Information geometry of quantum measurement processes

Jukka Kiukas
Department of Mathematics
Aberystwyth University

Collaboration: Madalin Guta
(Nottingham)

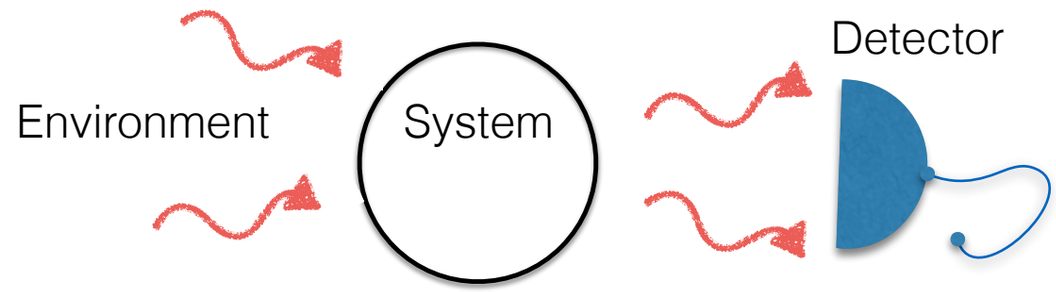
24 January 2017
Toulouse

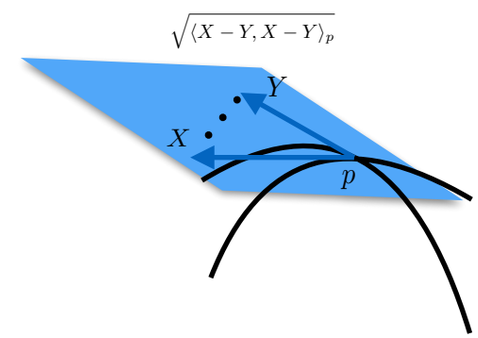
arXiv:1601.04355

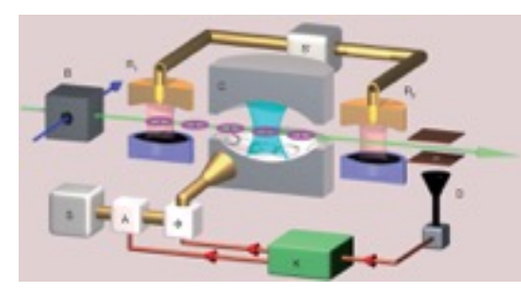
Comm. Math. Phys. 335 1397 (2015)

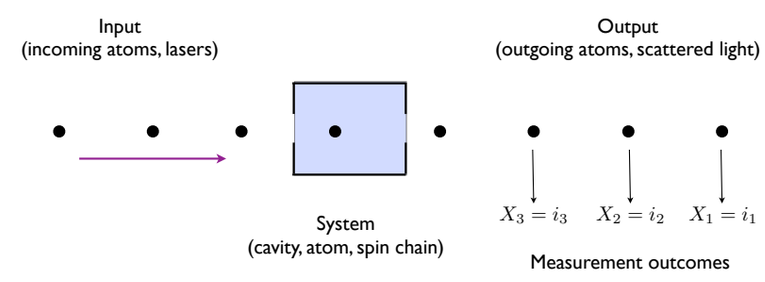
Overview

- **topic** — inference of noisy measurement processes with short memory
- **Intro: general statistical approach** — models, estimation theory, information (differential) geometry
- **Intro: quantum case** — motivation from quantum technologies, quantum vs classical statistics
- **Specific setting:** Quantum Markov processes
- **Results** — information geometry, Local Asymptotic Normality



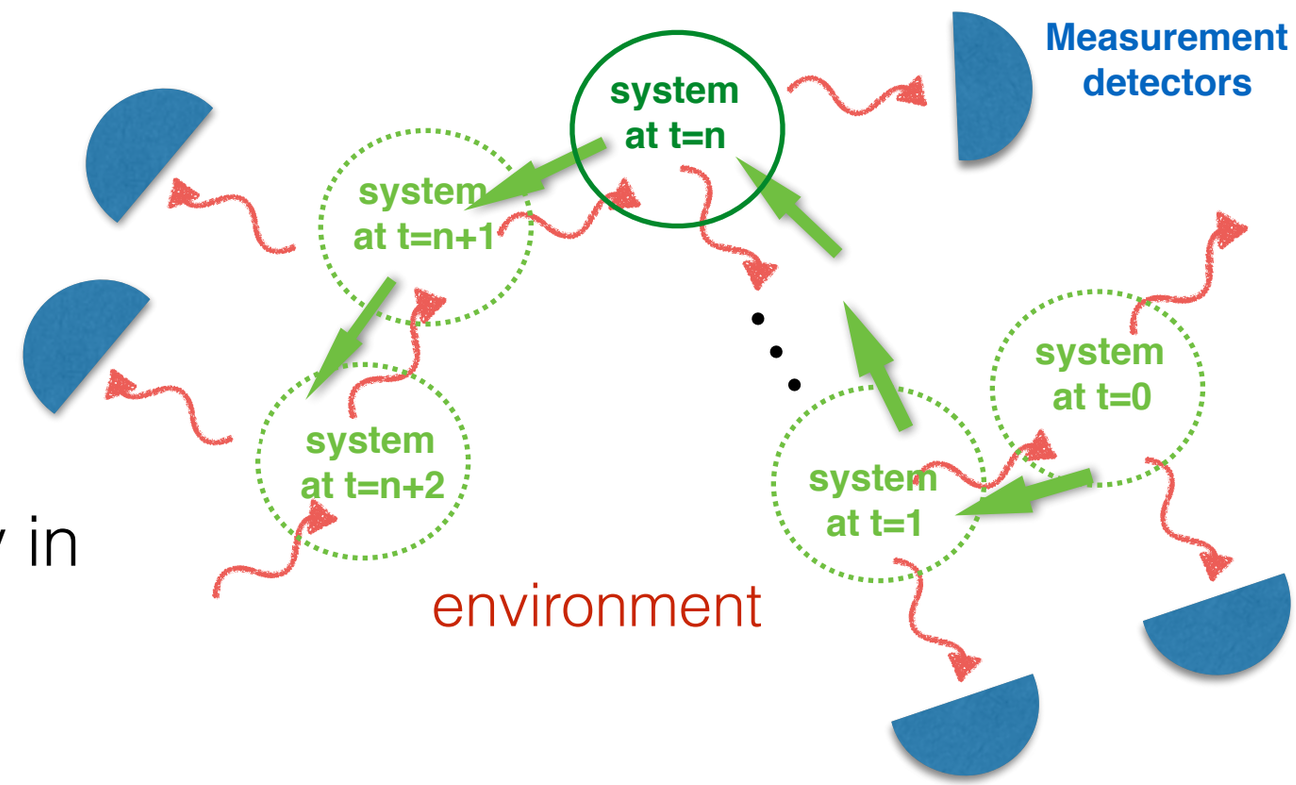


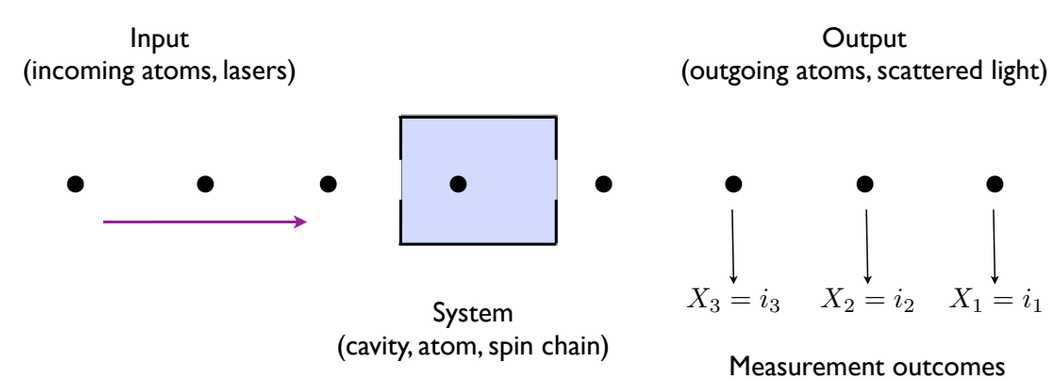




Continual measurement processes

- setting — (small) system of interest & environment
- system interacts with environment locally
- system "moves" dynamically in the environment
- environment "flows" through the system — "input-output"
- output contains information on the interaction process — accessed via measurements





Inference of measurement processes

- **Assumption** — partially unknown interaction

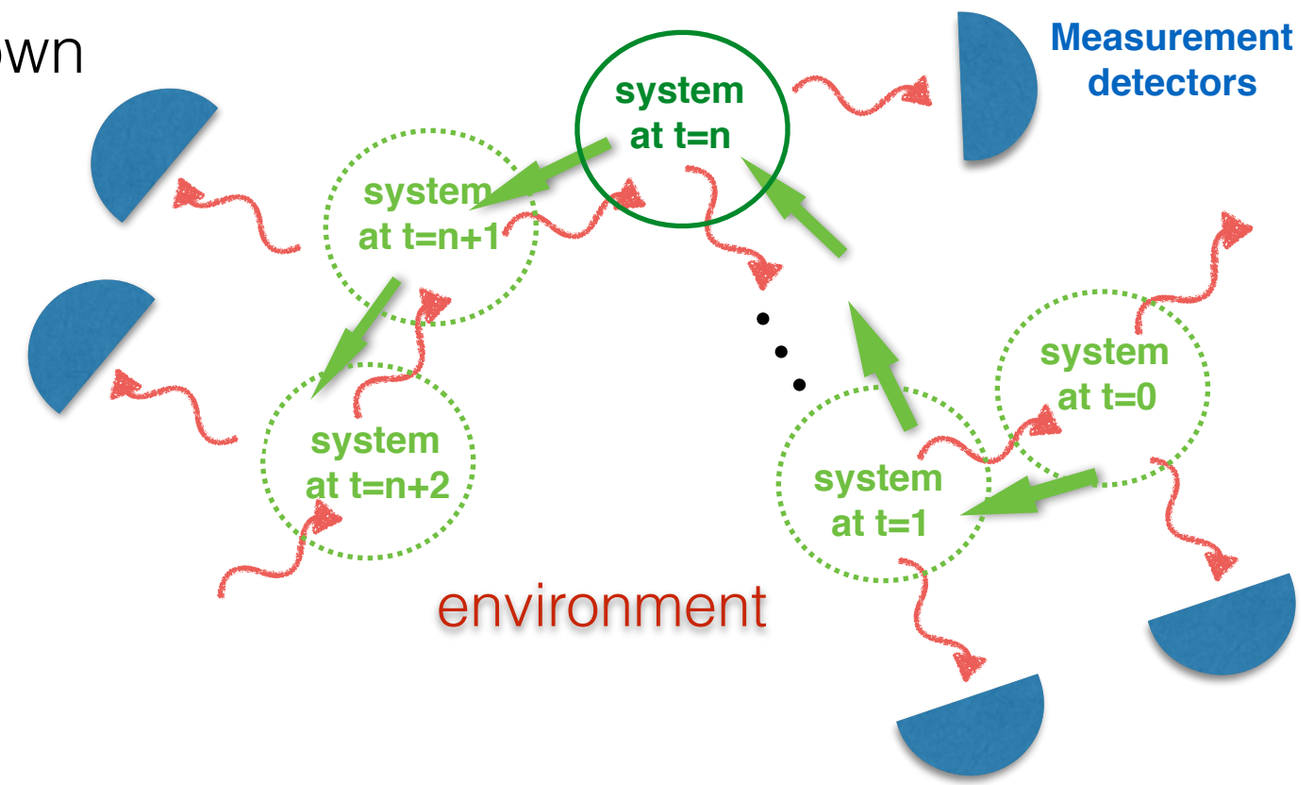
- **Goal** — infer the unknown parameters from detection statistics

