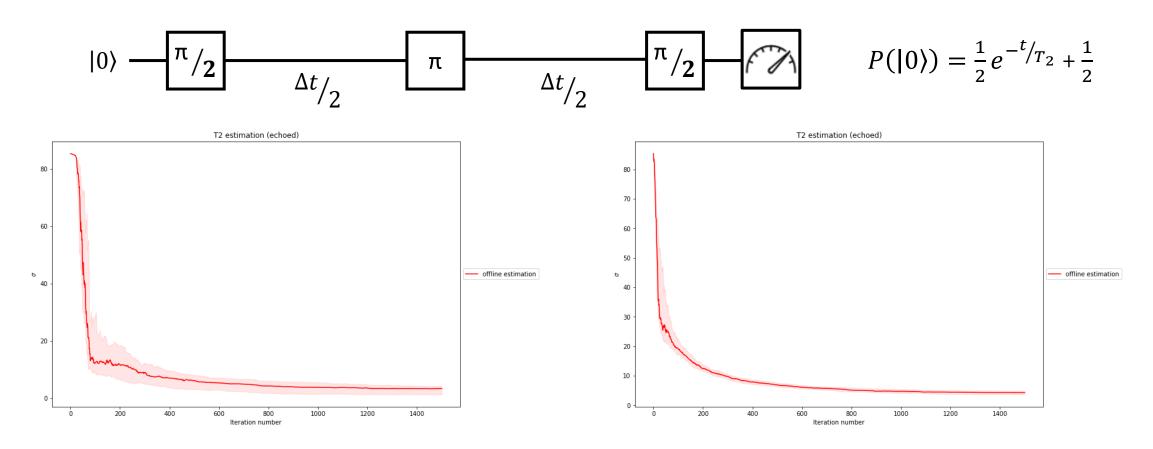
(3 Markov moves)

$$T_2^* = 21 \pm 1 \,\mu s$$

 $\Delta f \approx 1.83$ (apart from errors in the backend resonance frequency estimate) Same set of measurements for both; 75 evenly spaced times for t \in [0.2,5] μ s

$$nosc = 5$$
, $\Delta f = 0.99 MHz$

Hahn echo experiment on IBMQ device 'Rome'

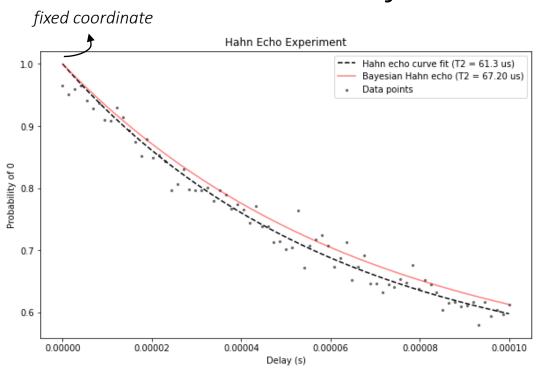


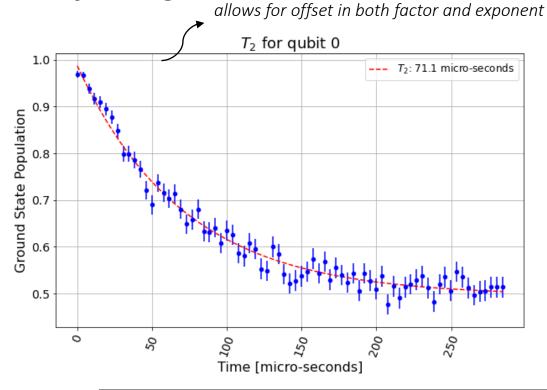
Benefits from measurements being fed to the sampler by reversed order of time evolution (R), oppositely to by increasing order (L) as in Ramsey-type experiments, probably due to having steeper slopes close to origin. This order yields more reliable estimation, less resampler calls, and a higher acceptance rate

Different learning pace from sinusoidal likelihood \rightarrow requires more data (here 1500), but less resampling steps (here 1%)

