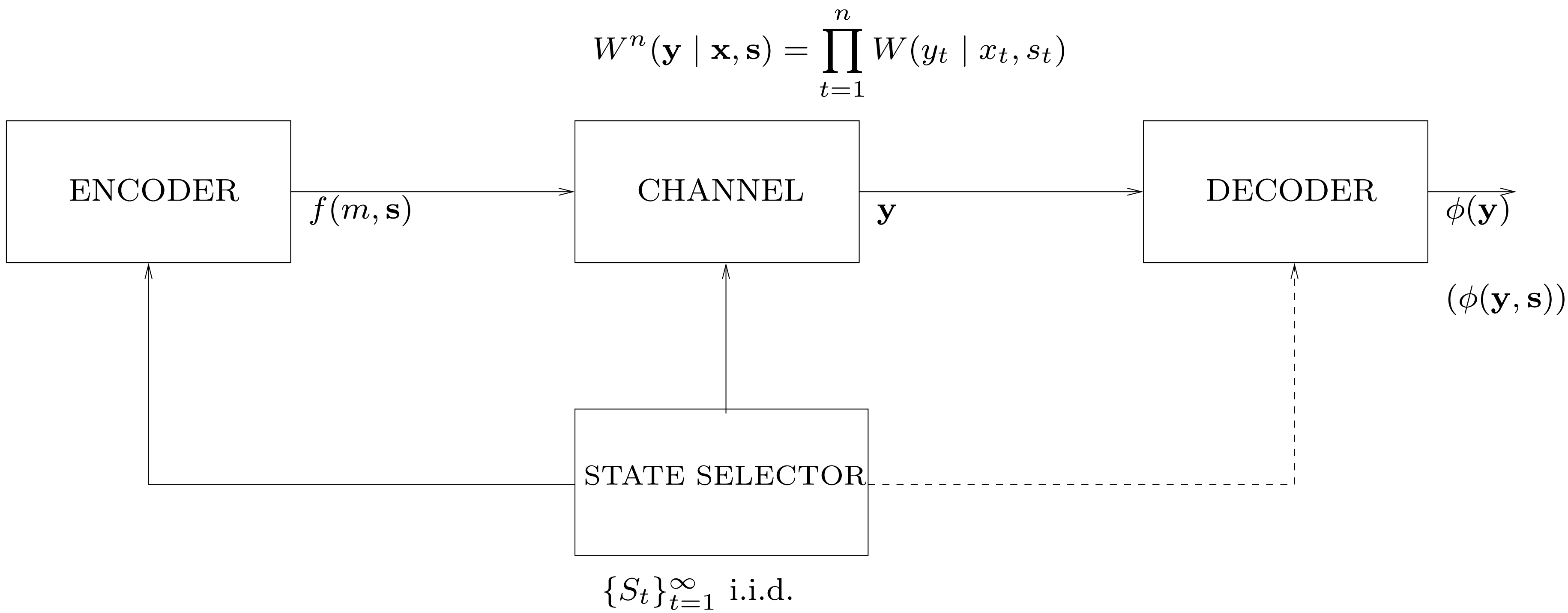
The Gelfand-Pinsker Channel: Strong Converse and Upper Bound for the Reliability Function

Himanshu Tyagi Prakash Narayan

The Gelfand-Pinsker Channel Model



$$e(f, \phi) = \max_{m \in \mathcal{M}} \sum_{\mathbf{s} \in \mathcal{S}^n} \mathtt{P}_S(\mathbf{s}) W^n((\phi^{-1}(m))^c \mid f(m, \mathbf{s}), \mathbf{s})$$

where

$$\phi^{-1}(m) = \{\mathbf{y} \in \mathcal{Y}^n : \phi(\mathbf{y}) = m\}.$$

Capacity

Receiver with no CSI [Gelfand-Pinsker, '80]:

$$C_{\mathrm{GP}} = \max_{\mathrm{P}_{USXY}} I(U \wedge Y) - I(U \wedge S)$$

where

$$U -\!\circ\!- S, X -\!\circ\!- Y, \qquad P_{Y|X,S} = W.$$

Receiver with full CSI [Wolfowitz, '60]:

$$C = \max_{\mathrm{P}_{X|S}} I(X \wedge Y \mid S).$$

Reliability Function

Definition: The *reliability function* $E(R)$, $R \geq 0$, of the DMC W with noncausal CSI, is the largest number $E \geq 0$ such that for every $\delta > 0$ and for all sufficiently large n, there exist n-length block codes (f, ϕ) of rate greater than $R - \delta$ and $e(f, \phi) \leq \exp\left[-n(E - \delta)\right]$.

