INFORMATION THEORY

Classical compression

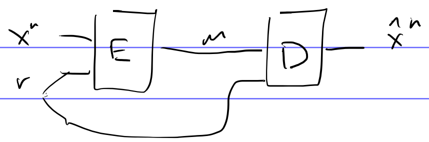
$H(p) = H(X)_p = -\sum_x p(x) \log p(x)$

$T^n_{p,\delta} = \{ x^n = (x_1, \dots, x_n) : \left| \frac{-1}{n} \log p^{\otimes n}(x^n) - H(x) \right| \leq \delta \}$

$\epsilon = 1 - p^{\otimes n}(T^n_{p,\delta}) \leq n^{O(1)} 2^{-n\delta} \to 0$ as $n \to \infty$

converse

- compress to k bits
- derandomize



pick r to maximize $\Pr[\hat{x}^n = x^n \mid r]$

- $S \subseteq \Sigma^n$ = set decoded correctly

$|S| \leq 2^k$

$p^n(S) \leq p^n(T^n_{p,\delta} \cap S) + p^n(\overline{T^n_{p,\delta}})$

$\leq 2^k \, 2^{-nH(p) + n\delta} + \epsilon$

$\to 0$ if $\frac{k}{n} < H(p)$

Q compression

$$S(\rho) = H(\operatorname{eig}(\rho)) = -\operatorname{tr}[\rho \log \rho]$$

$$0 \leq S(\rho) \leq \log d$$

$$S(\rho) = 0 \iff \operatorname{eig}(\rho) = (1, 0, \dots, 0) \iff \rho = |\psi\rangle\langle\psi|$$

$$S(\rho) = \log d \iff \rho = I/d$$

$$S(X) = \rho_X$$

$$S(X|Y) = S(XY) - S(Y)$$ (this defn carries over)

can be negative

$$I(X:Y) = S(X) + S(Y) - S(XY) = S(X) - S(X|Y)$$

classically $I(X;Y)$ = channel capacity

quantum $I(A';B)$ $|\psi\rangle \langle^{A}_{\boxed{N} - B}$ = entanglement-assisted classical capacity typical subspaces & projectors

$$\rho = \sum_{x=1}^{d} \lambda_x |v_x\rangle\langle v_x| \qquad \rho^{\otimes n} = \sum_{x^n} \lambda_{x_1} \cdots \lambda_{x_n} |v_{x^n}\rangle\langle v_{x^n}|$$

$$\Pi^n_{\rho,\delta} = \sum_{x^n \in T^n_{\lambda,\delta}} |v_{x^n}\rangle\langle v_{x^n}| \qquad |v_{x^n}\rangle = |v_{x_1}\rangle \otimes \cdots \otimes |v_{x_n}\rangle$$

typ. proj. supp $\Pi^n_{\rho,\delta}$ is typ. subspace

$$\mathrm{tr}\, \rho^{\otimes n} \Pi^n_{\rho,\delta} = \sum_{x^n} \lambda_{x_1} \cdots \lambda_{x_n} 1_{x^n \in T^n_{\rho,\delta}} = \lambda^n (T^n_{\lambda,\delta})$$

<u>q compression?</u>

1) $\rho^{\otimes n} - \boxed{\mathcal{E}} - \boxed{\mathcal{D}} - \sigma \approx \rho^{\otimes n}$

2) $|v_{x^n}\rangle - \boxed{\mathcal{E}} - \boxed{\mathcal{D}} - \sigma \qquad \mathbb{E}_{x^n \sim \lambda^n} F(|v_{x^n}\rangle\langle v_{x^n}|, \sigma) \approx 1$

3) $\rho = \sum_i p_i |w_i\rangle\langle w_i| \qquad |w_{i^n}\rangle - \boxed{\mathcal{E}} - \boxed{\mathcal{D}} - \sigma \qquad \mathbb{E}_{i^n \sim p^n} F(|w_{i^n}\rangle, \sigma) \approx 1$