Hahn echo experiment on IBMQ device 'Rome'

- inference vs. curve fitting





Curve fit results, **512 shots** (Scipy):

$$T_2 = 61 \pm 45 \,\mu s$$

Inference results, **20 shots** (so 1500 steps):

$$T_2 = 67 \pm 4 \,\mu s$$

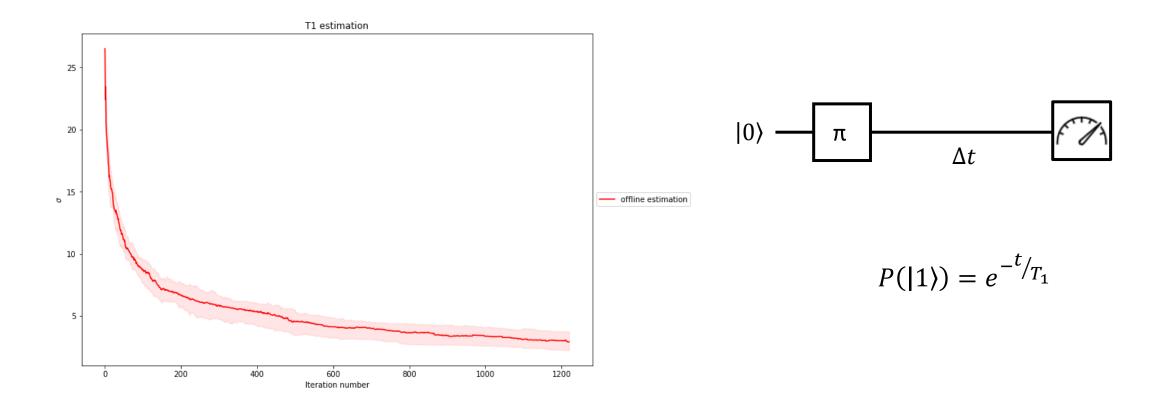
Same set of 75 evenly spaced measurement times for both $(t \in [0.2,100[~\mu s)$

Curve fit results, **512 shots** (Qiskit fitter): $T_2 = 71 \pm 3 \,\mu s$

(Scipy curve fitter is also used internally for the estimate, but the uncertainty is instead calculated from the variance of the success probability which seems to provide a more reasonable estimate)

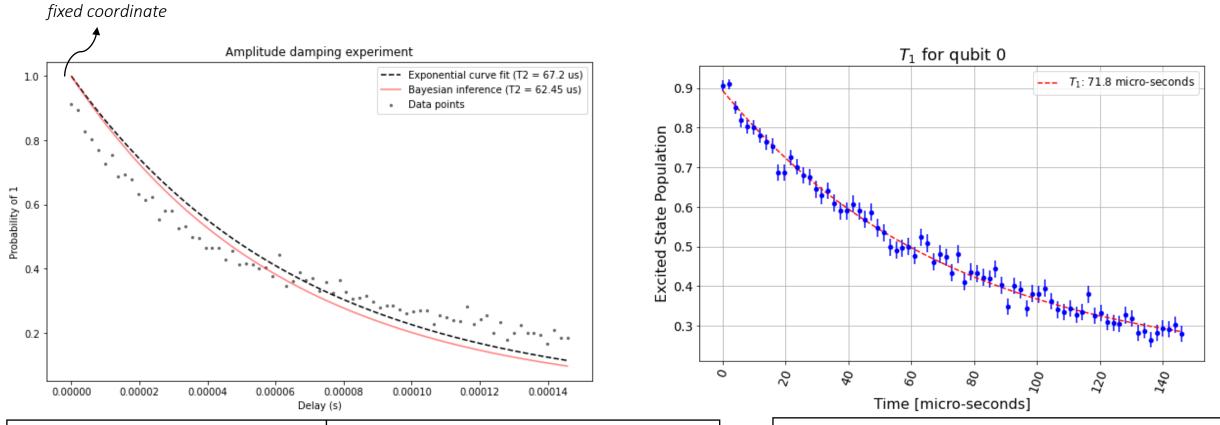
– Estimated backend $T_2 = 77 \,\mu s$ (not too reliable between calibrations)

