Quantum System Identification

Hamiltonian Estimation using Spectral and Bayesian Analysis

Sophie Schirmer (Cambridge, UK)

joint work with

Daniel Oi (Strathclyde, UK)

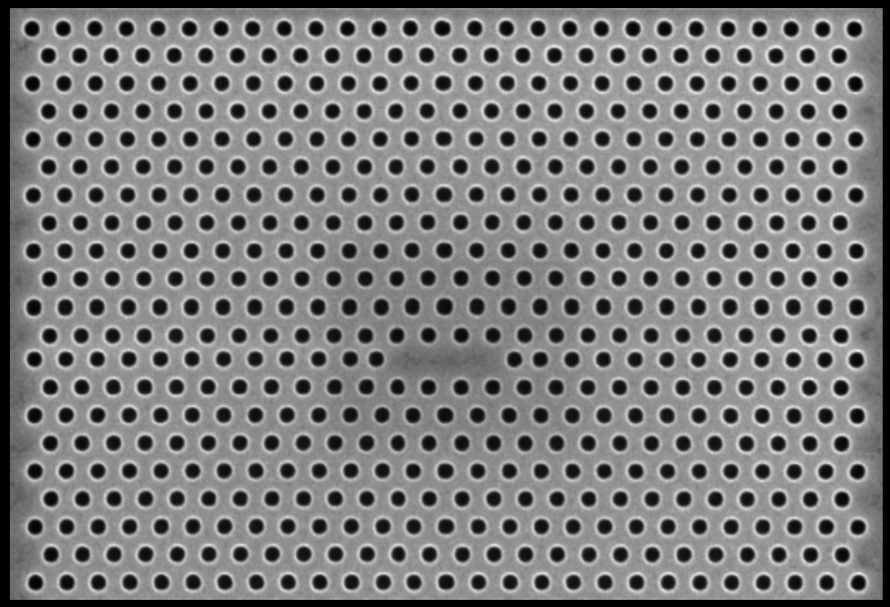
Frank Langbein (Cardiff, UK)

Jared Cole, Simon Devitt, Lloyd Hollenberg (U Melb., Australia)

Weiwei Zhou, Erling Gong, Ming Zhang (NUDT, China)

Quantum Engineering

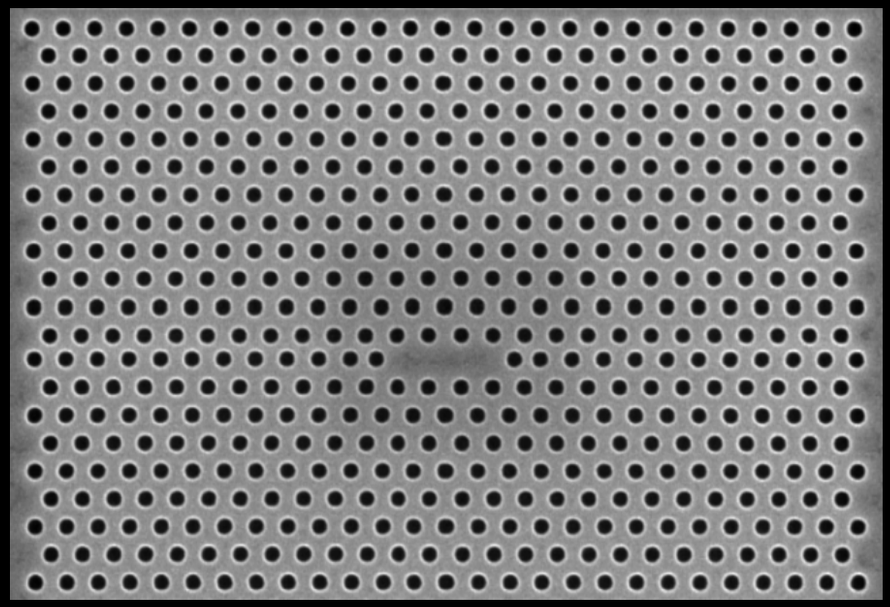
- Recent progress in areas such as
  - nano-fabrication
  - photonics
  - laser technology
  - ion-trapping
  - atom chips
  - Bose-Einstein condensation



  - ➔ Enables wide range of applications for quantum phenomena
- Quantum effects such as *tunnelling*, *coherence*, *interference* and *entanglement* hold the promise novel technological applications

Quantum Engineering

- Recent progress in areas such as
  - nano-fabrication
  - photonics
  - laser technology
  - ion-trapping
  - atom chips
  - Bose-Einstein condensation



  - ➔ Enables wide range of applications for quantum phenomena
- Quantum effects such as *tunnelling*, *coherence*, *interference* and *entanglement* hold the promise novel technological applications

- Quantum engineering requires *control of quantum effects*
  - Effective control requires *accurate models* of the system
  - ➔ Need for identification of quantum systems

Quantum System Identification

- System identification for quantum systems often interpreted as
  - quantum state tomography (or estimation)
  - quantum process tomography
- But for control a *dynamical model* of the system is required

Quantum System Identification

- System identification for quantum systems often interpreted as
  - quantum state tomography (or estimation)
  - quantum process tomography
  - But for control a *dynamical model* of the system is required

- System must satisfy evolution equation
  - Such as the *quantum Liouville equation*

$$\dot{\rho} = -i[H, \rho] + L_D(\rho)$$

  - Need to identify dynamical operators
    - H *(Hamiltonian)*
    - L_D *(dissipation)*
  - If is Hamiltonian *control dependent* ($H = H[\mathbf{u}]$), need to identify this dependency as well

Quantum System Identification

- System identification for quantum systems often interpreted as
  - quantum state tomography (or estimation)
  - quantum process tomography
  - But for control a *dynamical model* of the system is required

- System must satisfy evolution equation
  - Such as the *quantum Liouville equation*

$$\dot{\rho} = -i[H, \rho] + L_D(\rho)$$

  - Need to identify dynamical operators
    - H *(Hamiltonian)*
    - L_D *(dissipation)*
  - If is Hamiltonian *control dependent* ($H = H[\mathbf{u}]$), need to identify this dependency as well