4) $|\phi_\rho^{AR}\rangle^{\otimes n}$ — R ———— ; A — $\boxed{\mathcal{E}} - \boxed{\mathcal{D}}$ — $\} \sigma \qquad F(|\phi_\rho\rangle^{\otimes n}, \sigma) \approx 1$

$-\!\!/\!\!-$ means $\equiv$, i.e. n qubits

① is too weak   2–4 are all good, ≈ equiv

Schumacher – Jozsa compression

$$\{\Pi^n_{\rho,\delta},\ I - \Pi^n_{\rho,\delta}\} \qquad \mathrm{tr}\, \Pi^n_{\rho,\delta} \le \exp(n(S(\rho) + \delta))$$

alg. coding $\qquad \rho = \dfrac{I + \epsilon Z}{2}$

$$F\left(\sum_i p_i |i\rangle\langle i| \otimes \alpha_i, \sum_i p_i |i\rangle\langle i| \otimes \beta_i\right) = \left\| \left(\sum_i \sqrt{p_i} |i\rangle\langle i| \otimes \sqrt{\alpha_i}\right)\left(\sum_j \sqrt{p_j} |j\rangle\langle j| \otimes \sqrt{\beta_j}\right) \right\|_1 = \sum_i p_i \|\sqrt{\alpha_i}\sqrt{\beta_i}\|_1$$

④ → ③ → ②

4→3 measure R and get ensemble.

3→2 pick eigenbasis

$$F\left(\sum_{x^n} |x^n\rangle\langle x^n| \otimes \lambda_{x^n} |v_{x^n}\rangle\langle v_{x^n}|, \sum_{x^n} |x^n\rangle\langle x^n| \otimes \lambda_{x^n} \mathcal{D}(\mathcal{E}(|v_{x^n}\rangle\langle v_{x^n}|))\right) = \sum_{x^n} \lambda_{x^n} F(|v_{x^n}\rangle, \mathcal{D}(\mathcal{E}(|v_{x^n}\rangle)))$$

$$\sum_{x^n} \lambda_{x^n} \langle v_{x^n}| \mathcal{D}(\mathcal{E}(|v_{x^n}\rangle\langle v_{x^n}|)) |v_{x^n}\rangle$$

Relative entropy

Surprise $(x) := \log \frac{1}{p(x)}$

Huffman coding $E(x)$ has length $\lceil \log \frac{1}{p(x)} \rceil$ possible with prefix-free encoding

e.g.

x	$p(x)$	$E(x)$
a	$1/2$	0
b	$1/4$	10
c	$1/8$	110
d	$1/8$	111

$\lceil \cdot \rceil \Rightarrow$ some inefficiency which $\to 0$ if we encode blocks of letters

$$H(p) = \mathbb{E}_x [\text{surprise}(x)] = \sum_x p(x) \log \frac{1}{p(x)}$$

wrong codebook

$$\mathbb{E}_{x\sim p} |\mathrm{Enc}_q(x)| = \sum_x p(x) \log \frac{1}{q(x)} \geq \sum_x p(x) \log \frac{1}{p(x)}$$

$$\text{excess} = D(p\|q) = \sum_x p(x) \log \frac{p(x)}{q(x)} \quad \text{relative entropy}$$

claim: $D(p\|q) \geq 0$

fact
$$1 + z \leq e^z \quad \forall z \in \mathbb{R} \qquad \text{Pf } f(z) = e^z - 1 - z \text{ is convex}$$
$$\log y \leq y - 1 \qquad f(0) = f'(0) = 0$$
$$\log \frac{1}{y} \geq 1 - y$$

$$D(p\|q) = \sum_x p(x) \log \frac{p(x)}{q(x)}$$
$$\geq \sum_x p(x) \left(1 - \frac{q(x)}{p(x)}\right) = \sum_x p(x) - q(x) = 0$$