➡ **Statistical inference problem**

  - use estimators

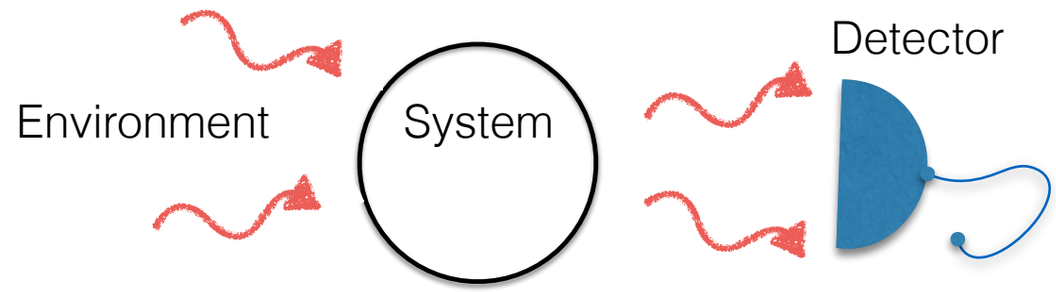
  - “volume” of the output trail — sample size

  - estimator accuracy increases dynamically

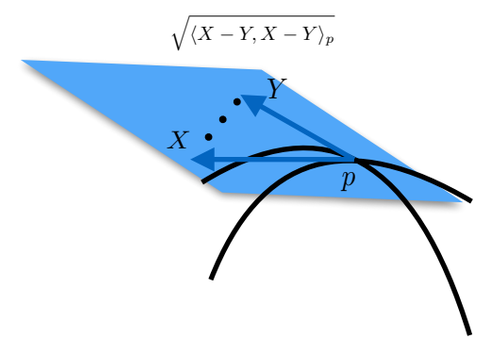


Overview

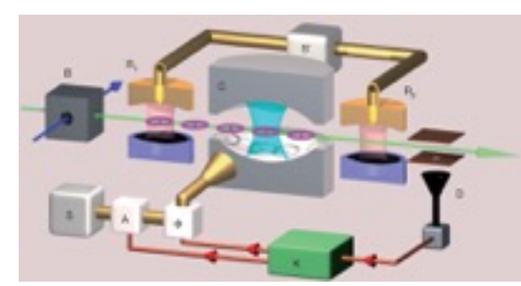
- **topic** — inference of continual noisy measurement processes with short memory



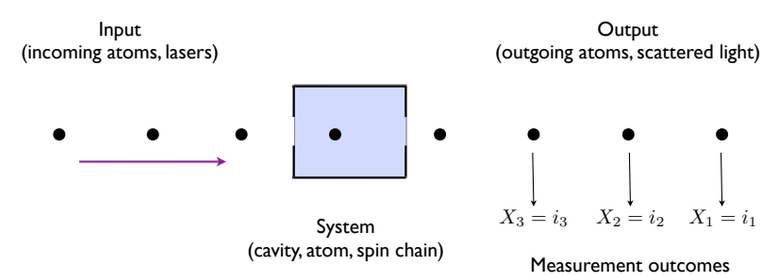
- **Intro: general statistical approach** — models, estimation theory, information geometry



- **Intro: quantum case** — motivation from quantum technologies, quantum vs classical statistics



- **Specific setting:** Quantum Markov processes



- **Results** — information geometry, Local Asymptotic Normality

Statistical inference - brief reminder

- **Statistical model**: a family of probability distributions
  - describes our ***assumptions*** of the system under study
  - contains ***parameters*** whose ***true values*** are unknown
- Parameters can be inferred by devising a statistic, **estimator**

statistical model

$$\{f_p(x) \mid p \in \Theta\}$$

estimator for p

$$\hat{p}(x)$$

random variable

Statistical inference

- Estimation error (typically) reduced by increasing sample size

$$f_p(x) \quad \rightarrow \quad f_p^{(n)}(x)$$

$$\hat{p}(x) \qquad\qquad \hat{p}^{(n)}(x)$$

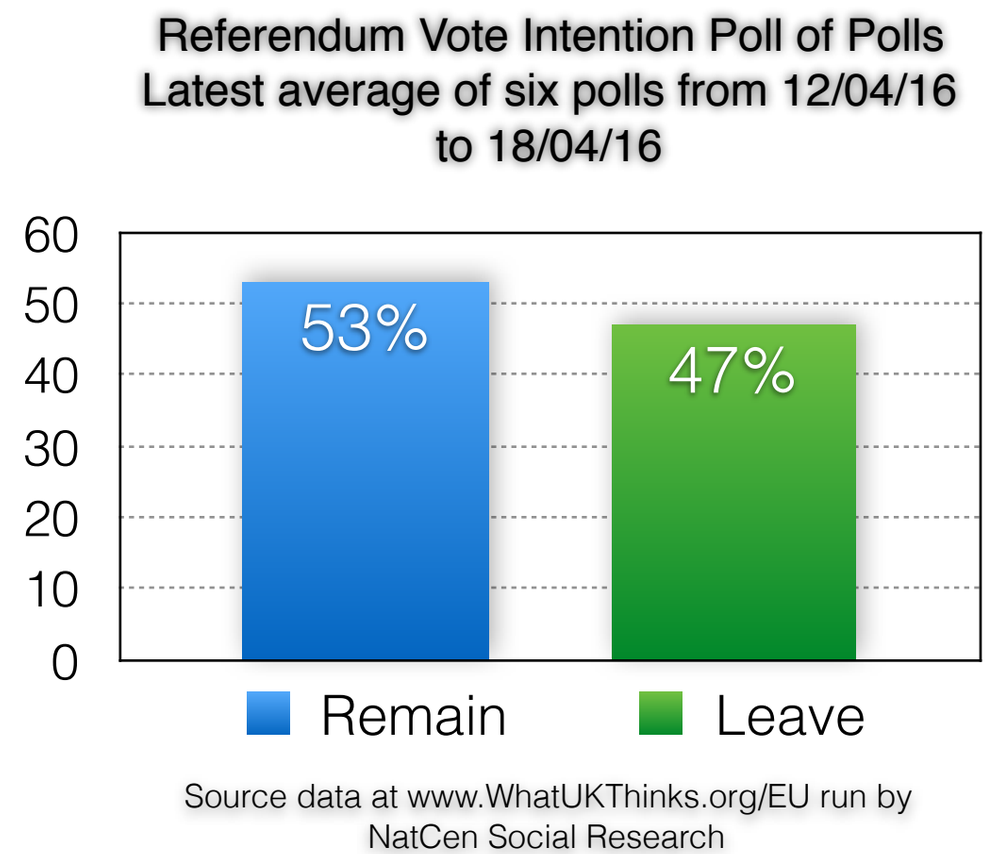
- The usual binomial example:

statistical model

$$f_p^{(n)}(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

estimator for p

$$\hat{p}^{(n)}(x) = \frac{x}{n}$$



Statistical inference

- Standard scaling

$$\mathrm{Var}(\hat{p}^{(n)}(x)) \sim 1/n$$

- Cramer-Rao bound:

$$\mathrm{Var}(\hat{p}^{(n)}(x)) \geq \frac{1}{nF}$$

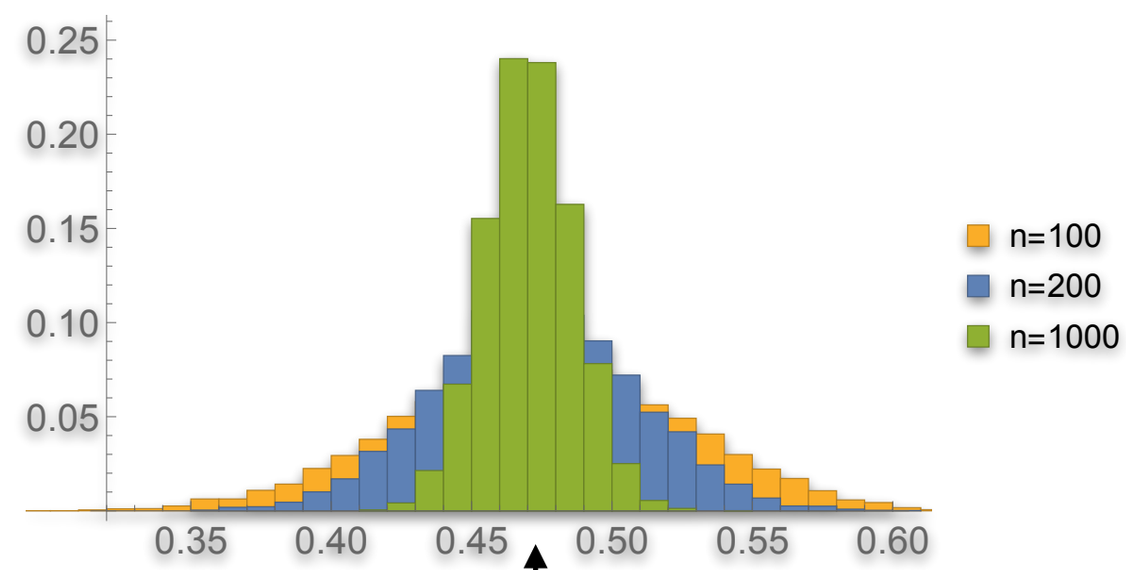
asymptotic **Fisher information** rate

- Optimal estimator achieves this bound

The binomial example:

$$\mathrm{Var}(\hat{p}^{(n)}(x)) = \frac{p(1-p)}{n}$$

$$F = \frac{1}{p(1-p)}$$



True value of p

Statistical inference

- Asymptotic Fisher information rate is the constant in the standard scaling of the error of an *optimal* unbiased estimator
- <u>Definition</u>: logarithmic derivative

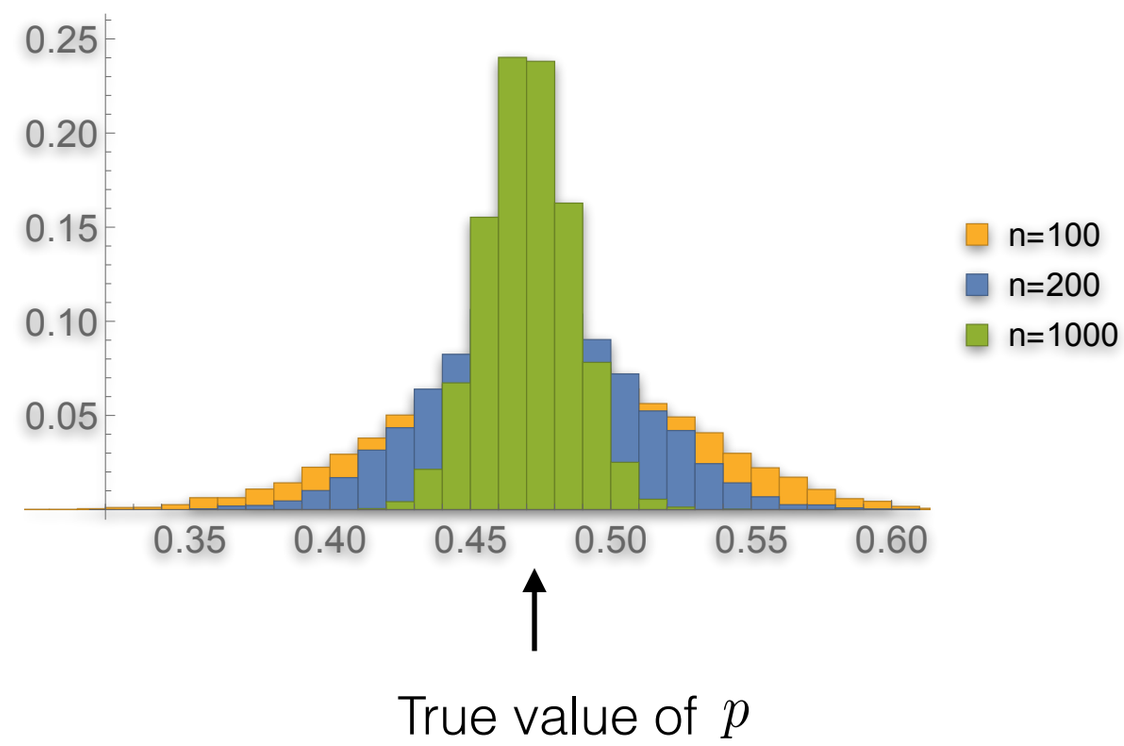
Fisher information per sample:

$$F^{(n)}(p) = \frac{1}{n}\mathbb{E}\left[\left(\frac{\partial}{\partial p}\log f_p^{(n)}(x)\right)^2\right]$$