T_1 experiment (amplitude damping channel) on IBMQ device 'Guadalupe'



Similar to T_1 (same class of models)

T_1 experiment (amplitude damping channel) on IBMQ device 'Guadalupe' - inference vs. curve fitting



Curve fit results, 512 shots (Scipy):

 $T_1 = 67 \pm 26 \,\mu s$

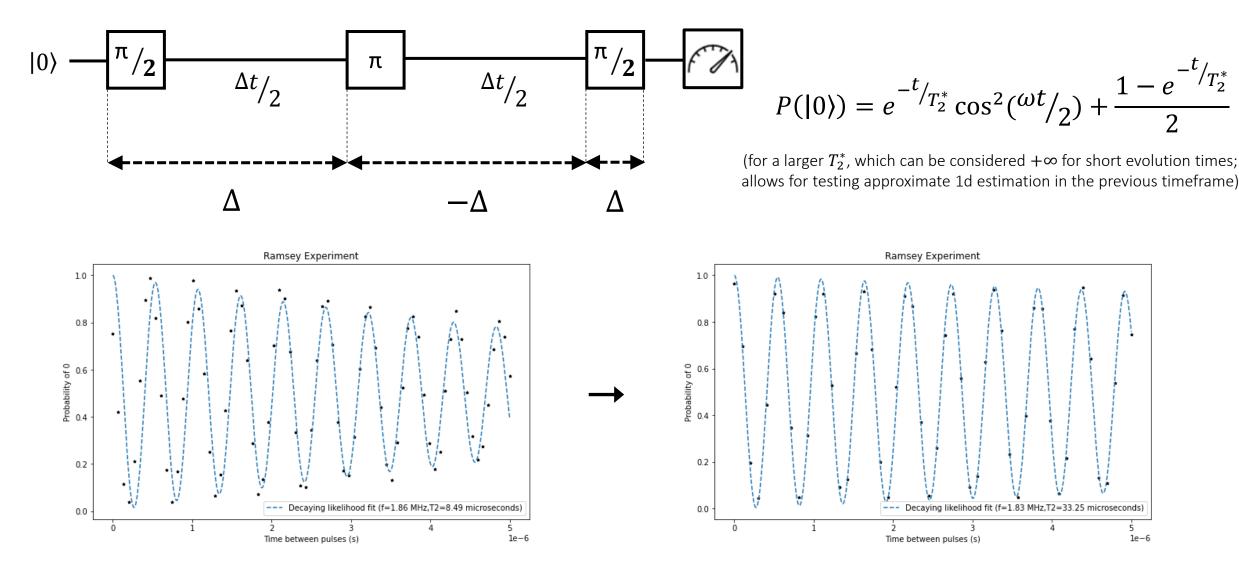
Inference results, 20 shots (so 1500 steps):