$D = 0 \iff \frac{p(x)}{q(x)} = 1$ always $\iff p = q$ (later we'll make this robust)

not a metric, not symmetric, no Δ ineq, etc

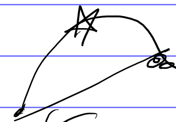
Cor $H(p) \leq \log d$

Pf $u = (1/d, \ldots, 1/d)$ $\quad 0 \leq D(p\|u) = \sum_x p(x) (\log p(x) + \log d)$
$$= \log d - H(p)$$

Cor $\quad I(X;Y) \geq 0$

$$D(p_{XY} \,\|\, p_X \otimes p_Y) = \sum_{x,y} p_{XY}(x,y)\left(\log p(x,y) - \log p_X(x) - \log p_Y(y)\right)$$
$$= -H(XY) + H(X) + H(Y)$$
$$= I(X;Y) \geq 0$$

Application $\quad$ concavity



$$\sum_x \pi_x H(p_x) \leq H\Big(\sum_x \pi_x p_x\Big)$$

$p(x,y) = \pi_x p_x(y) \qquad H(Y|X) \leq H(Y)$

$$H(Y) - H(Y|X) = I(X;Y) \geq 0$$

Hypothesis testing

$x \sim p$ or q

$\alpha = \Pr[\text{guess } q \mid x \sim p] \qquad$ type 1

$\beta = \Pr[\text{guess } p \mid x \sim q] \qquad$ type 2

symmetric $\qquad \min \frac{\alpha+\beta}{2} = \frac{1}{2}\|p-q\|_1$

Bayesian $\qquad \min \pi\alpha + (1-\pi)\beta = \|\alpha p - (1-\pi) q\|_1$

asymmetric $\qquad \beta_\epsilon = \min\{\beta : \alpha \leq \epsilon\}$

we will study $\quad \beta_\epsilon^n = \beta_\epsilon$ for p^n vs $q^n \sim \exp(-nR)$

Chernoff - Stein

$$\lim_{n\to\infty} \frac{-1}{n} \log \beta_\epsilon^n = D(p\|q) \qquad \forall \epsilon \in (0,1)$$

Examples

1) $p = q \Leftrightarrow D(p\|q) = 0 \quad$ i.e. exp. small error always possible

2) $q = U \qquad D(p\|U) = \log d - H(p)$

$T^n_{p,\delta}$ guess $p \qquad \overline{T^n_{p,\delta}}$ guess U

$$\alpha = p^n\left(\overline{T^n_{p,\delta}}\right) \to 0 \qquad \text{as} \quad n \to \infty$$

$$\beta = u^n\left(T^n_{p,\delta}\right) = \frac{|T^n_{p,\delta}|}{d^n} \leq \exp\left(nH(p) + n\delta - n\log d\right)$$
$$= \exp\left(-n\left(D(p\|u) - \delta\right)\right)$$

3) $D(p\|q) = \infty \iff \exists x$ s.t. $p(x) > 0$, $q(x) = 0$

$\iff \operatorname{supp} p \not\subseteq \operatorname{supp} q$

If we see $x \in \operatorname{supp} p - \operatorname{supp} q$, guess p, otherwise q

$$\alpha = p(\operatorname{supp} q)^n \to 0 \qquad \beta = 0$$

Pf sketch of Chernoff-Stein thm

LRT (likelihood ratio test)

$X^n \sim p^n$ or q^n $\qquad W(x^n) = \log \frac{p^n(x^n)}{q^n(x^n)}$

$\mathbb{E}_{x^n \sim p^n} W = n D(p\|q)$ $\qquad \mathbb{E}_{x^n \sim q^n} W = -n D(q\|p)$

$A = \{x^n : W \geq n(D(p\|q) - \delta)\}$ $\qquad x^n \in A \Leftrightarrow q^n(x^n) \leq e^{-n(D-\delta)} p^n(x^n)$