Asymptotic rate:

$$F(p) = \lim_{n\to\infty} F^{(n)}(p)$$

Exists for MLE with sum of i.i.d. variables — but also in more general cases



“Classical” information geometry[1]

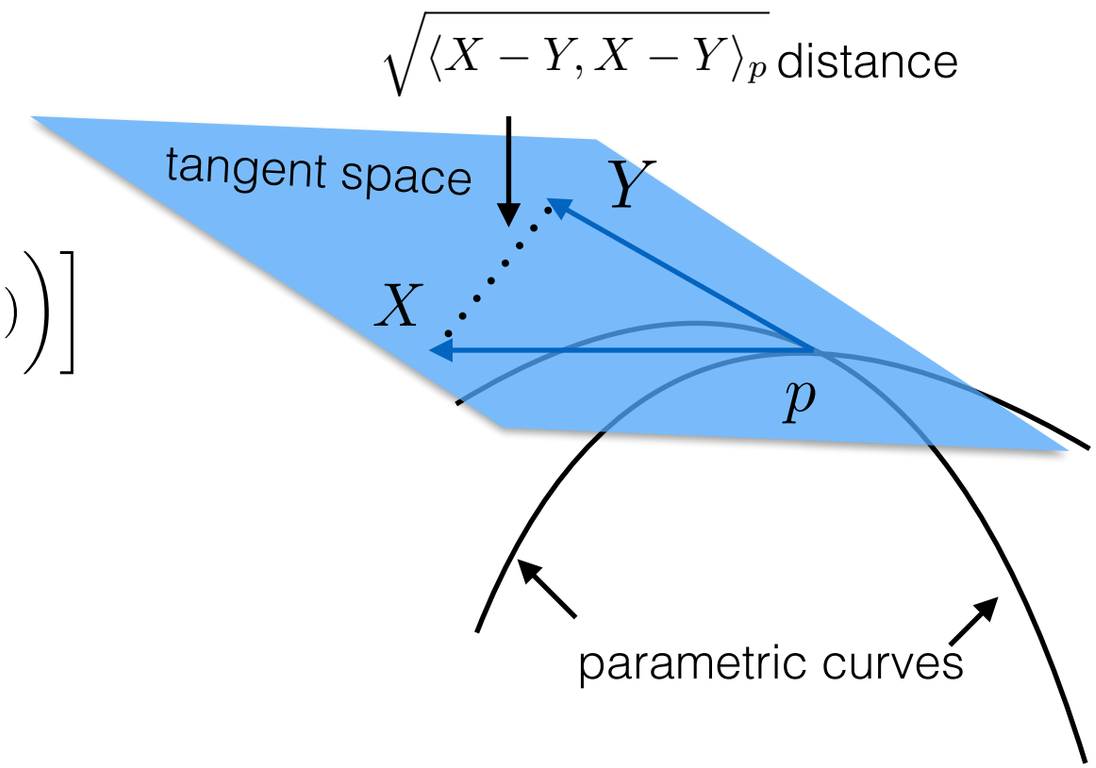
- <u>Multi-parameter setting</u>

  $p \in \mathcal{S}$ $\quad m$ -manifold

- Fisher information matrix:

$$F_{ij}^{(n)}(p) = \frac{1}{n}\mathbb{E}\left[\left(\frac{\partial}{\partial p_i}\log f_p^{(n)}(x)\right)\left(\frac{\partial}{\partial p_j}\log f_p^{(n)}(x)\right)\right]$$

$$F_{ij}(p) = \lim_{n\to\infty} F_{ij}^{(n)}(p)$$



➡ metric in ***full rank case:***

$$\langle X, Y\rangle_p = \sum_{ij} X^i Y^j F_{ij}(p)$$

- distinguishability distance for the inference problem

$$T_p(\mathcal{S}) = \left\{ X \mid X = \sum_{i=1}^{M} X^i \frac{\partial}{\partial p_i} \right\} \simeq \mathbb{R}^m$$ tangent space

[1] S. Amari. Differential-geometrical methods in statistics, vol 28 Lect. Notes Stat. (Springer, New York, 1985)

Degenerate case

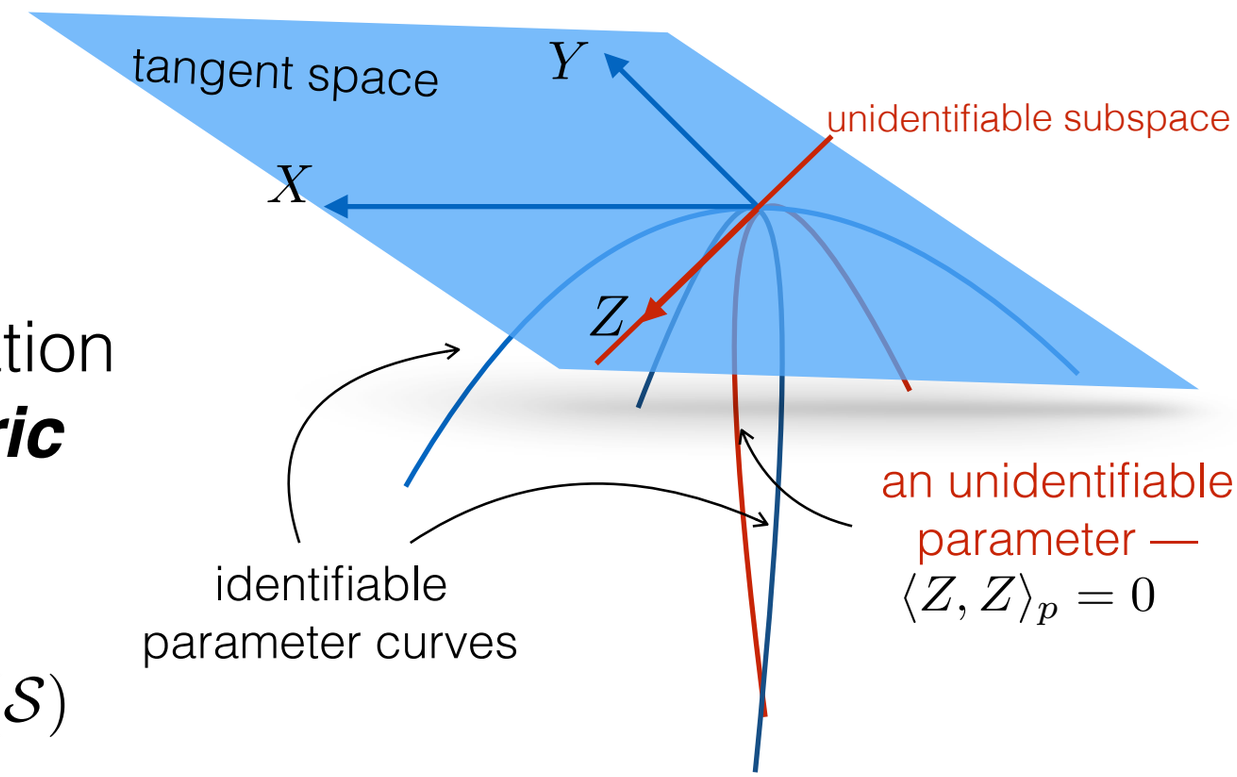
- **Inference problem**: identification of local parameter changes

- Interesting case: some parameters are unidentifiable

➡ they have zero Fisher information — $F_{ij}(p)$ ***not a proper metric***

➡ unidentifiable subspace:
$\mathcal{T}_p^{\text{nonid}} = \{Z \mid \langle Z, Z\rangle_p = 0\} \subset T_p(\mathcal{S})$

- A common instance: ***symmetries*** of the model form a group



Degenerate case — symmetries

- Suppose: symmetries of the model form a group G:

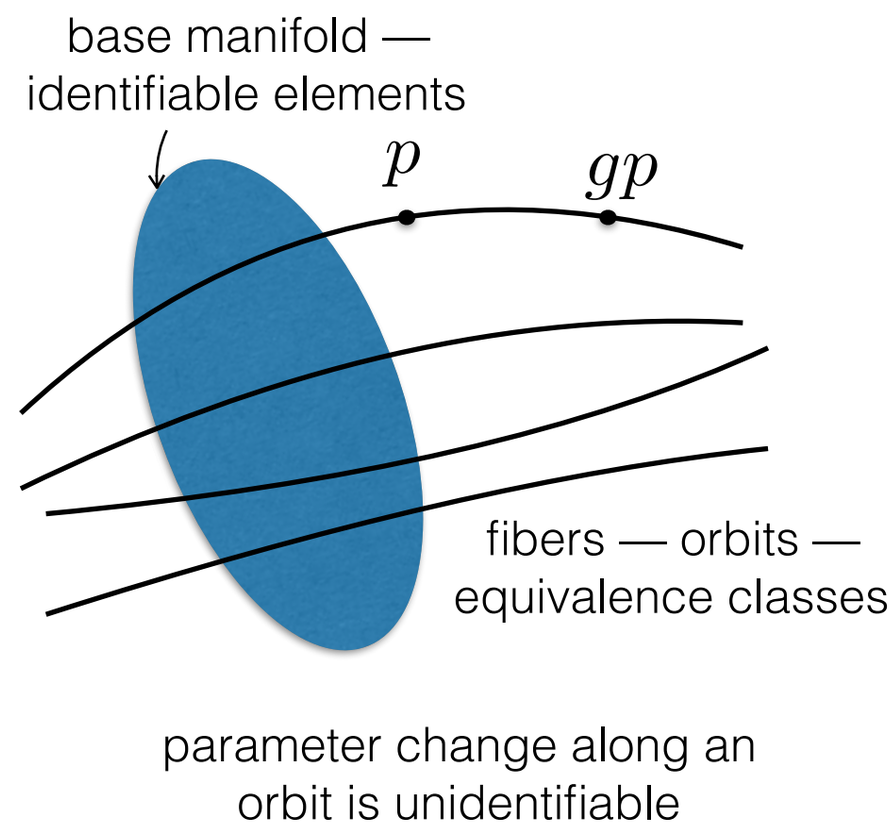
$$(\forall n : f_p^{(n)} = f_{p'}^{(n)}) \Leftrightarrow (\exists g \in G : p' = gp)$$

➡ only elements of $\mathcal{S}/G$ are identifiable

- principal bundle $\pi : \mathcal{S} \to \mathcal{S}/G$

- is there a natural way to describe the structure of $\mathcal{S}/G$ within the structure of $\mathcal{S}$?