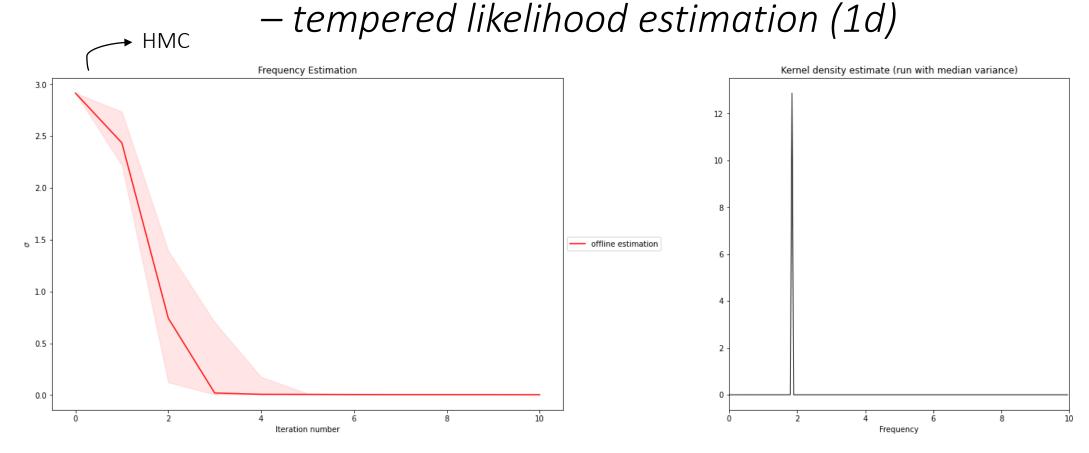
$$T_1 = 62 \pm 3 \,\mu s$$

Curve fit results, 512 shots (Qiskit fitter):

$$T_1 = 72 \pm 4 \,\mu s$$

- Estimated backend $T_1 = 49 \, \mu s$ (not too reliable between calibrations)

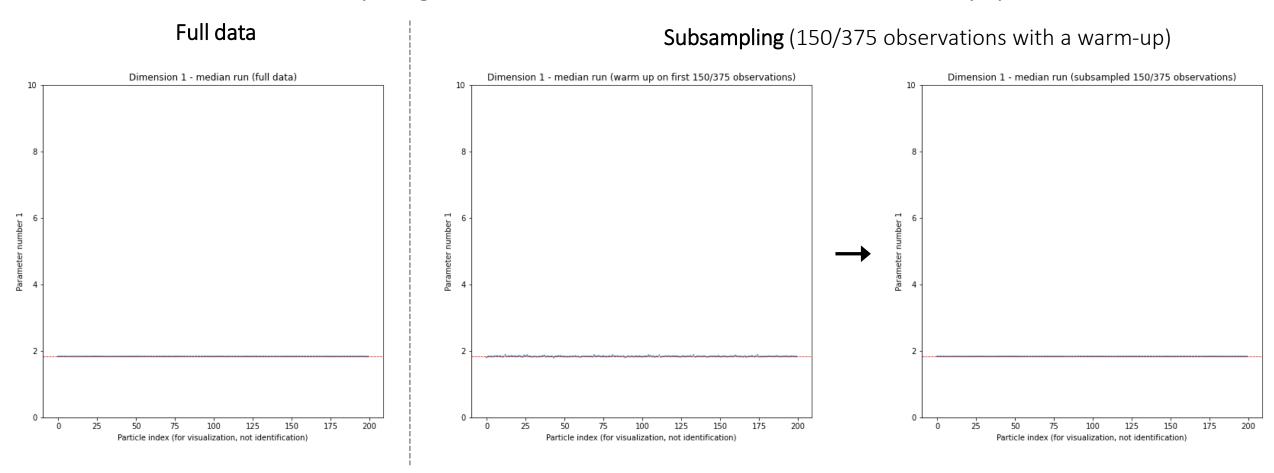




Parallel Markov chains are inefficient for long evolution times and/or wide priors (too steep likelihood function causes trajectory divergence for HMC, low acceptance for RWM)

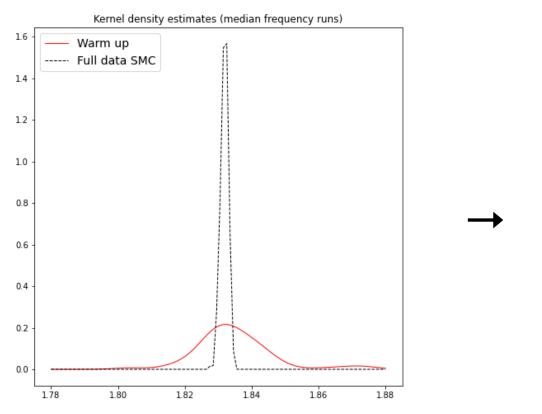
If taking a non-sequential approach, coupling particles and tempering the likelihoods works better and allows for scale-independent calculations of the Bayes' factors, while being more suitable to assymptotically correct subsampling schemes