Protocols for Hamiltonian Estimation

Protocol 1: Continuous weak measurements

- *Problems*: measurement-backaction and accuracy

Protocols for Hamiltonian Estimation

Protocol 1: Continuous weak measurements

- *Problems*: measurement-backaction and accuracy

Protocol 2: Projective measurements at discrete times

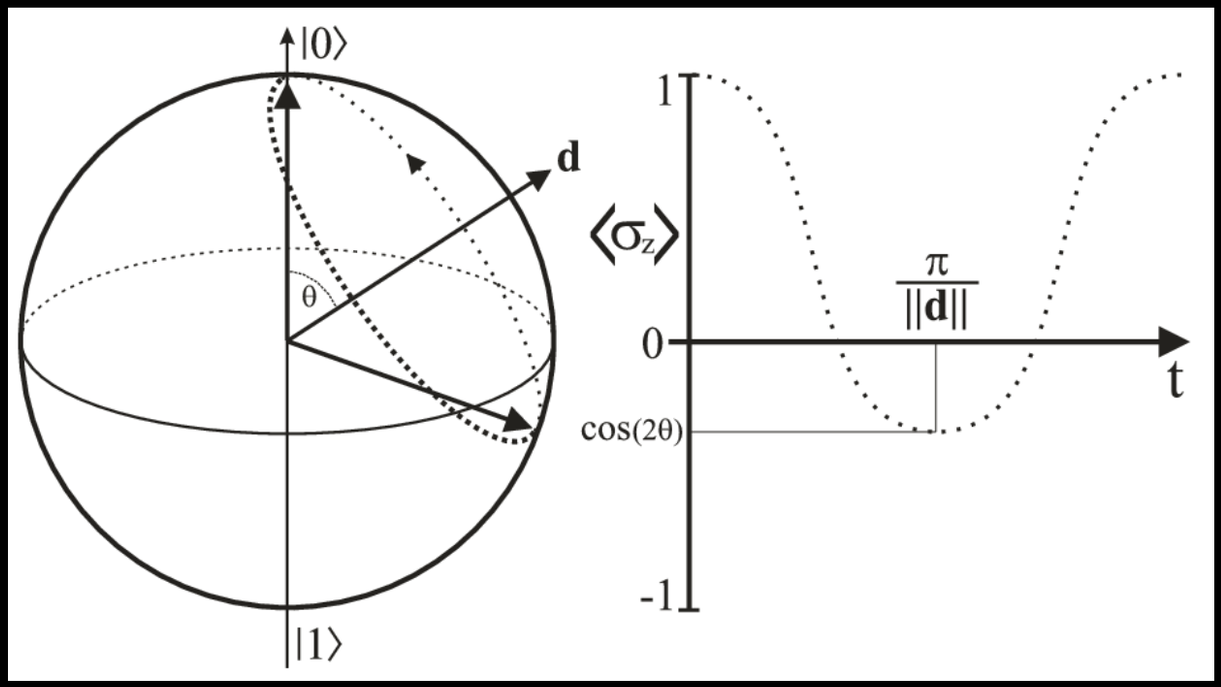
- Initialize system in a known state ρ_0
- Let system evolve for time t under (unknown) Hamiltonian
- Perform projective measurement Π
  - Output state $\Pi[\rho(t)]$
- Repeat many times for given t, then vary t

- *Challenge*

  As outcome of a single experiment is random, many repetitions of individual experiment necessary to build up statistics, but total number of experiments should be minimized