➡ connections



Metric & connection

- principal bundle $\pi : \mathcal{S} \to \mathcal{S}/G$

- unidentifiable parameters — *vertical bundle*

  $\mathcal{T}^{\mathrm{nonid}} = \ker \pi_*$

- Is there a natural way to select complementary subspaces for identifiable directions?

➡ a *horizontal bundle* compatible with the metric:

$$\mathcal{T}^{\mathrm{id}} = P(T(\mathcal{S}))$$

$$P_p : T_p(\mathcal{S}) \to T_p(\mathcal{S}) \qquad P_p = P_p^2$$

$$\langle X, Y \rangle_p = \langle P_p X, P_p Y \rangle_p$$

tangent space — vertical bundle — horizontal bundle

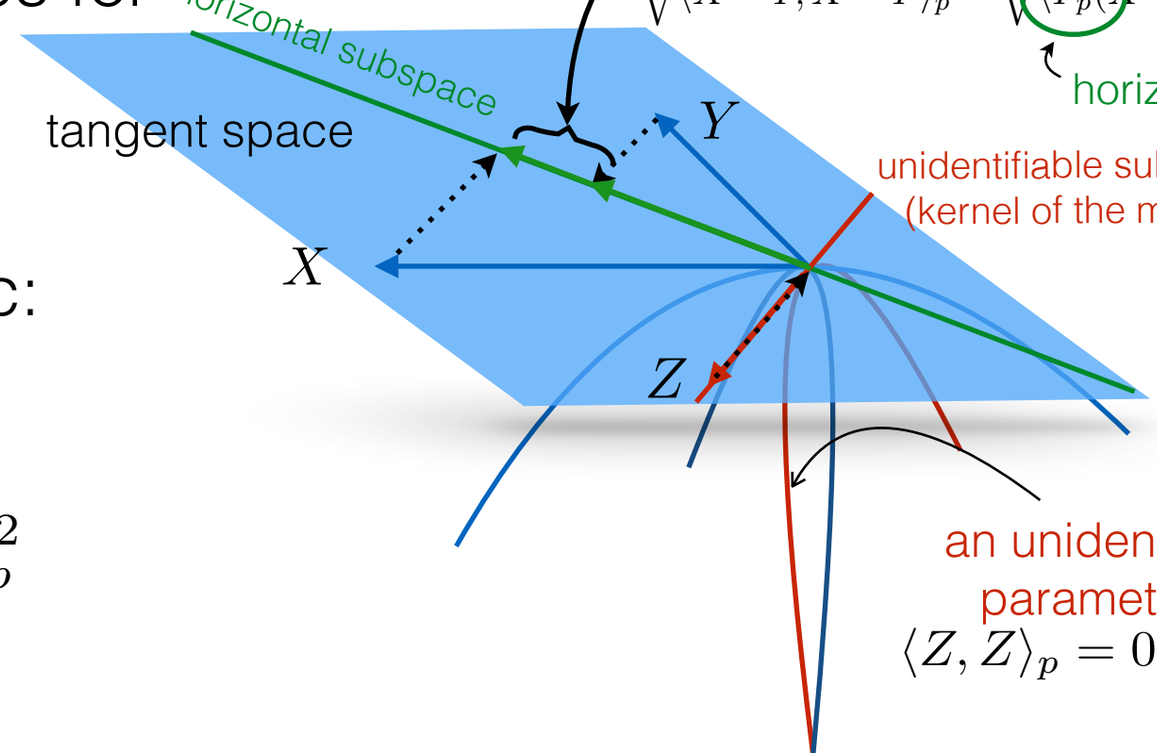
$$T_p(\mathcal{S}) = \mathcal{T}_p^{\mathrm{nonid}} \oplus \mathcal{T}_p^{\mathrm{id}}$$

metric is proper here!

distance on the identifiable part

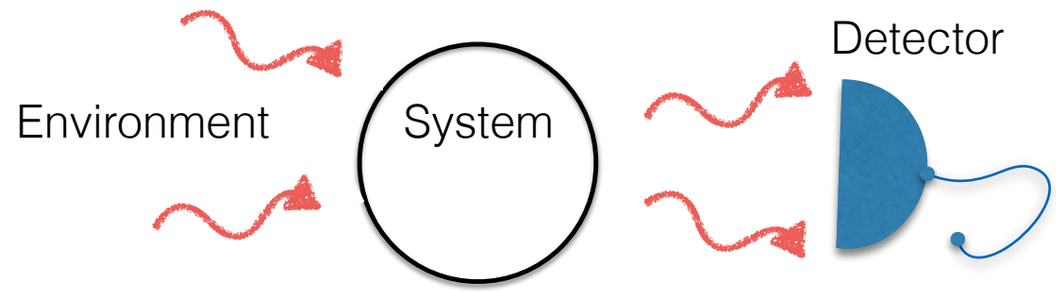
$$\sqrt{\langle X - Y, X - Y \rangle_p} = \sqrt{\langle P_p(X - Y), P_p(X - Y) \rangle_p}$$

horizontal projection

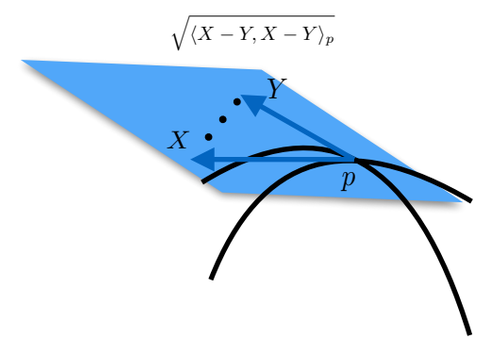


Overview

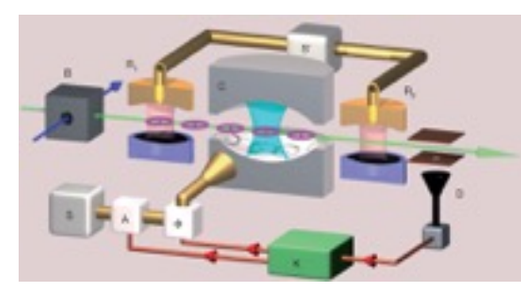
- **topic** — inference of continual noisy measurement processes with short memory



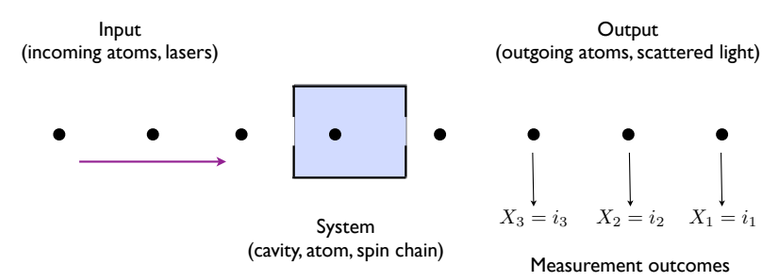
- **Intro: general statistical approach** — models, estimation theory, information geometry



- **Intro: quantum case** — motivation from quantum technologies, quantum vs classical statistics



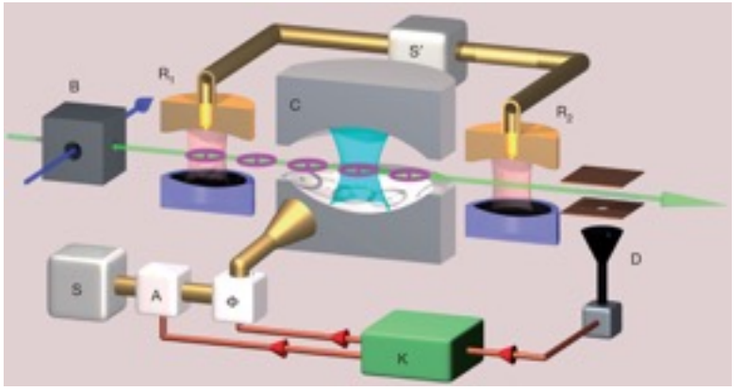
- **My specific setting:** Quantum Markov processes



- **Results** — information geometry, Local Asymptotic Normality

Application — quantum technology

- Quantum Technologies — a “1 billion future industry for the UK”[2]
  - High-precision devices
  - Secure communication
  - Quantum computation
  - **Relies heavily on statistical methods**



A quantum device: feedback control in atom maser (Sayrin et al. Nature 2011)

- Challenge — modelling noise
  - current experimental progress — need more accurate theory
  - **different aspects** (“quantum resource”, control, **system identification**)

- Topic of this talk — system identification
  - statistical models for quantum devices
  - inference of the unknown parameters of the model
  - **<u>question</u>: which quantum system actually sits inside a given device?**

[2] UK national quantum technologies programme http://uknqt.epsrc.ac.uk

Estimation in quantum statistics

quantum statistical model ρ_p → Measure-ment → classical statistical model $f_p(x) = \mathrm{tr}[\rho_p M(x)]$ → estimator $\hat{p}(x)$

- **Quantum statistical model**: a family of density matrices describing our assumptions on the state of a quantum system

quantum statistical model

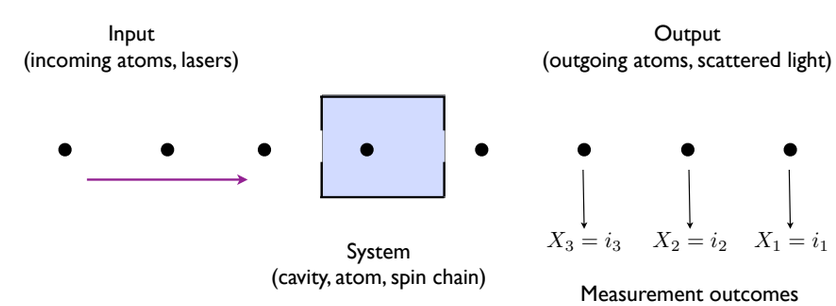
$$\rho_p \geq 0, \quad \mathrm{tr}[\rho_p] = 1$$

- **Quantum Fisher Information (QFI)** = maximal Fisher information over all classical models (measurements)

quantum measurement

$$0 \leq M(x) \leq \mathbb{I}, \ \sum_x M(x) = \mathbb{I}$$

- **Asymptotic estimation** (vague idea):
  1. prepare weakly correlated states $\rho_p^{(n)}$ on an extending system (say *n*-fold tensor product)
  ➡ QFI "per sample" $F^{(n)}$ should converge at $n \to \infty$
  2. construct an optimal sequence of measurements
  3. apply a (classical) estimator on the data