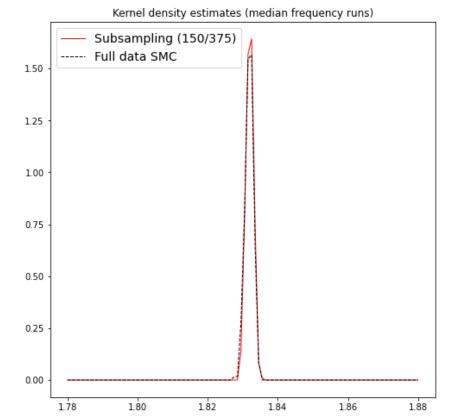
- subsampling with control variates and a warm-up phase



Gradient-based control variates (Taylor expansions) don't work well due to high target curvature, which can be controlled by either keeping the evolution times short (inefficient because they bring little information after a point) or narrowing the prior. The latter can be done by a warm-up phase on the first (lower times) observations if prior knowledge is insufficient to justify it (the prior may be flattened before proceeding if the marginal likelihoods are required)

- subsampling with control variates and a warm-up phase





 $\Delta f \approx 1.83 \, MHz$

200 particles; prior on f ∈ [0,10[

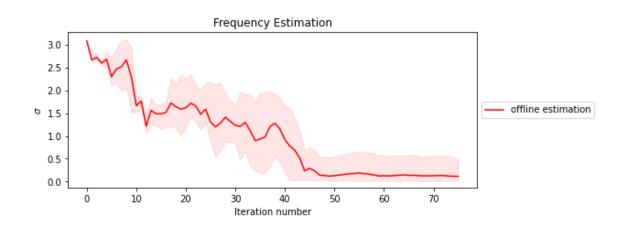
Median results: (distribution means and standard deviations)

- Full data: $\Delta f = 1.83184 \pm 0.0012 \, MHz$
- Subsampling: $\Delta f = 1.83197 \pm 0.0010 \, MHz$ (Warm up: $\Delta f = 1.8348 \pm 0.013 \, MHz$)

Median deviations: (when subsampling, reference is full data)

- Frequency: +0.0001615(+0.0%)
- Standard deviation: -0.0002680 (-22.2%)

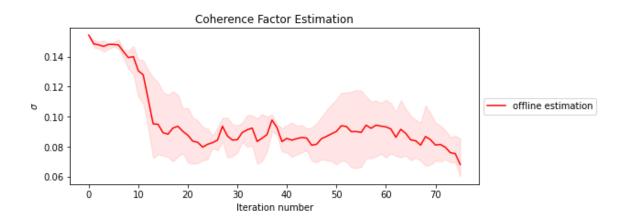
- inference with Liouvillian integration based likelihoods



$$\frac{d}{dt}\rho(t) = -i\left[H, \rho\right] + \Gamma\left(\sigma^{-}\rho\sigma^{+} - \frac{1}{2}\left\{\sigma^{+}\sigma^{-}, \rho\right\}\right)$$

$$\mathcal{L} = \begin{pmatrix} 0 & i\Omega & -i\Omega & \Gamma \\ i\Omega & -iE - \frac{\Gamma}{2} & 0 & -i\Omega \\ -i\Omega & 0 & -iE - \frac{\Gamma}{2} & i\Omega \\ 0 & -i\Omega & i\Omega & -\Gamma \end{pmatrix}$$