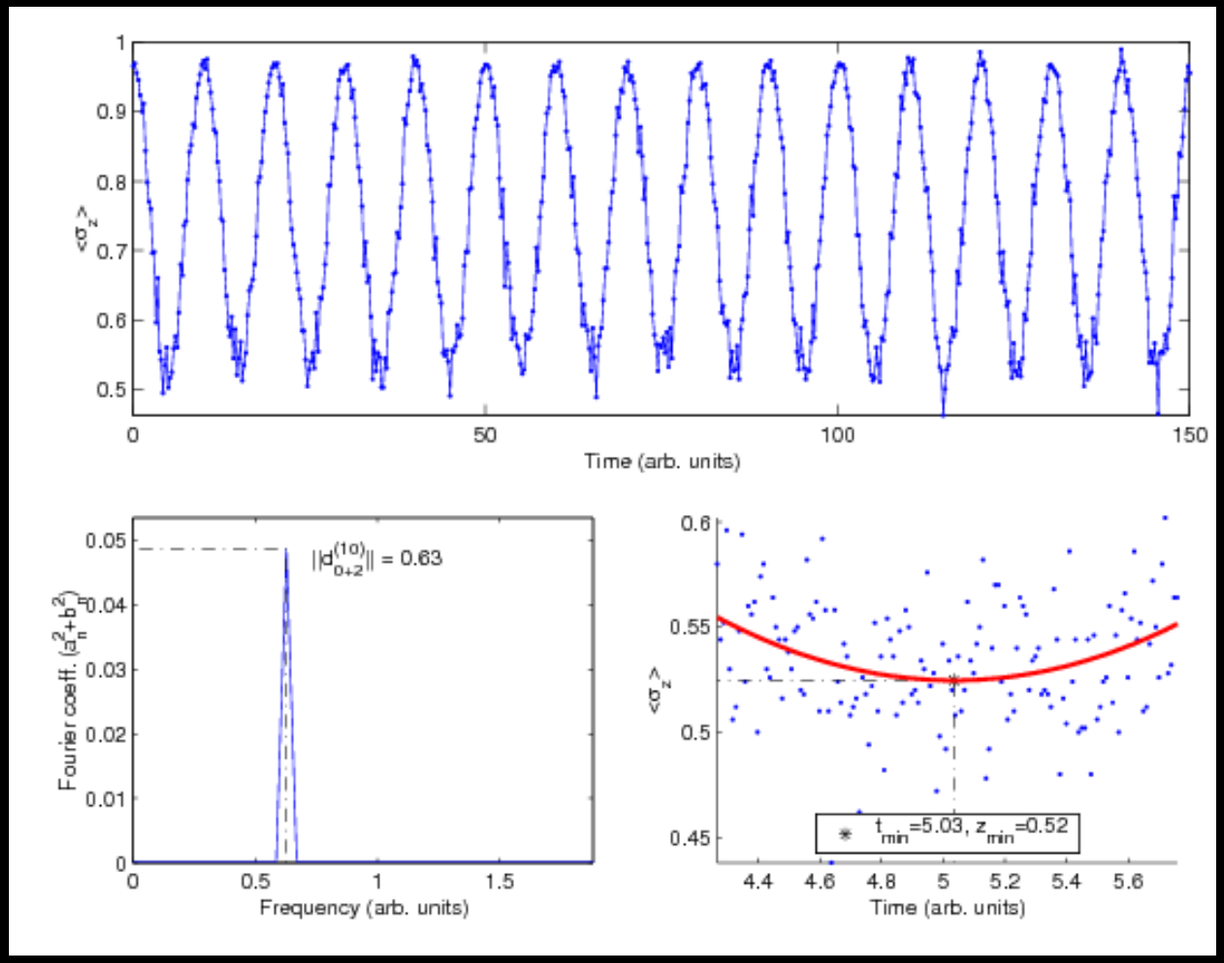
Qubit Hamiltonian Tomography

	Hilbert space picture	Bloch sphere picture
State	density op $\rho \in \mathfrak{D}(\mathbb{H})$	Bloch vector $\mathbf{s} \in \mathbb{R}^3$
	$\rho = \frac{1}{2}(I + \mathbf{s} \cdot \sigma)$	$s_x = \mathrm{Tr}(\rho\sigma_x)$, etc
	$\sigma = (\sigma_x, \sigma_y, \sigma_z)$	$\sigma_x, \sigma_y, \sigma_z$ Pauli matrices
	$\mathrm{Tr}(\rho^2) \leq 1$	$\|\mathbf{s}\|^2 = 2\,\mathrm{Tr}(\rho^2) - 1 \leq 1$
Evolution	$\frac{d}{dt}\rho(t) = -i[H, \rho(t)]$	$\frac{d}{dt}\mathbf{s}(t) = R(\mathbf{d})\mathbf{s}(t)$
	$H = \frac{1}{2}(d_0 I + \mathbf{d} \cdot \sigma)$	$R(\mathbf{d})$ rotation about $\mathbf{d}$



Fixed Hamiltonian identification

- Find Rotation Frequencies and Declination of Rotation Axes
- Rotation of the Bloch vector $\mathbf{s}(t)$ with $\mathbf{s}(0) = (0, 0, 1)^T$ about axis $\mathbf{d}$
- Observable oscillations $\langle \sigma_z(\alpha_t) \rangle = z(\alpha_t) = \cos(\alpha_t) \sin^2\theta + \cos^2\theta$



Rotation frequency $\|\mathbf{d}\|$ and declination angle θ determined by

- position and height of Fourier peaks, $\cos\theta = \sqrt{F(0)}$
- minimum (t_{min}, z_{min}) in the oscillation data
- maximum likelihood (best strategy)

Multiple Hamiltonians

- Single Hamiltonian: suffices to determine Rabi frequency angle θ
- Multiple rotations about different axes $\mathbf{d}_k \Rightarrow$ must determine the relative azimuthal angles $\phi_k = \phi'_k - \phi'_r$ w.r.t. a reference axis $\mathbf{d}_r$
- Finding azimuthal angles between rotation axes:
  - Having determined $||\mathbf{d}||$ and θ of the axes:
    Choose reference axis $\mathbf{d}_r$ (some conditions apply) with $\phi_r = 0$
  - Repeat all Rabi oscillation experiments with new initial state $s_1 = (\cos\beta, \sin\beta, 0)^T$ obtained by rotating s_0 about reference axis by angle by suitable angle α_r.

$$\alpha_r = \arccos\left(\frac{\cos(2\theta_r)+1}{\cos(2\theta_r)-1}\right), \quad \beta = \arctan(-\sqrt{|2\cos(2\theta_r)|}\sec\theta_r)$$

  - Spectral or Bayesian analysis of Rabi data $\Rightarrow$ horizontal angles

$$z(\alpha) = C(1-\cos\alpha) + D\sin\alpha$$

$$C = 1/2\sin(2\theta_k)\cos(\phi_k - \beta), \qquad D = \sin\theta_k \sin(\phi_k - \beta).$$

General Strategy

(1) Select set of fixed values for controls

$$\mathbf{f}^{(\ell)} = (f_m^{(\ell)}), \quad m = 1, 2, \ldots, M, \quad \ell = 1, 2, \ldots, L.$$

(2) Find rotation frequencies $||\mathbf{d}_{0+m}^{(\ell)}||$ and declination angles $\theta_{0+m}^{(\ell)}$ of the rotation axes

$$\mathbf{d}_{0+m}^{(\ell)} = \mathbf{d}_0 + f_m^{(\ell)} \mathbf{d}_m, \qquad \forall \ell, m$$

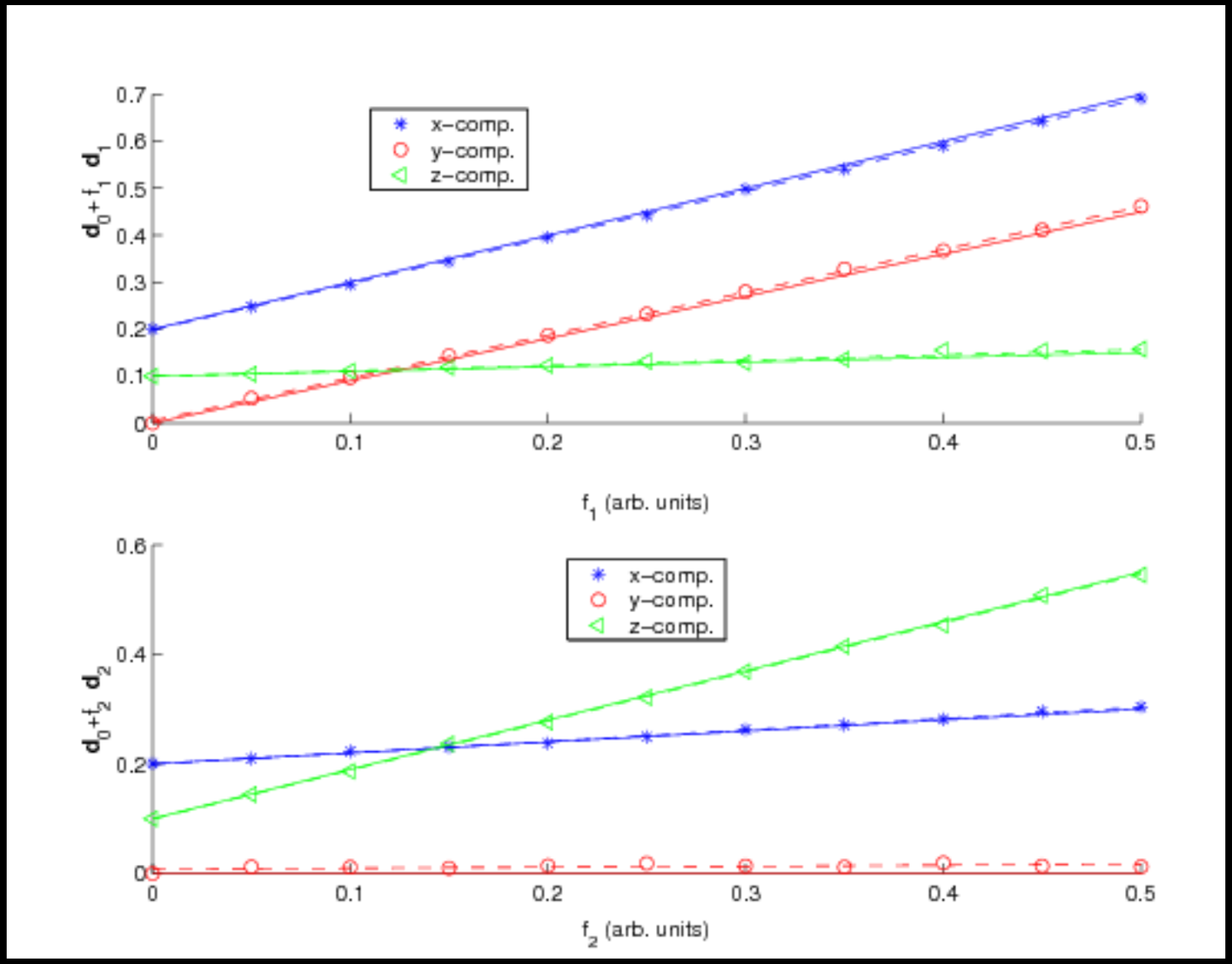
(3) Find horizontal angles $\phi_{0+m}^{(\ell)}$ between rotation axes