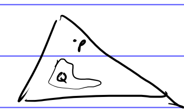
$q^n(A) \leq e^{-n(D-\delta)} p^n(A) \leq e^{-n(D-\delta)}$

$p^n(A) \to 1$ by law of large numbers

multiple hypothesis testing

$x^n \sim p^n$ or q^n for $q \in Q$

rate $= \min_{q \in Q} D(p\|q)$



$\boxed{\text{quantum relative entropy}}$

$D(\rho\|\sigma) = \mathrm{tr}[\rho(\log\rho - \log\sigma)]$

$\beta_\epsilon^n = \min\{\mathrm{tr}\, M\sigma^{\otimes n} : \mathrm{tr}\, M\rho^{\otimes n} \geq 1 - \epsilon\}$

q Stein $\qquad \lim_{n\to\infty} -\frac{1}{n}\log\beta_\epsilon^n = D(\rho\|\sigma)$

$D(\rho\|\sigma) = \infty \Leftrightarrow \mathrm{supp}\,\rho \not\subseteq \mathrm{supp}\,\sigma$

e.g. $\rho = |\psi\rangle\langle\psi|$ $\qquad \sigma = |\phi\rangle\langle\phi|$

$A = |\psi^\perp\rangle\langle\psi^\perp|$ $\qquad B = I - A$ $\qquad M = A \otimes B^{\otimes n-1} + B \otimes A \otimes B^{\otimes n-2} + \dots + B^{\otimes n-1} \otimes A$

Application to entropies

$D(\rho\|\sigma) \geq 0$ $\qquad S(A) \leq \log d$

$I(A;B) \geq 0$ $\qquad S(A) \geq S(A|B)$

Pf of q Stein

want M s.t. $\operatorname{tr} \rho^{\otimes n} M \geq \alpha$

$\operatorname{tr} \sigma^{\otimes n} M \lesssim e^{-nR}$ $\quad R \approx D(\rho\|\sigma)$

optimal $M = [\alpha^{-1} \rho^{\otimes n} - 2^{nR} \sigma^{\otimes n} \geq 0]$

following Bjelakovic et al.

$\rho = \sum_x r_x |\alpha_x\rangle\langle\alpha_x|$ $\quad \sigma = \sum_x s_x |\beta_x\rangle\langle\beta_x|$

$\Pi^n_{\rho\|\sigma,\delta} = \sum_{x^n: \left|\frac{1}{n}\sum_{i=1}^n \log s_{x_i} - \operatorname{tr}\rho\log\sigma\right| \leq \delta} \beta_{x^n}$ $\quad \beta_{x^n} = \beta_{x_1} \otimes \cdots \otimes \beta_{x_n}$

cf. $\Pi^n_{\rho,\delta} = \sum_{x^n \text{ s.t. } \left|\frac{1}{n}\sum_{i=1}^n -\log r_{x_i} - S(\rho)\right| \leq \delta} |\alpha_{x^n}\rangle\langle\alpha_{x^n}|$

$\operatorname{tr} \rho^{\otimes n} \Pi_{\rho\|\sigma} \geq 1-\epsilon$ $\quad$ law of large numbers

$[\Pi, \sigma^{\otimes n}] = 0$

$2^{n(\operatorname{tr}\rho\log\sigma - \delta)} \Pi \leq \Pi\sigma^{\otimes n}\Pi \leq 2^{n(\operatorname{tr}\rho\log\sigma + \delta)} \Pi$

Achievability: $M = \Pi^n_{\rho\|\sigma,\delta} \Pi^n_{\rho,\delta} \Pi^n_{\rho\|\sigma,\delta}$