$$\frac{d}{dt}\rho\left(t\right) = \mathcal{L}\rho(t)$$

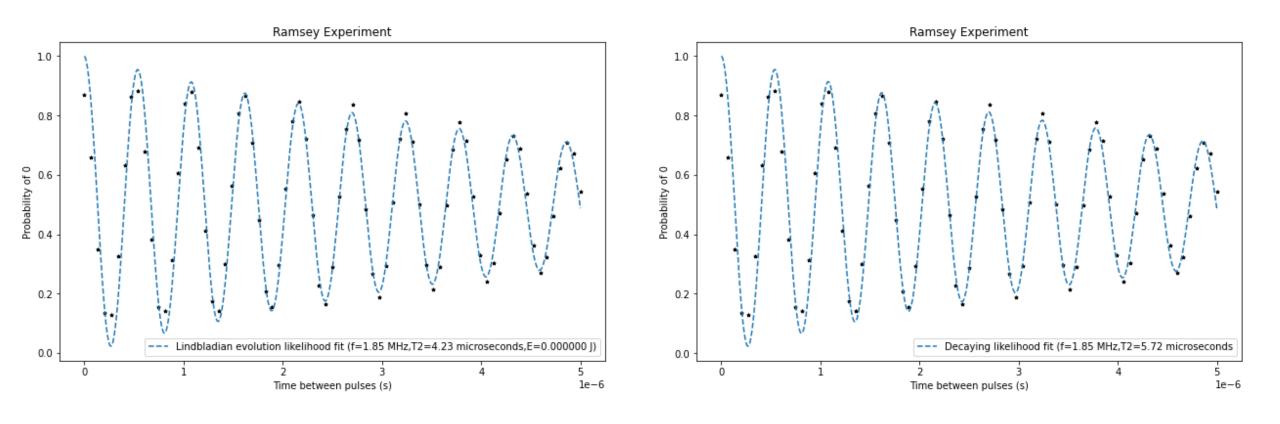


$$\Delta f = 1.85 \pm 0.1 \, MHz$$

 $T_2^* = 3.3 \pm 0.8 \, \mu s$

«Statistics» over 3 runs and with all but latest 30 data dropped when resampling (integration takes very long)

Ramsey experiment on IBMQ device 'Armonk' - curve fitting with Liouvillian integration based likelihoods



Master equation time evolution (L) vs. the damped oscillator model (R) given the same data (black markers)

Integration methods for simulating the time evolution

Integrator results for P(1):

Runge-Kutta 2(3): 0.8495685717649359+0j Runge-Kutta 4(5): 0.8578922098165880+0j Runge-Kutta 8: 0.8565401440306653+0j Backward diff: 0.8567935084404129+0j

High precision RK8*: 0.8571408221863202+0j

Time taken (same tolerances except for last):

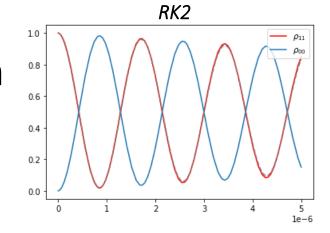
Runge-Kutta 2(3): 13.9 ms

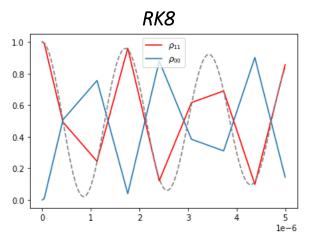
Runge-Kutta 4(5): 5.0 ms

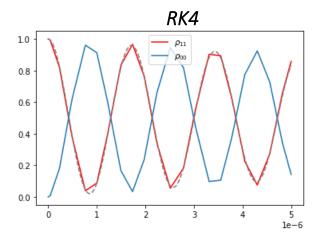
Runge-Kutta 8: 6.8 ms

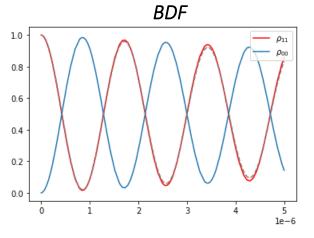
Backward diff: 36.3 ms

- High precision RK8: 20.5 ms
 - Plotting only the computed points
 - Dashed line gives ρ_{11} for the damped oscillator model
 - E ≈ 0

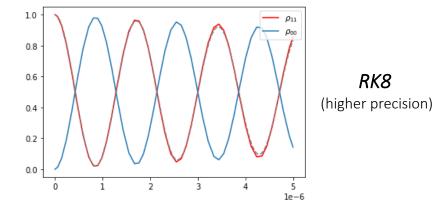








RK8



^{*}Reference