Quantum information geometry

- <u>Multi-parameter setting</u>

  $p \in \mathcal{S}$ $\quad m$ -manifold

- Asymptotic QFI matrix:

$$F_{ij}(p) = \lim_{n\to\infty} F_{ij}^{(n)}(p)$$

➡ similar structure as in the classical case

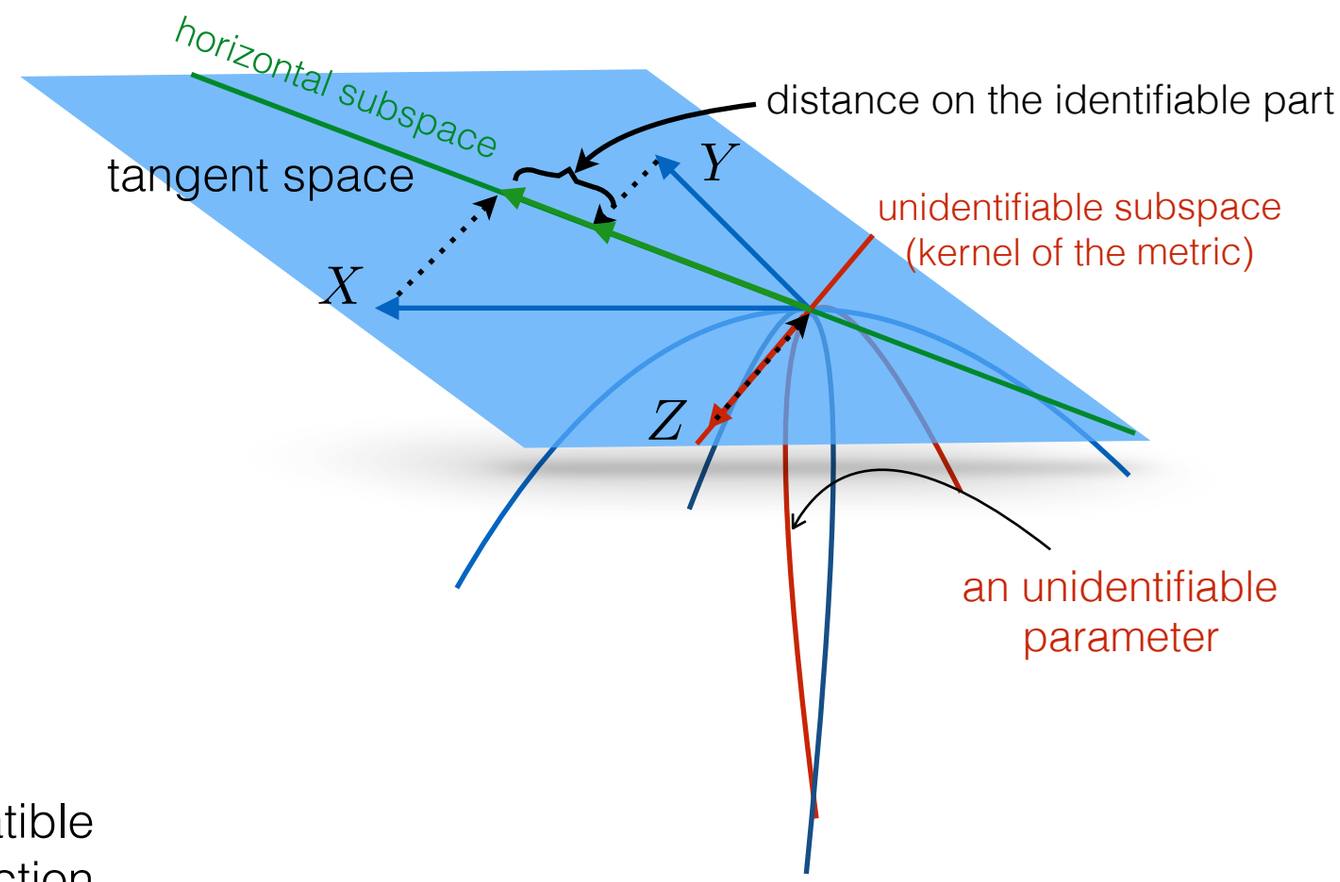
$$\langle X, Y\rangle_p = \sum_{ij} X^i Y^j F_{ij}(p)$$ metric

$$P_p : T_p(\mathcal{S}) \to T_p(\mathcal{S})$$

$$\langle X, Y\rangle_p = \langle P_p X, P_p Y\rangle_p$$

$$\mathcal{T}^{\mathrm{id}} = P(T(\mathcal{S}))$$

compatible connection — identifiable and non-identifiable directions

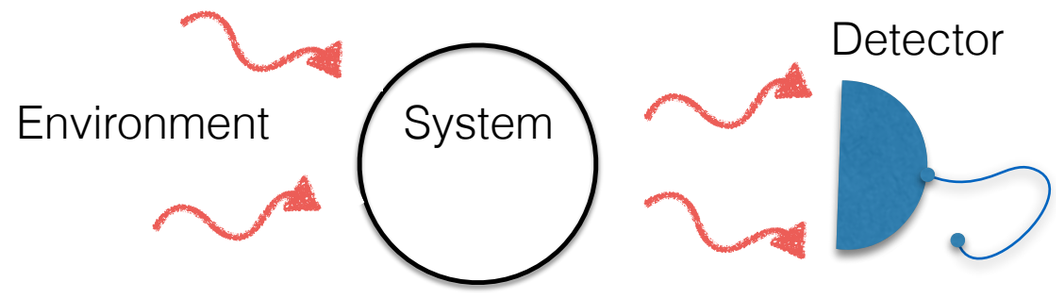


$$T_p(\mathcal{S}) = \mathcal{T}_p^{\mathrm{nonid}} \oplus \mathcal{T}_p^{\mathrm{id}}$$

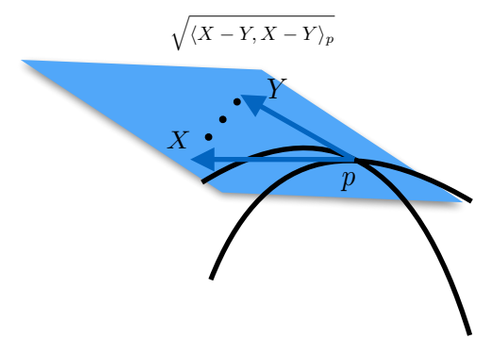
vertical bundle horizontal bundle

Overview

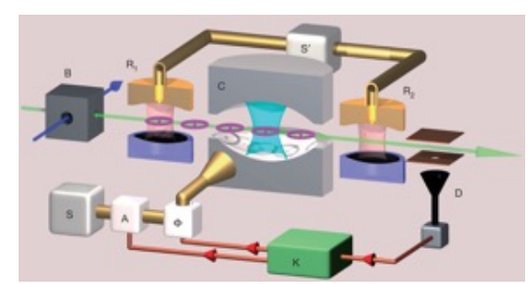
- **topic** — inference of noisy measurement processes with short memory



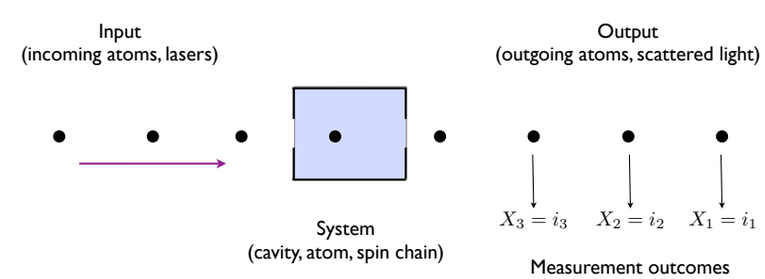
- **Intro: general statistical approach** — models, estimation theory, information geometry



- **Intro: quantum case** — motivation from quantum technologies, quantum vs classical statistics



- **Specific setting:** Quantum Markov processes
- **Results** — information geometry, local asymptotic normality

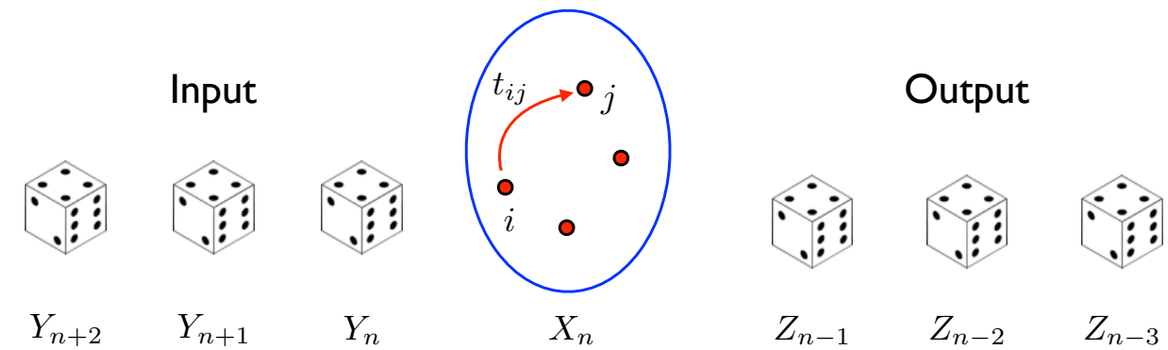


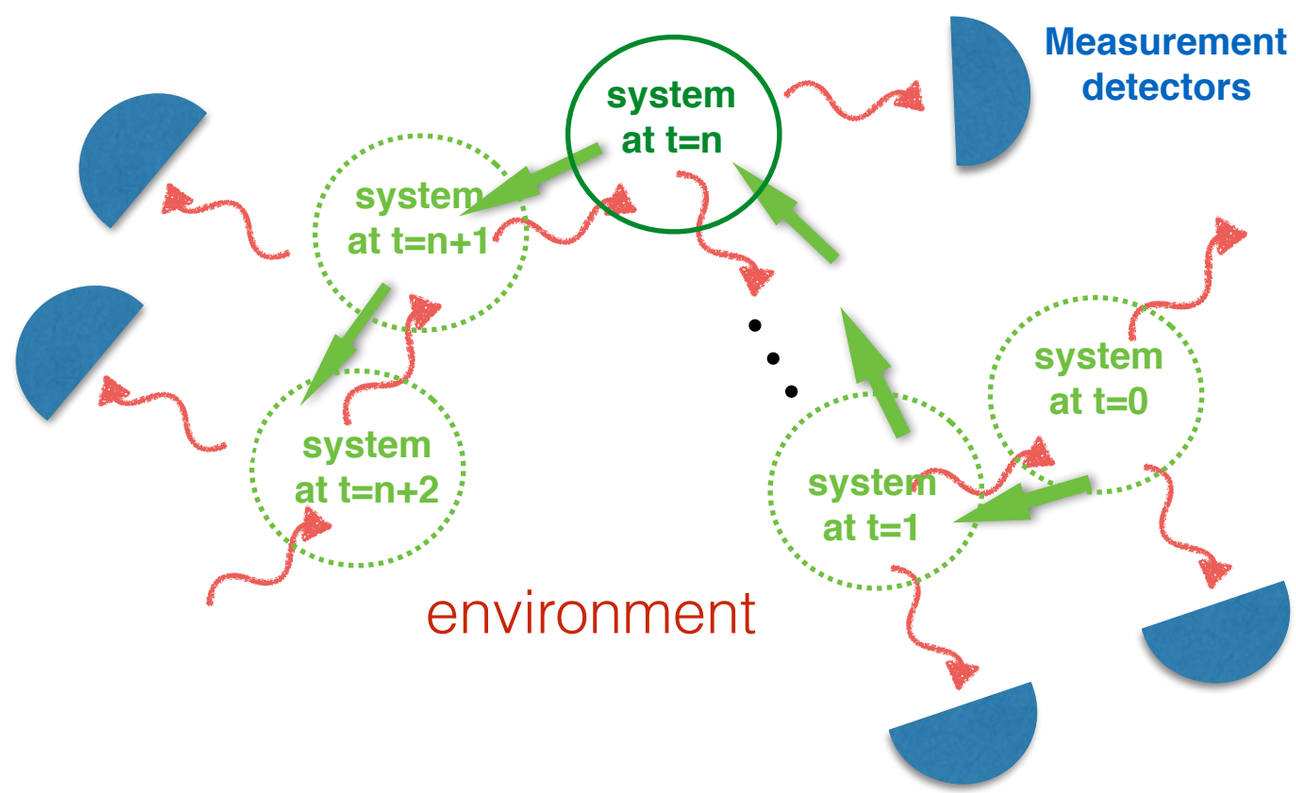
Intuition— classical hidden Markov chain