(4) Identify control dependence, e.g., by plotting $\mathbf{d}_{0+m}^{(\ell)}$ versus $f_m^{(\ell)}$

- If plot suggests linear dependence find best straight line fit using linear regression:
  - y-intercepts determine d_{0x}, d_{0y} and d_{0z}
  - slopes determine d_{mx}, d_{my} and d_{mz}
- Otherwise find non-linear fit

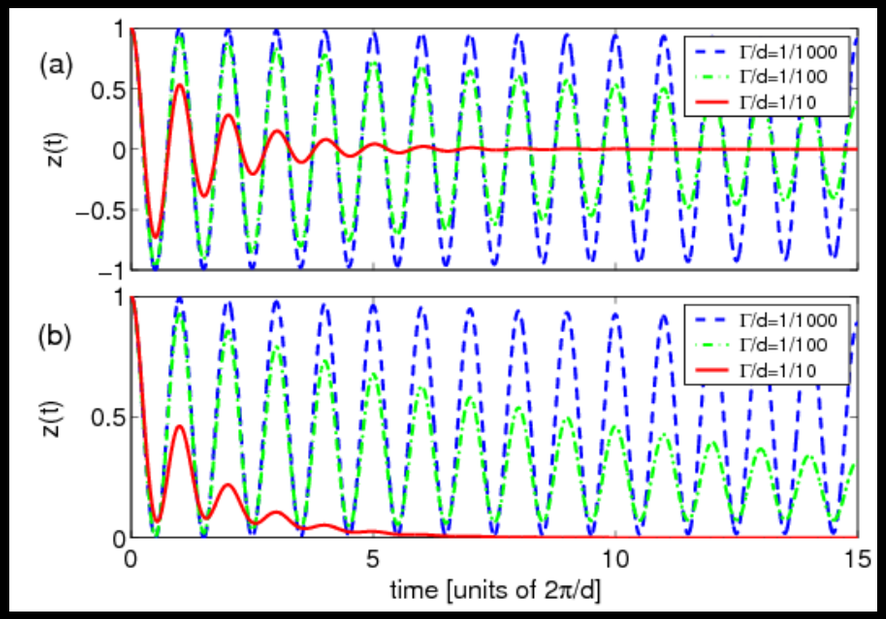
Phys Rev A 69, 050306 (2004)

Example for Linear Control Dependence



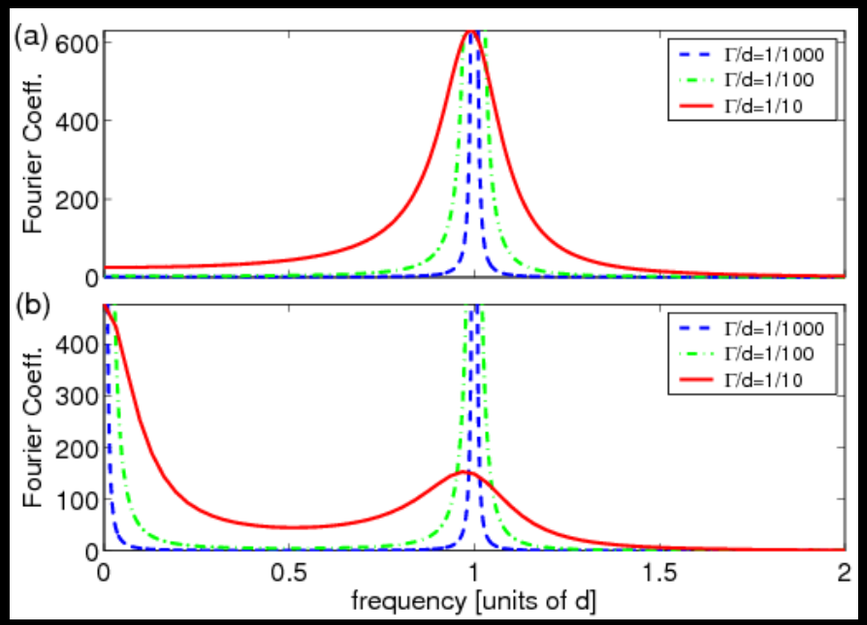
Dissipation Characterization

- Dissipation $\Rightarrow$ Damping of observed coherent oscillations



$$z(t) = a + be^{-\gamma t}\sin(\omega_0 t)$$

Lorentzian broadening of δ-like peaks in Fourier spectrum



$$|F(\omega)| = \frac{\omega_0}{[\gamma^2+(\omega_0-\omega)^2][\gamma^2+(\omega+\omega_0)^2]}$$

- $|F(\omega)|$ — maximum at $\omega = \sqrt{\omega_0^2 - \gamma^2}$ with peak value $(2\gamma)^{-1}$
- Can estimate frequency ω_0 and dephasing rate γ from position ω_* peak height $|F|_*$ of peak: $\gamma = (2|F|_*)^{-1}$ and $\omega_0 = \sqrt{\omega_*^2 + \gamma^2}$.