$\operatorname{tr} \rho^{\otimes n} M \to 1$ using gentle measurement lemma

$\operatorname{tr} \sigma^{\otimes n} M \leq \operatorname{tr} \Pi^n_{\rho,\delta} 2^{n(\operatorname{tr}\rho\log\sigma + \delta)} \Pi^n_{\rho\|\sigma,\delta} \leq 2^{n(S(\rho)+\delta)} 2^{n(\operatorname{tr}\rho\log\sigma+\delta)} = \exp(-n(D(\rho\|\sigma) - 2\delta))$

Converse

suppose $\operatorname{tr} \rho^{\otimes n} M \geq \alpha$

$\sigma^{\otimes n} \geq \Pi_{\rho\|\sigma} 2^{n(\operatorname{tr}\rho\log\sigma - \delta)}$ $\quad \rho^{\otimes n} \Pi^n_{\rho,\delta} \leq 2^{-n(S(\rho)-\delta)} \Pi^n_{\rho,\delta}$

$\operatorname{tr} M\sigma^{\otimes n} \geq \operatorname{tr}(M\Pi_{\rho\|\sigma}) 2^{n(\operatorname{tr}\rho\log\sigma - \delta)}$

$\operatorname{tr} M\Pi_{\rho\|\sigma} = \operatorname{tr} \Pi_{\rho\|\sigma} M \Pi_{\rho\|\sigma}$

$\geq \operatorname{tr} \Pi_{\rho\|\sigma} M \Pi_{\rho\|\sigma} \Pi_\rho$

$\geq \underbrace{\operatorname{tr} \Pi_{\rho\|\sigma} M \Pi_{\rho\|\sigma} \Pi_\rho \rho^{\otimes n}}\, 2^{-n(S(\rho)-\delta)}$ $\quad \rho^{\otimes n} = \overbrace{\Pi_\rho \rho^{\otimes n}}^{A} + \overbrace{(I-\Pi_\rho)\rho^{\otimes n}}^{B}$

$\operatorname{tr} M\left(\Pi_{\rho\|\sigma} \left(\rho^{\otimes n} - \underbrace{(I-\Pi_\rho)\rho^{\otimes n}}_{\operatorname{tr} \leq \epsilon}\right) \Pi_{\rho\|\sigma}\right)$

$\geq (\alpha - 2\sqrt{\epsilon} - \epsilon)$

$\operatorname{tr} M\sigma^{\otimes n} \geq (\alpha - 2\sqrt{\epsilon} - \epsilon)\, 2^{-n(D(\rho\|\sigma)+2\delta)}$

Cor $D(\rho\|\sigma) \geq D(\mathcal{E}(\rho)\|\mathcal{E}(\sigma))$

Cor strong subadditivity

$I(A:C|B) = I(A:BC) - I(A:B) = D(\rho_{ABC}\|\rho_A\otimes\rho_{BC}) - D(\rho_{AB}\|\rho_A\otimes\rho_B) \geq 0$

monotonicity under tr_C

Applications of information theory to physics

given H

ρ has energy $\mathrm{tr}[\rho H]$

free energy $F(\rho) = \mathrm{tr}[\rho H] - T S(\rho)$

minimized by $\sigma = \dfrac{e^{-H/T}}{\mathrm{tr}\, e^{-H/T}}$

This is robust. $D(\rho\|\sigma) = F(\rho) - F(\sigma)$ (Pinsker's ineq)

so $F(\rho) \le F(\sigma) + \epsilon \Leftrightarrow D(\rho\|\sigma) \le \epsilon \Rightarrow \|\rho - \sigma\|_1^2 \le \sqrt{2\epsilon}$

Thermalization process

$$\dot{\rho} = \mathcal{L}(\rho) \qquad \mathcal{L}(\sigma) = 0$$

MLSI $\quad \frac{d}{dt} D(\rho\|\sigma) \le -2\alpha\, D(\rho\|\sigma)$

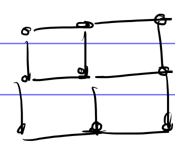
compare with spectral gap $\lambda = \min -\mathrm{Re}\,\mu$, $\mathcal{L}(X) = \mu X$

$\Rightarrow$ Var decays as $\exp(-\lambda t)$

Structure of Gibbs state

classically $p(x) = \dfrac{e^{-E(x)/T}}{Z}$

$E(x) = \sum_S h_s(x_S)$ e.g.