- Scattering interaction between input (environment) and system

  $(X_n, Y_n) \mapsto F(X_n, Y_n) = (X_{n+1}, Z_n)$

- Input: i.i.d. random variables $Y_1, Y_2, \ldots$
- System: induced Markov process $X_1, X_2, \ldots$
- Output: ***trajectory*** of weakly correlated variables $Z_1, Z_2, \ldots$
- **Statistical problem**: estimate dynamics F by observing the output*





* T. Petrie, Ann. of Math. Stat. (1969). P.J. Bickel, Y. Ritov, and T. Ryden, Ann. Stat. (1998)

Quantum Markov processes

- environment = Bosonic Fock space over $L^2(\mathbb{R}_+, dt) \otimes \mathcal{L}$

  time space — outcome space $\mathcal{L} = \mathrm{span}\{|i\rangle \mid i \in \mathcal{I}\}$

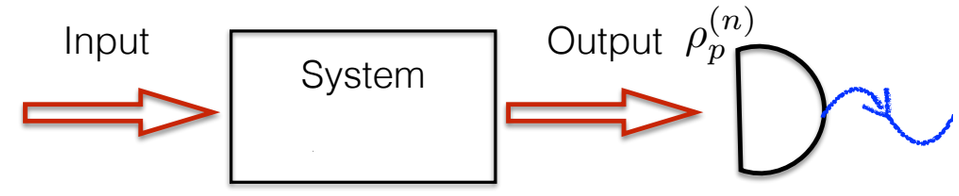
- unitary dynamics

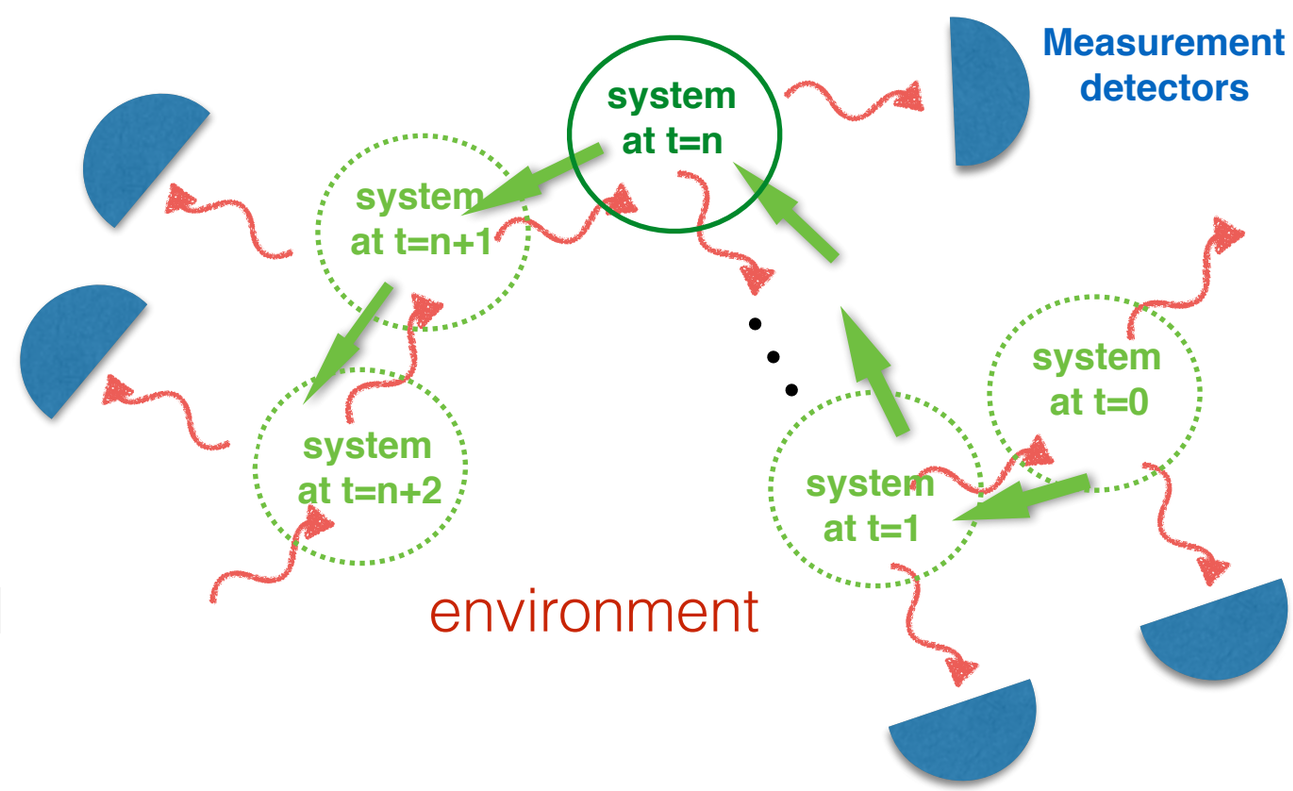
- input = field vacuum

- output = weakly correlated field states $\rho_p^{(n)}$

- **Statistical problem**: estimate dynamics from output

➡ **Our goal**: derive information geometry for the asymptotic QFI of the output states





Quantum jump trajectories

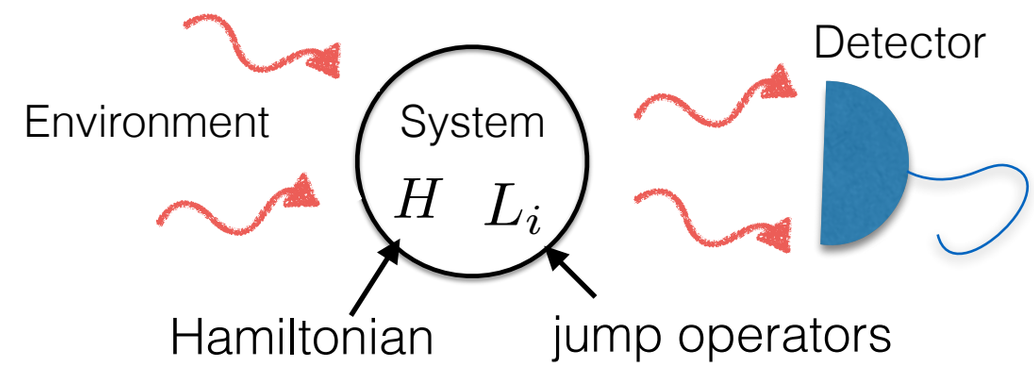
- Observations at random times induce "quantum jumps" in the system

➡ Measurement record = trajectory

$$((t_1, i_1), \ldots, (t_n, i_n))$$

jump time

outcome — corresponds to jump $\psi \mapsto L_i \psi$



system evolution between jumps is contractive:

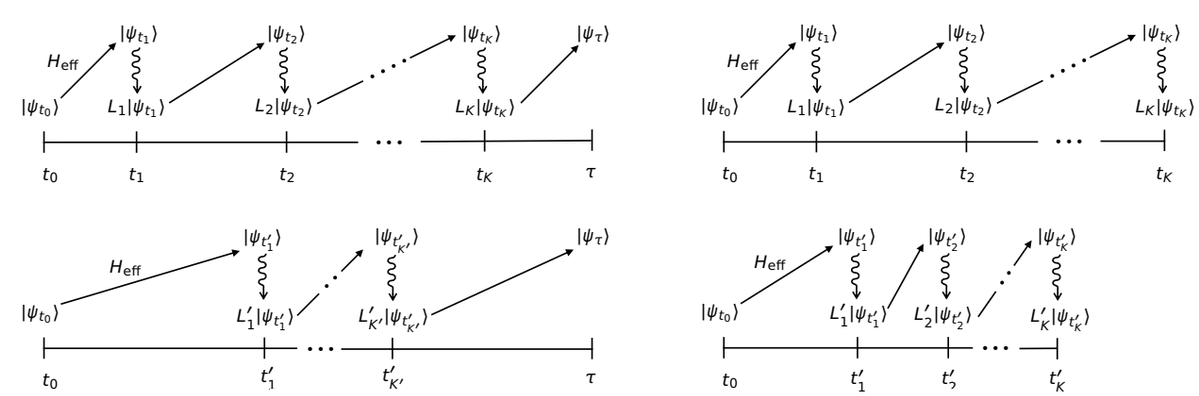
$$H_{\text{eff}} = H - \frac{i}{2} \sum_{i=1}^{M} L_i^* L_i$$

- Parameter manifold: $\mathcal{S} = \{(H, L_1, \ldots, L_M) \mid H \in M_{d,sa}(\mathbb{C}),\ L_i \in M_d(\mathbb{C})\}$

- **Two distinct observation schemes***:

1. observe until a given time:
   $((t_1, i_1), \ldots, (t_n, i_n)), \quad n \text{ arbitrary}, \quad t_n \leq t$

2. observe until given number of jumps:
   $((t_1, i_1), \ldots, (t_n, i_n)), \quad n \text{ fixed}, t_i \text{ arbitrary}$

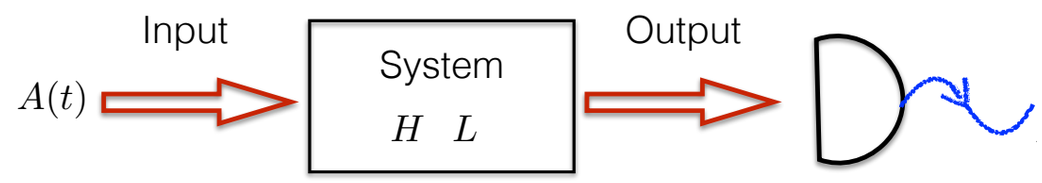


(1.) (2.)