Dissipation Characterization

- But peak height identification inaccurate.
- Better estimate for γ using width of the peak. Let $\omega_{1,2}$ be the (positive) frequencies for which $|F(\omega)|$ assumes half its maximum or $1/(4\gamma)$. Then the full-width-half-maximum $2d$ of $|F(\omega)|$ is $|\omega_2 - \omega_1|$ or

$$d = \left[\sqrt{\omega_*^2 + 2\sqrt{3}\gamma\sqrt{\omega_*^2 + \gamma^2}} - \omega_*\right]$$

Given ω_* and half-width d of the peak we can solve for γ

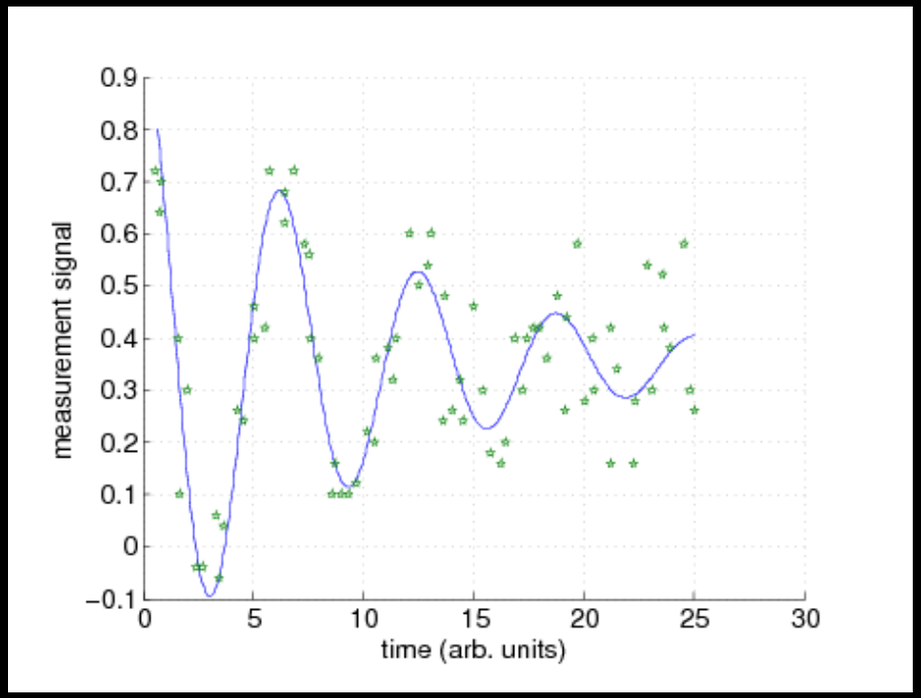
$$\gamma = \frac{1}{6}\sqrt{6g(\omega_*, d) - 18\omega_*^2}$$

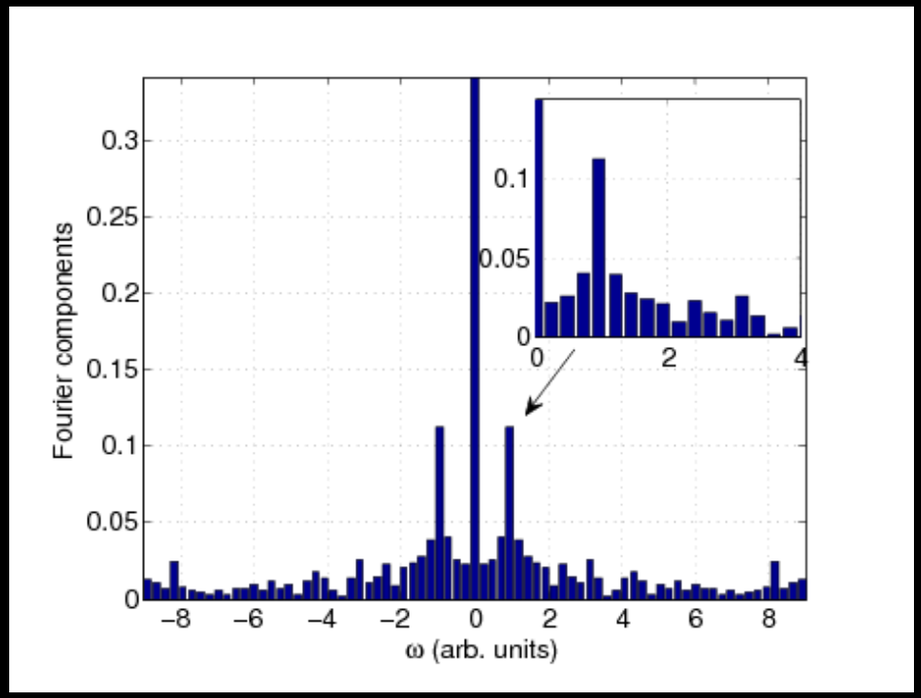
where $g(\omega_*, d) = \sqrt{9\omega_*^4 + 12d^2\omega_*^2 + 12d^3\omega_* + 3d^4}$. Thus, in principle we can determine both the frequency and the dephasing rate by estimating the position and width of the fourier peak.

Phys. Rev. A 71 062312 (2005); Phys. Rev. A 73, 062333 (2006)

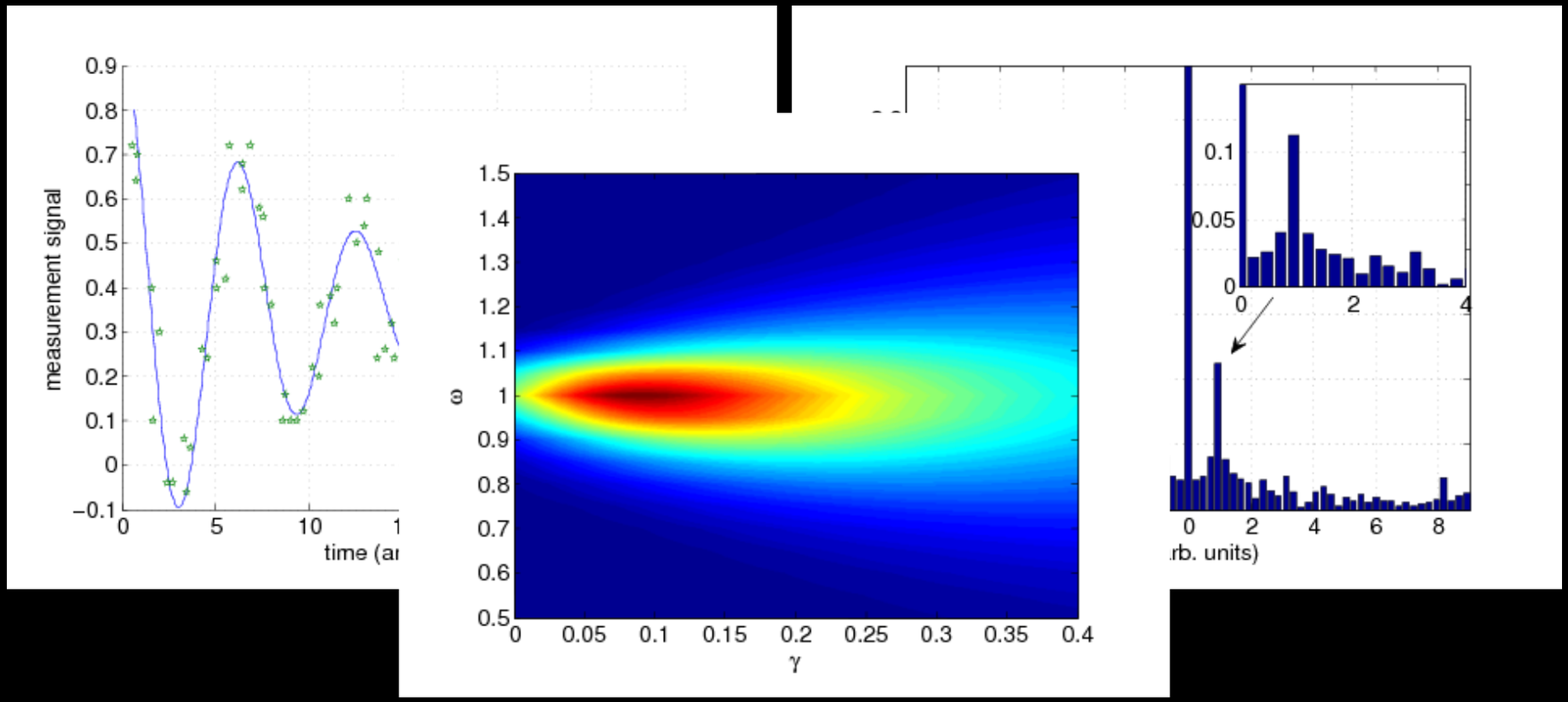
- Even better estimates using Bayesian Estimation!