$p(x)$ is Markovian

i.e. $p(x_A, x_B, x_C) = p(x_B)\, p(x_A|x_B)\, p(x_C|x_B)$

$\Leftrightarrow I(A;C|B)_p = 0$

quantumly

$I(A;C|B) = 0 \Leftrightarrow \rho = \exp(\log \rho^{AB} + \log \rho^{BC} - \log \rho^{B})$

$\Leftrightarrow B \cong \bigoplus_\alpha B_\alpha^L \otimes B_\alpha^R$ $\quad \rho = \bigoplus_\alpha p_\alpha\, \sigma_\alpha^{AB_\alpha^L} \otimes \omega_\alpha^{B_\alpha^R C}$

^ classical variable

$\Leftrightarrow \exists R_{B\to BC}$ s.t. $\rho_{ABC} = (\mathrm{id}_A \otimes R_{B\to BC})(\rho_{AB})$

CMI is nonzero for Gibbs states but small

small CMI $\Rightarrow$ approx recovery maps

$$H = \sum_{S \subset V} h_S$$

How does locality of H affect locality of e^{-iHt}, etc..

<u>Lieb – Robinson bound</u>

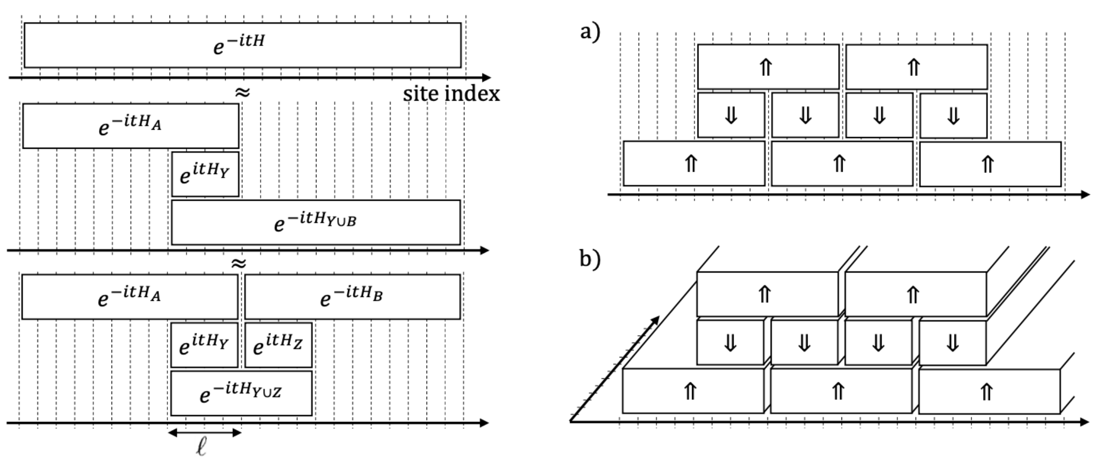
$\mu > 0 \qquad \alpha := \max_i \sum_{S \ni i} \|h_S\| \cdot |S| \, e^{\mu \, \mathrm{diam}(S)}$

A_X, B_Y supported on $X, Y \subset V$

$$\| [A_X(t), B_Y] \| \leq 2 \, \|A_X\| \, \|B_Y\| \, |X| \exp(-\mu \, \mathrm{dist}(X,Y)) \left(e^{2\alpha |t|} - 1\right)$$

$$\exp\Big(-\mu\big(\mathrm{dist} - \underbrace{\tfrac{2\alpha}{\mu}}_{V_{LR}} t\big)\Big) \Rightarrow \text{approx light-cone with speed } V_{LR}$$