* J. Kiukas, M. Guta, J.P. Garrahan, and I. Lesanovsky. Phys. Rev. E, 92:012132, 2015

Scheme 1 — time-extensive

- Evolve until time *t*



- The dynamics $U_p(t)$ is up to a phase given by*

$$dU_p^{(t)} = \left(-iHdt + \sum_i (L_i dA_i^* - L_i^* dA_i - \frac{1}{2} L_i^* L_i dt)\right) U_p^{(t)}$$

Quantum Stochastic Differential Equation

➡ **output + system state**:

$$\Psi_p^{(t)} = U_p^{(t)} |\psi_0 \otimes \mathrm{vac}\rangle$$

$$[a_k(t), a_{k'}^*(t')] = \delta_{kk'}\delta(t-t')$$

creation and annihilation operators

$$A_k(t) = \int_0^t a_k(t) dt$$

quantum Wiener processes

- average system dynamics follows the master equation

$$\frac{d}{dt}\rho_t = \mathbb{W}(\rho_t)$$

system density matrix at time t

$$\mathbb{W}_p(X) = \mathbb{G}_p(X) + \sum_i L_i^* X L_i$$

evolution generator

$$\mathbb{G}_p(X) = iXH_{\mathrm{eff}} - iH_{\mathrm{eff}}^* X$$

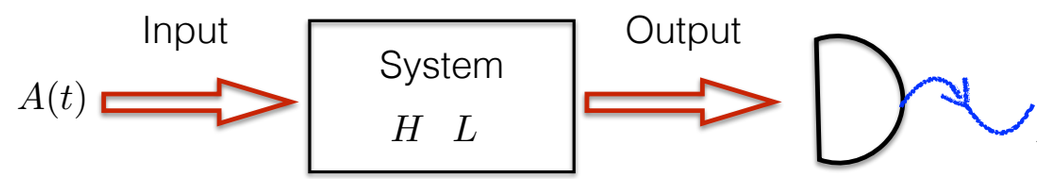
"drift" generator

assume strong ergodicity: there is a unique stationary state with $\lim_{t\to\infty} e^{t\mathbb{W}_p}(X) = \mathrm{tr}[\rho_p^{\mathrm{time}} X]\mathbb{I}$

*R.L. Hudson, C.Gardiner, P.Zoller, Quantum Noise (Springer, Berlin, 2004); V. P. Belavkin, J. Multiv. Anal. 42 171 (1992); R. S. Hudson, K. R. Parthasarathy, Comm. Math. Phys. 93 301 (1984)

Scheme 2 — count-extensive

- Evolve until a given number of jumps *n*



- At each jump record the time & outcome in the state $V_p^{(n)}\psi_0 \in \mathcal{H} \otimes L^2(\Delta_n, dt_1 \cdots dt_n) \otimes \mathcal{L}^{\otimes n}$

System time space (ordered)

$$(V_p^{(n)}\psi_0)(t_1,\ldots,t_n) = \sum_{i_1,\ldots,i_n} L_{i_n} e^{-i(t_n - t_{n-1})H_{\text{eff}}} \cdots L_{i_1} e^{-it_1 H_{\text{eff}}} \psi_0 \otimes |i_1,\ldots,i_n\rangle$$

➡ **output + system** state $\Psi_p^{(n)} = V_p^{(n)}\psi_0$

- Average system dynamics is a *discrete* Markov chain

transition channel: $T_p(X) = [V_p^{(1)}]^*(X \otimes \mathbb{I}_{\text{out}})V_p^{(1)}$

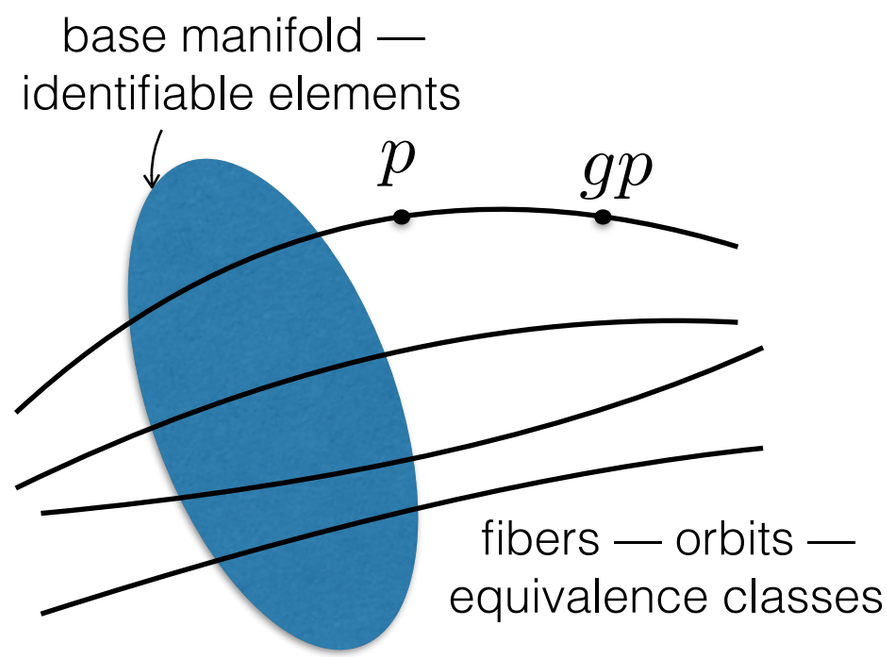
assume strong ergodicity: unique stationary state with $\lim_{n\to\infty} T_p^n(X) = \text{tr}[\rho_p^{\text{count}} X]\mathbb{I}$

The stationary states of the two schemes are different but related:

$$\rho_p^{\text{count}} = r_p[\mathbb{G}_p]_*(\rho_p^{\text{time}}) \qquad \mathbb{G}_p(X) = iXH_{\text{eff}} - iH_{\text{eff}}^* X$$

Symmetries

- The parameters $p = (H, L_1, \ldots, L_M)$ are *not* fully identifiable from the output



- **Result**: the symmetry groups coincide *up to a phase*:

  1. Time-extensive scheme

     $$G = PU(d) \times \mathbb{R}$$
     $$p = (H, L) \mapsto gp = (W^* H W + c\mathbb{I}, W^* L W), \quad g = (W, c)$$

  2. Count-extensive scheme

     $$G = PU(d) \times U(1)$$
     $$p = (H, L) \mapsto gp = (W^* H W, cW^* L_1 W, \ldots, cW^* L_M W), \quad g = (W, c)$$

Obtaining the asymptotic geometry

- Show that the QFI rate converges as the extensive parameter goes to infinity

Scheme 1: $F_{ij}(p) = \lim_{t\to\infty} F_{ij}^{(t)}(p)$

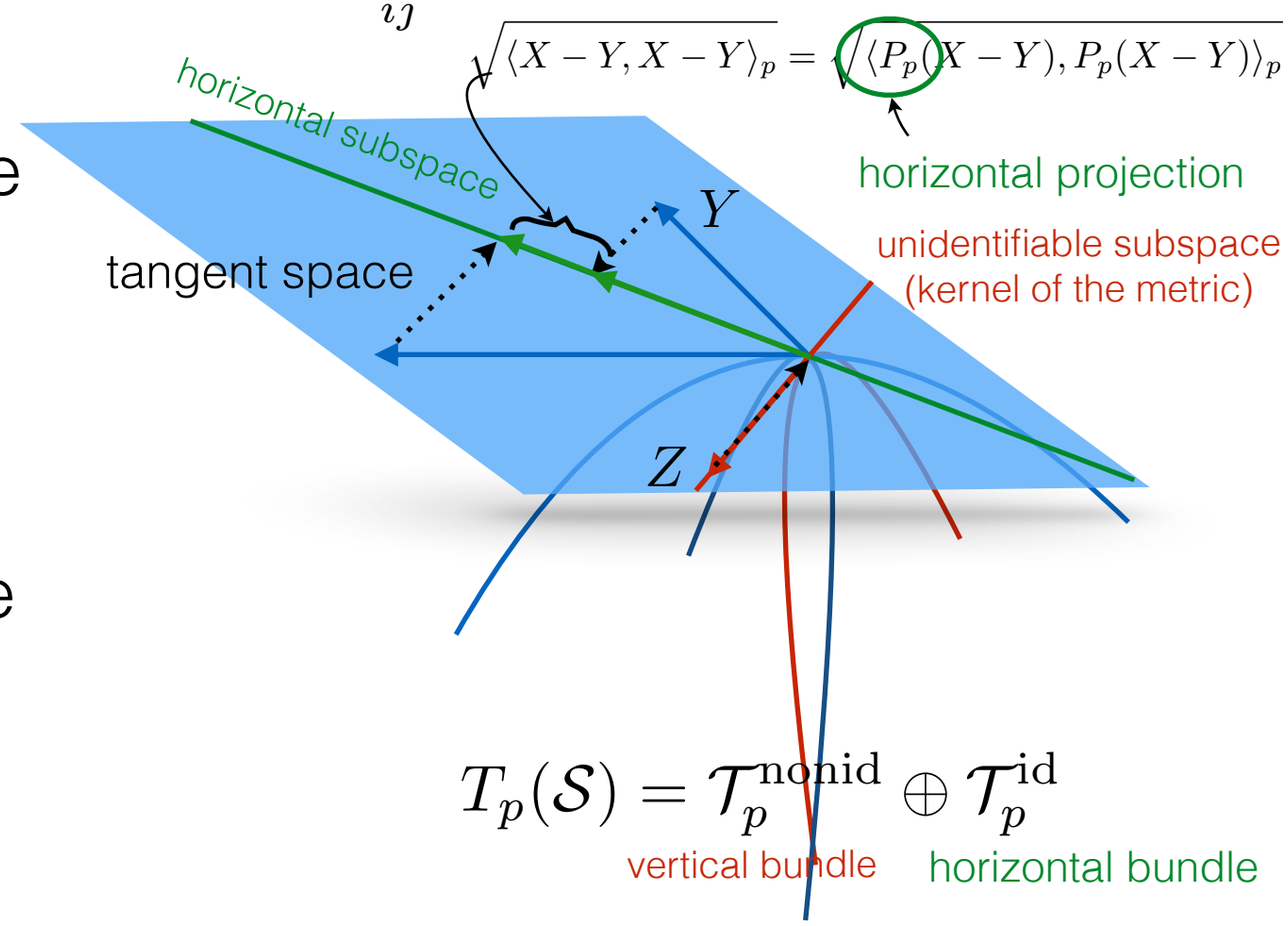
Scheme 2: $F_{ij}(p) = \lim_{n\to\infty} F_{ij}^{(n)}(p)$

- Write the metric in terms of a projection

$$\langle X, Y\rangle_p = \sum_{ij} X^i Y_j F_{ij}(p) = \langle P_p X, P_p Y\rangle_p$$

- Metric should be simple in the range of the projection (horizontal bundle)

➡ the nontrivial part of the geometry encapsulated in the connection



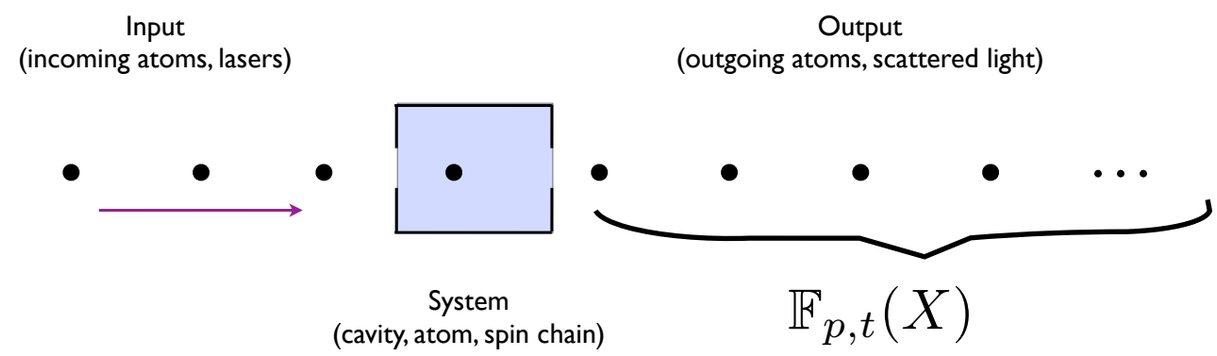
Obtaining the asymptotic Fisher metric — time-extensive case

- *How do we take the limit?*

- Define a sesquilinear form

$$(X,Y)_p(t) = \langle \varphi \otimes \mathrm{vac} | \mathbb{F}_{p,t}(X)^* \mathbb{F}_{p,t}(Y) \varphi \otimes \mathrm{vac} \rangle$$

$$X, Y \in T_p(\mathcal{S})$$



- covariances of quantum stochastic fluctuators

Quantum stochastic fluctuator

$$\mathbb{F}_{p,t}(X) = \frac{1}{\sqrt{t}} \int_0^t \left(i \sum_{k=1}^N X^k(t) dA_k^\dagger(t) + \mathcal{E}_p(X)(t) dt \right)$$

$$\mathcal{E}_p(X) = X^0 + \mathrm{Im} \sum_k [X^k]^* L_k$$

- **Prop 1.** (from Q Ito calculus):

$$\langle X, Y \rangle_p(t) = \mathrm{Re}(X,Y)_p(t)$$

QFI metric fluctuator covariance

- **Prop 2.** Fluctuator covariance converges at $t \to \infty$

<u>Compare</u>: CLT for i.i.d mean zero variables

$$\lim_{t\to\infty} \mathrm{Var}\left(\frac{1}{\sqrt{t}} \sum_{i=1}^t X_i \right) = \sigma^2$$