Open System Identification — Bayesian





Open System Identification — Bayesian

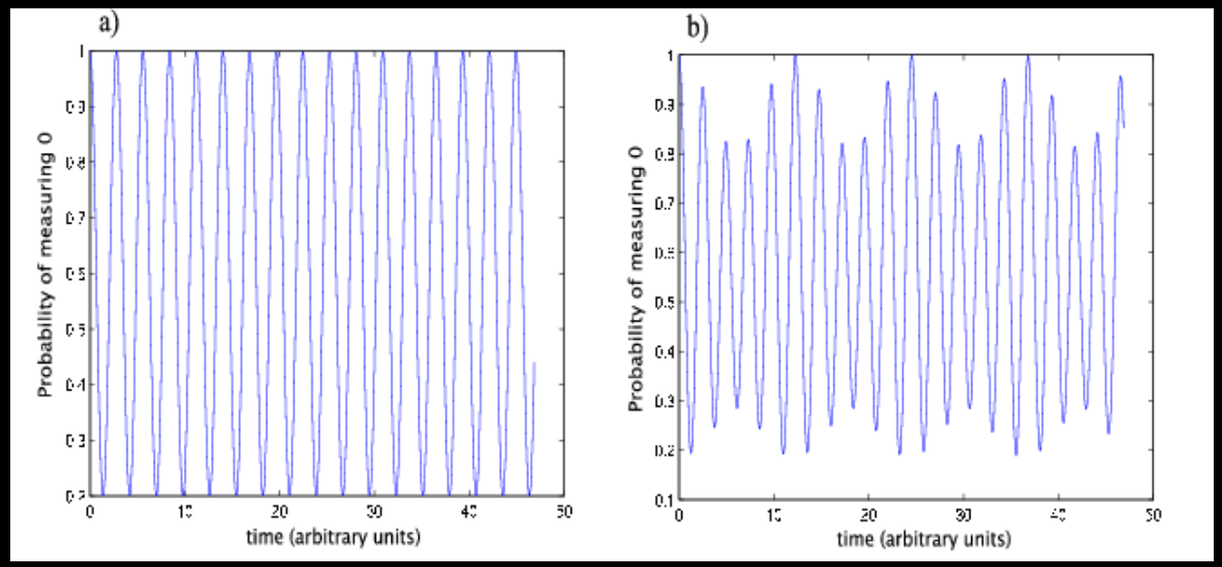


- Noisy sparse signal ⇒ spectral estimation inaccurate
- log-Likelihood function still has sharp peak
- Bayesian estimation possible even if initial state and measurement basis uncertain, provides estimates for both!

arXiv:1012.4593

Characterizing Subspace Confinement

- Qubit characterization assumed dynamics confined to 2D subspace
- Reality: many degrees of freedom — dynamics may not be confined
- Experimental characterization of subspace leakage possible
- Two scenarios: Measurement discriminates
  - $|0\rangle$, $|1\rangle$ and "other" $\Rightarrow$ confinement characterization easy
  - Only $|0\rangle$ and "other" $\Rightarrow$ special protocols needed
    Detect modulations in Rabi oscillation data NJP 9, 384 (2007)



Visual detection sometimes possible but not reliable

Assessing Subspace Leakage

- Cumulative amplitudes of non-qubit states for given eigenstate of H

$$\epsilon = 1 - \mathrm{Tr}[\Pi\rho], \qquad \Pi = |0\rangle\langle 0| + |1\rangle\langle 1|$$

- Can be calculated from peak heights in spectrum. Assume

$$f(t) = |\langle 0|A^\dagger e^{-iH_d t} A|0\rangle|^2 = \sum_{k,\ell} h_{k\ell} e^{i\omega_{k\ell} t}$$

- $h_{k\ell}$ height of Fourier peak for transition $k \to \ell$ with frequency $\omega_{k\ell}$

$$\epsilon = \sum_k \sqrt{\frac{h_{k\ell} h_{km}}{h_{\ell m}}}, \quad b, c \neq a$$

- Exact calculation theoretically possible but generally impractical — signal/noise ratio problem (signal 'lost' in noise floor)
- Can give upper and lower bounds on subspace confinement based only on two main peaks h_0 and h_{01}

$$1 - \sqrt{h_0 + 2h_{01}} \leq \epsilon \leq \frac{1}{2}(1 - \sqrt{2h_0 + 4h_{01} - 1})$$

- Improvement using Bayesian estimation?

Two-qubit Hamiltonian Tomography

- Interaction identification by Concurrence Spectroscopy
- **Aim:** Characterize interaction Hamiltonian H_I for two coupled qubits
  - single qubit control Hamiltonian known, fixed coupling
- **Idea:** Use concurrence (measure of entanglement)

$$C(t) = \rho(t)(\sigma_y \otimes \sigma_y)\rho(t)^*(\sigma_y \otimes \sigma_y)$$

  σ_y usual Pauli matrix, $\rho(t)$ system density operator
- Outline of protocol for experimental characterization of H_I
  - Initialize system in different separable basis states
  - Measure time-evolved states $\Rightarrow \rho(t)$
  - Calculate concurrence $C(t)$ time-series for each input state
  - Compute concurrence spectra
  - Peak frequencies/heights determine interaction Hamiltonian
- Limitations: only non-local part of interaction Hamiltonian

J. Phys A 39, 14649 (2006); PRA 73, 052317 (2006)

General Hamiltonians — Identifiability

- *What information* can we hope to extract about a given system with a certain limited set of resources?