<u>Application</u> Simulating lattice Hamiltonians [Haah, Hastings, Kothari, Low '18]



$$\| U_t^{A\cup B} (U_t^B)^\dagger U_t^{B\cup C} - U_t^{A\cup B\cup C} \| \leq O(e^{-\mu \, \mathrm{dist}(A,C)}) \sum_{S \text{ on bdy}(AB:C)} \|h_S\|$$

$$W_t := e^{it(H_{AB} + H_C)} e^{-itH_{ABC}} = (U_t^{AB+C})^\dagger U_t^{ABC}$$

$$\partial_t W_t = U_t^{AB+C} \underbrace{(H_{ABC} - H_{AB} - H_C)}_{H_{bd}} (U^{AB+C})^\dagger W_t$$

$$\approx U_t^{B+C} H_{bd} (U_t^{B+C})^\dagger W_t$$

$\exp(-\mu \, \mathrm{dist}(A,C))$

Application: correlation decay in gapped ground states

$|\psi_0\rangle = |\psi_{gs}\rangle$

Suppose H has gap g, i.e. energies $E_0 < E_1 = E_0 + g \le E_2 \le \dots$

then $|\langle A_x B_y\rangle - \langle A_x\rangle\langle B_y\rangle| \lesssim e^{-\frac{\ell g}{2v_{LR}}} \|A\| \|B\|$

example TFIM $\quad H = -J \sum_i Z_i Z_{i+1} - B \sum_i X_i$

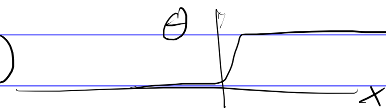
$B \gg 0$ unique g.s. $|+\rangle^{\otimes N} = |\to\to\to\to \cdots \to\rangle$ $\quad$ no correlations

$|B| \ll J$ $\quad \frac{|0\rangle^{\otimes N} \pm |1\rangle^{\otimes N}}{\sqrt{2}}$ have exp. small gap $\quad$ long-range correlations $\langle Z_i Z_j\rangle = 1$

Pf of correlation decay

Assume WLOG that $\langle A\rangle = \langle B\rangle = 0$ $\quad E_0 = 0$

$\langle E_i | B^+ | E_j\rangle := \langle E_i | B | E_j\rangle \, \theta(E_i - E_j)$



$B^+|\psi_0\rangle = B|\psi_0\rangle \qquad \langle\psi_0| B^+ = 0$

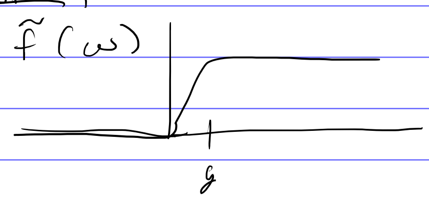
$\langle AB\rangle = \langle AB^+\rangle = \langle [A, B^+]\rangle$

but B^+ is nonlocal

$$\int f(t)\, B(t)\, dt = \sum_{i,j} |E_i\rangle\langle E_j|\, B_{ij} \int dt\, f(t)\, e^{it(E_i - E_j)}$$

$$= \sum_{i,j} |E_i\rangle\langle E_j|\, B_{ij}\, \tilde{f}(E_i - E_j)$$

goals for f



$\tilde{f}(\omega) \approx 0 \quad \omega < 0$

$\tilde{f}(\omega) \approx 1 \quad \omega > g/2$

say erf

$|f(t)|$ small if $t \gtrsim 1/g$

then $B^+ \approx \int dt\, f(t)\, B(t)$

$$[A, B^+] \approx \int dt\, f(t)\, [A, B(t)]$$

small if $t \gtrsim 1/g$ $\qquad \lesssim e^{-\mu \ell}$ if $t \le 1/g$

balancing terms $\Rightarrow$ error $\exp\left(\frac{-\ell g}{2 v_{LR}}\right)$