Equivalence of information geometries

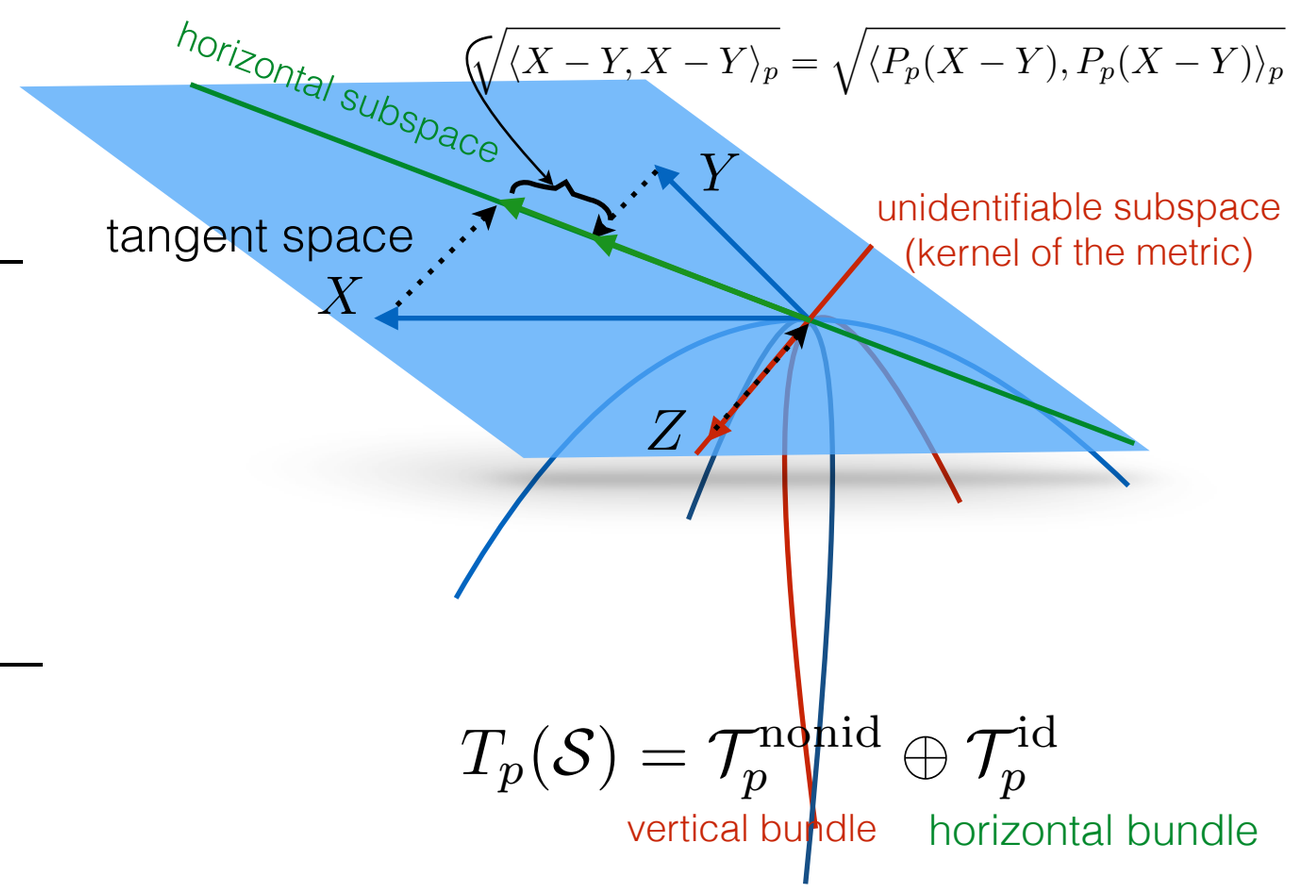
- The two schemes coincide in $\mathcal{M}_p = \{X \in T_p(\mathcal{S}) \mid \mathrm{tr}[\rho_p^{\mathrm{time}} \mathcal{E}_p(X)] = 0\}$

- Explicitly:

$$P_p(X) = X - (i[H, \mathbb{W}_p^{-1} \circ \mathcal{E}_p(X)], i[L_1, \mathbb{W}_p^{-1} \circ \mathcal{E}_p(X)], \ldots, i[L_M, \mathbb{W}_p^{-1} \circ \mathcal{E}_p(X)])$$

$$\langle X, Y \rangle_p = \mathrm{Re} \sum_{i=1}^{M} \mathrm{tr}[\rho_p^{\mathrm{time}} X_i^* Y_i] \quad \text{for all} \quad X, Y \in \mathrm{Ran}\, P_p$$

- Summary of the structure:
  - unidentifiable parameters — subspace $\mathcal{T}_p^{\mathrm{nonid}} = \ker P_p$
  - identifiable parameters — subspace $\mathcal{T}_p^{\mathrm{id}} \simeq \mathrm{ran}\, P_p$
  - compatible with symmetry — $\mathfrak{g} \simeq \mathcal{T}_p^{\mathrm{nonid}}$
    Lie algebra



Symplectic structure

- We have $\langle X, Y\rangle_p = \mathrm{Re}\sum_{i=1}^{M} \mathrm{tr}[\rho_p^{\mathrm{time}} X_i^* Y_i]$ on $\mathcal{T}_p^{\mathrm{id}} \simeq \mathrm{ran}\, P_p$
- What about the imaginary part?

➡ **symplectic form** $\Sigma(X, Y)_p = \mathrm{Im}\sum_{i=1}^{M} \mathrm{tr}[\rho_p^{\mathrm{time}} X_i^* Y_i]$

  - Non-degenerate on $\mathcal{T}_p^{\mathrm{id}} \simeq \mathrm{ran}\, P_p$

➡ a **CCR algebra:**

Weyl operators

$$W(X)W(Y) = e^{-i\Sigma(X,Y)_p} W(X+Y)$$

Vacuum state

$$\langle\Omega|W(X)|\Omega\rangle = e^{-\langle X,X\rangle_p/2}$$

- Why is this useful?
  - geometry is imprinted in a quantum phase space
  - quantum LAN

Reduction of the inference problem (LAN)

➡ We get a CCR algebra on the identifiable subspace $\mathcal{T}_p^{\mathrm{id}} \simeq \operatorname{ran} P_p$

$$W(X)W(Y) = e^{-i\Sigma(X,Y)_p} W(X+Y)$$

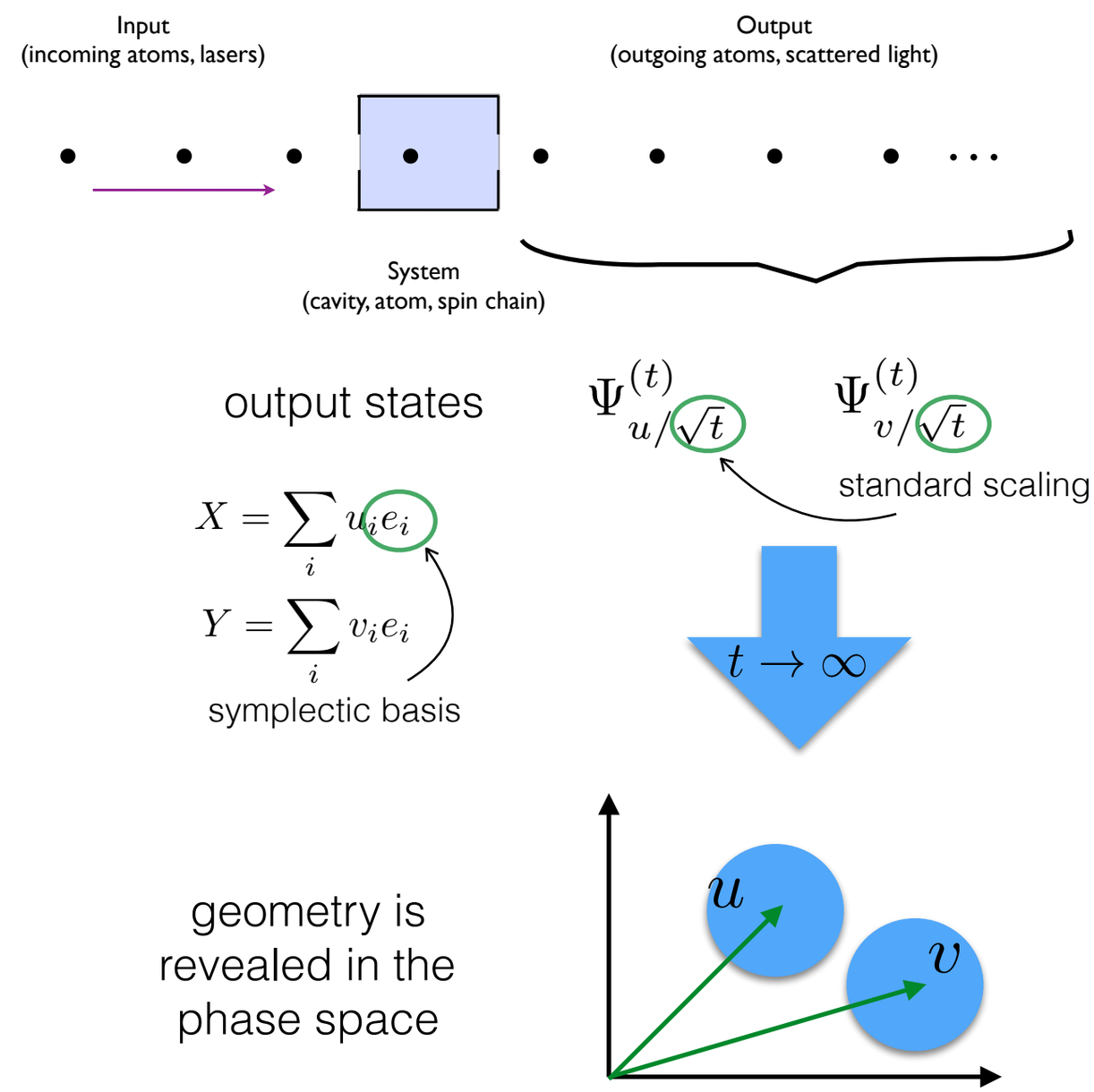
$$\langle \Omega | W(X) | \Omega \rangle = e^{-\langle X, X \rangle_p / 2}$$

- quantum version of a Gaussian statistical model — coherent states:

$$|X\rangle = W(X)|\Omega\rangle$$

- **Theorem (Local Asymptotic Normality)**. The output quantum model reduces to this Gaussian model at $t \to \infty$
  - simple and tractable model
  - <u>information is encapsulated in the symplectic geometry</u>

$$\langle X, Y \rangle_p \qquad \Sigma(X,Y)_p$$



Conclusion

- **Context**: inference of continual quantum measurement processes
- **Mathematical structure**:
  - *Statistical inference* (models, estimators)
  - *Riemannian geometry* — Fisher information
  - *principal connection* — “gauge” group for unidentifiable parameters
  - *symplectic structure* — quantum inner product
- **Result**: Model reduction to quantum Gaussian states (linear inference)