General Hamiltonians — Identifiability

- *What information* can we hope to extract about a given system with a certain limited set of resources?

- E.g., suppose we can
  - only prepare system in a set of computational basis states $|n\rangle$, $n = 1, \ldots, N = \dim \mathcal{H}$
  - measure the system in these basis states $\Pi_n = |n\rangle\langle n|$

- ➔ Resources *not* sufficient for
  - Process tomography, even unitary case
  - Full Hamiltonian identification, even for constant H
- ➤ Nonetheless a significant amount of information about Hamiltonian and even dissipative effects (e.g. decoherence rate) can be obtained
  - Stroboscopically map evolution over time

Limitations on Identifiability

Theorem

- Let H and M be Hermitian operators on $\mathcal{H}$ [Hamiltonian and measurement]
- Let ρ_0 be a positive operator with $\mathrm{Tr}(\rho_0) = 1$ [initial state of the system]
- If M, H, ρ_0 are simultaneously blockdiagonalizable
  - i.e. there exists a decomposition of the Hilbert space $\mathcal{H} = \oplus_{s=1}^{S>1} \mathcal{H}_s$ such that

    $$M = \mathrm{diag}(M_s), \quad H = \mathrm{diag}(H_s), \quad \rho_0 = \mathrm{diag}(\rho_s),$$

    where M_s, H_s, ρ_s are operators on the Hilbert spaces $\mathcal{H}_s$
- Then we can at most identify H up to $\sum_s \lambda_s \mathbb{I}_s$
  - where $\mathbb{I}_s$ is the identity on the subspace $\mathcal{H}_s$

Identifiability and A-priori Information

- Generic case: H and M *not simultaneously block-diagonizable*
  - Can identify H at most up to a diagonal unitary matrix $D = (1, e^{i\phi_2}, \ldots, e^{i\phi_N})$ and a global energy shift $\lambda_0 \mathbb{I}$

$$\tilde{H} \simeq H = D^\dagger \tilde{H} D + \lambda_0 \mathbb{I}$$

Identifiability and A-priori Information

- Generic case: H and M *not simultaneously block-diagonizable*
  - Can identify H at most up to a diagonal unitary matrix $D = (1, e^{i\phi_2}, \ldots, e^{i\phi_N})$ and a global energy shift $\lambda_0 \mathbb{I}$

$$\tilde{H} \simeq H = D^\dagger \tilde{H} D + \lambda_0 \mathbb{I}$$

- *Off-diagonal elements are real and positive* in computational basis

$$H_{k\ell} = \langle k|H|\ell\rangle = |\langle k|H|\ell\rangle|$$

  - Then $$|\langle k|H|\ell\rangle| = |\langle k|\tilde{H}|\ell\rangle| \qquad (\tilde{H} \text{ as above})$$

➜ With this constraint Hamiltonian is effectively uniquely determined
(up to a global energy level shift and global inversion of the energy levels)

Identifiability and A-priori Information

- Generic case: H and M *not simultaneously block-diagonizable*
  - Can identify H at most up to a diagonal unitary matrix $D = (1, e^{i\phi_2}, \ldots, e^{i\phi_N})$ and a global energy shift $\lambda_0 \mathbb{I}$

$$\tilde{H} \simeq H = D^\dagger \tilde{H} D + \lambda_0 \mathbb{I}$$

- *Off-diagonal elements are real and positive* in computational basis

$$H_{k\ell} = \langle k|H|\ell\rangle = |\langle k|H|\ell\rangle|$$

  - Then $$|\langle k|H|\ell\rangle| = |\langle k|\tilde{H}|\ell\rangle|$$ ($\tilde{H}$ as above)

➡ With this constraint Hamiltonian is effectively uniquely determined
(up to a global energy level shift and global inversion of the energy levels)

- Example: Hamiltonian for spin-networks, many atomic systems

Full Hamiltonian Tomography

- Full control Hamiltonian tomography for generic N-level systems
  - assume N-outcome measurement in fixed basis
  - initialization in measurement basis states possible
- Stroboscopically map coherent oscillations for different controls
  - Initialize qubits via simulataneous single-qubit measurements
  - Evolve under (fixed) control Hamiltonian H_c for time t
  - Measure both qubits (simple measurement)
  - ➔ N^2 coherent osciallation traces $p_{k\ell}(t)$
- Spectral and Bayesian analysis of coherent oscillation data
- Level-structure identification
- Hamiltonian reconstruction
- Generically, complete identification possible if system controllable

PRA 80, 022333 (2009); Laser Physics 20, 1203 (2010)

Simpler protocols using a-priori info

- Availablity of further a-priori information may significantly reduce resource requirements
- Example: $H_{\text{eff}} = \begin{bmatrix} 0 & \Omega_1 & 0 \\ \Omega_1 & 0 & \Omega_2 \\ 0 & \Omega_2 & 0 \end{bmatrix}$ with $\Omega_{1,2} \in \mathbb{R}^+$ gives

$$U(t,0) = \exp(-itH_{\text{eff}}) = \begin{bmatrix} c^2\cos(\Omega t) + s^2 & -ic\sin(\Omega t) & cs[\cos(\Omega t) - 1] \\ -ic\sin(\Omega t) & \cos(\Omega t) & -is\sin(\Omega t) \\ cs[\cos(\Omega t) - 1] & -is\sin(\Omega t) & s^2\cos(\Omega t) + c^2 \end{bmatrix}$$

  - with $\Omega = \sqrt{\Omega_1^2 + \Omega_2^2}$, $\alpha = \arctan(\Omega_2/\Omega_1)$, $c = \cos\alpha$ and $s = \sin\alpha$

➜ Single measurement trace $p_{k\ell}(t) = |\langle \ell | U(t,0) | k \rangle|^2$ contains information about both parameters, except $p_{22}(t)$

➜ Single trace should be sufficient to fully identify the Hamiltonian

Proc. ISCCSP 2010 (arXiv:0911.5429)

Bayesian Parameter Estimation

- Signal is linear combination of m_b basis functions, e.g., here $g_0 = 1, \quad g_1(t) = \cos(\Omega t), \quad g_2(t) = \cos(2\Omega t)$
- Define log-likelihood

$$P(\omega|\mathbf{d}) \propto \frac{m_b - N_t}{2} \log_{10}\left[1 - \frac{m_b \langle \mathbf{h}^2 \rangle}{N_t \langle \mathbf{d}^2 \rangle}\right],$$

  - where N_t is the number of data points, and

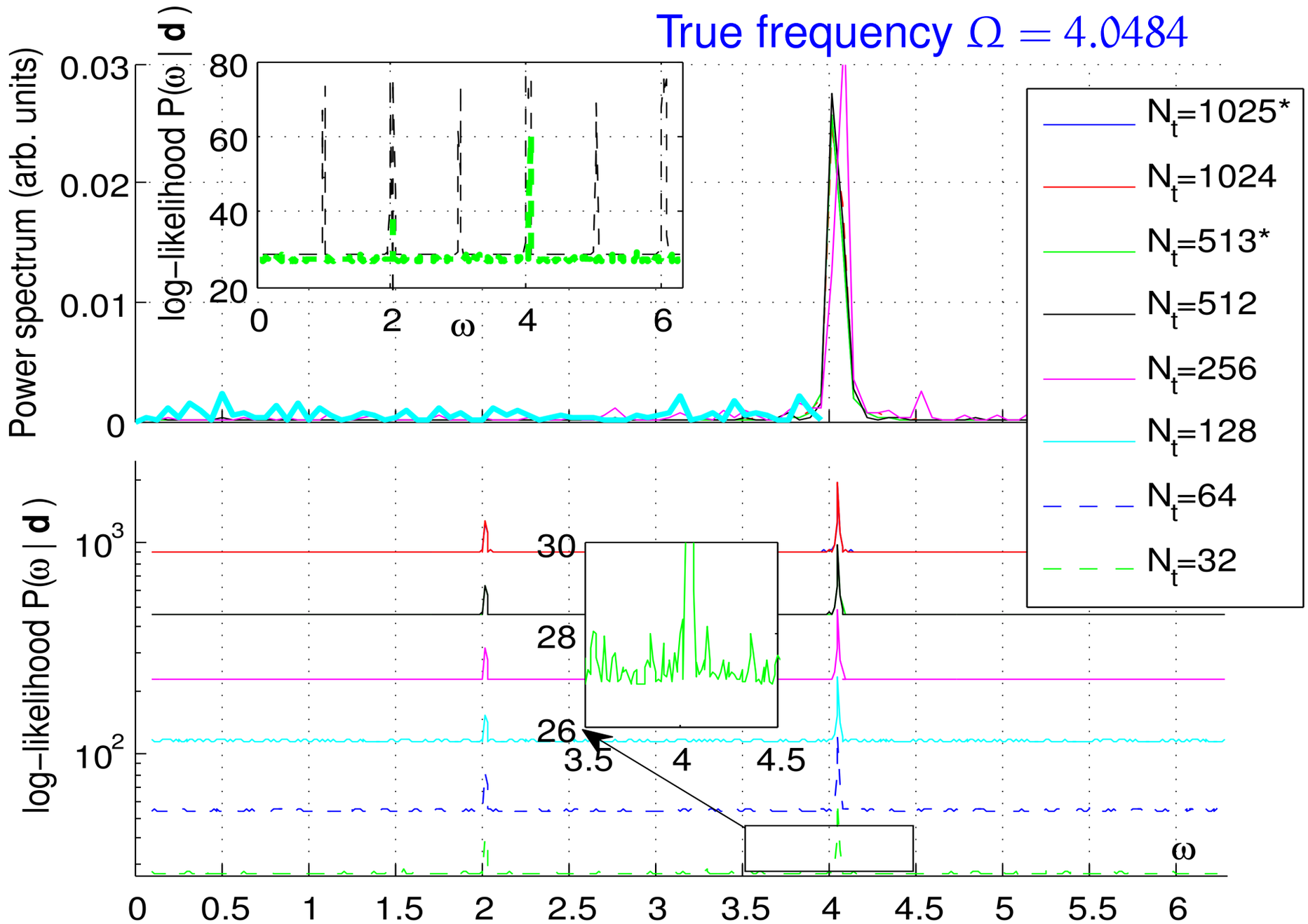
$$\langle \mathbf{d}^2 \rangle = \frac{1}{N_t} \sum_{n=0}^{N_t-1} d_n^2, \qquad \langle \mathbf{h}^2 \rangle = \frac{1}{m_b} \sum_{m=0}^{m_b-1} h_m^2,$$

  - where elements h_m of $(m_b, 1)$-vector $\mathbf{h}$ are projections of $(1, N_t)$-data vector $\mathbf{d}$ onto a set of orthonormal basis vectors derived from the non-orthogonal basis functions $g_m(t)$ evaluated at the respective sample times t_n

Maximizing Log-Likelihood

- $P(\omega|\mathbf{d})$ is function of single parameter
  - Ω is frequency for which $P(\omega|\mathbf{d})$ achieves global maximum
  - Coefficient vector $\mathbf{a}(\Omega)$ gives best estimate for $\cos^2\alpha$ and thus α
  - ➔ Minimize $||\mathbf{a}(x) - \mathbf{a}(\Omega)||$ with $a_m(x)$ as defined above
- ➔ Problem of finding the most likely model (Ω, α) reduced to finding global maximum of $P(\omega|\mathbf{d})$
- Finding maxima is difficult as likelihood function sharply peaked
  - In 1D exhaustive search possible
  - In general power spectrum can be used as pre-estimator
  - *Problem*: requires large number of sample points, ideally on regular grid, i.e. many experiments
  - ➔ Irregular sampling can substantially reduce number of data points required
    (low-discrepancy/quasi-random sampling provides optimal coverage by minimising gaps for given number of samples)

Power Spectra and Log-Likelihood



Results

- Data sampled at different times t in $[0, 100]$
  - For $N_t \geq 128$:
    - Power spectra have a single peak in the plotted range
    - Reasonable estimate for Ω
  - For $N_t \leq 64$:
    - Main peak is outside the range of the power spectrum
    - Power spectra no longer contain any useful information
- Log-likelihood still has a clearly identifiable global maximum at Ω even for data vectors with as few as 32 data points, provided a non-uniform sampling is used
- For uniform sampling with $N_t = 32$ the top inset shows that $P(\omega|\mathbf{d})$ has many peaks of approximately equal height due to aliasing effects (dashed black line)

Conclusions

- Power spectrum
  - Many data points and long signals required
  - regular sampling best
  - Peak detection easy
- Log-likelihood
  - Function sharply peaked even for very small number of sample points
  - Irregular low-discrepancy sampling strongly preferred
  - Peak detection difficult