Prior Results and New Contributions

Prior Results:

- Somekh Baruch-Merhav, '04, Moulin-Wang, '04: Lower bounds on $E(R)$ for Gelfand-Pinsker channel.
- Shannon-Gallagher-Berlekamp, '67: Upper bounds for $E(R)$ for DMC without states.
- Csiszár-Körner-Marton, '77: Alternative proof of upper bounds using strong converse for fixed type codewords.
- Wolfowitz, '60: Strong converse for DMC with states, causal transmitter CSI and no receiver CSI. An analog for noncausal CSI was not available.
- Haroutunian, '01: Upper bound for $E(R)$ for noncausal transmitter CSI and full receiver CSI.

Contributions:

- Strong converse for the Gelfand-Pinsker channel.
- Upper bound for $E(R)$ for the Gelfand-Pinsker channel.

Key Idea for Strong Converse

- Upper bound for the rate of codes with codewords that are conditionally typical over large *message dependent* subsets of a typical set of state sequences.
- **Note:** A direct extension of the Csiszár-Körner-Marton approach would have entailed a claim over a subset of typical state sequences *not depending* on the transmitted message; however, its validity is unclear.
- For a DMC without states, the Csiszár-Körner-Marton approach provides an image size characterization of a good codeword set. In the same spirit, our key technical lemma provides an image size characterization for good codeword sets for the noncausal DMC model, which now involves auxiliary rvs.

Results

Theorem: *(Strong Converse)* Given $0 < \epsilon < 1$ and a sequence of (M_n, n) codes (f_n, ϕ_n) with $e(f_n, \phi_n) < \epsilon$, it holds that

$$\limsup_n \frac{1}{n} \log M_n \leq C_{\mathrm{GP}}.$$

Theorem: (*Sphere Packing Bound*) For $0 < R < C_{\mathrm{GP}}$, it holds that

$$E(R) \leq E_{SP}(R),$$

where

$$E_{SP}(R) = \min_{\tilde{\mathsf{P}}_S} \max_{\tilde{\mathsf{P}}_{X|S}} \min_{V \in \mathcal{V}(R, \tilde{\mathsf{P}}_S \tilde{\mathsf{P}}_{X|S})} \left[D(\tilde{\mathsf{P}}_S \| \mathsf{P}_S) + \ D(V \| W \mid \tilde{\mathsf{P}}_S \tilde{\mathsf{P}}_{X|S}) \right]$$

with

$$\mathcal{V}(R, \tilde{\mathsf{P}}_{SX}) = \left\{ V : \mathcal{X} \times \mathcal{S} \to \mathcal{Y} \text{ s.t. } \max_{\mathsf{P}_{USXY} = \mathsf{P}_{U|SX} \tilde{\mathsf{P}}_{SX} V} I(U \wedge Y) - I(U \wedge S) < R \right\}.$$

Remark 1. For the case when the receiver, too, possesses (full) CSI, the sphere packing bound above coincides with that obtained earlier in [Haroutunian, '01].

Remark 2. The terms $D(\tilde{\mathsf{P}}_S \| \mathsf{P}_S)$ and $D(V \| W \mid \tilde{\mathsf{P}}_S \tilde{\mathsf{P}}_{X|S})$, respectively, account for the shortcomings of a given code for the corresponding “bad” state pmf and “bad” channel.

Results

Technical Lemma: Given a pmf $\tilde{\mathsf{P}}_S$ on S and conditional pmf $\tilde{\mathsf{P}}_{X|S}$, let (f, ϕ) be a (M, n)-code. For each $m \in \mathcal{M}$, let $A(m)$ be a subset of $\mathcal{S}^n$ which satisfies the following conditions

$$\begin{aligned} A(m) &\subset \mathcal{T}^n_{[\tilde{\mathsf{P}}_S]}, \\ \frac{1}{n} \log \|A(m)\| &\cong H(\tilde{\mathsf{P}}_S), \\ f(m, \mathbf{s}) &\in \mathcal{T}^n_{[\tilde{\mathsf{P}}_{X|S}]}(\mathbf{s}), \qquad \mathbf{s} \in A(m). \end{aligned}$$

Furthermore, let (f, ϕ) satisfy one of the following two conditions:

$$\begin{aligned} W^n(\phi^{-1}(m) \mid f(m, \mathbf{s}), \mathbf{s}) &\cong 1, \qquad \mathbf{s} \in A(m), \\ \frac{1}{\|A(m)\|} \sum_{\mathbf{s} \in A(m)} W^n(\phi^{-1}(m) \mid f(m, \mathbf{s}), \mathbf{s}) &\cong 1. \end{aligned}$$

Then for all n sufficiently large,

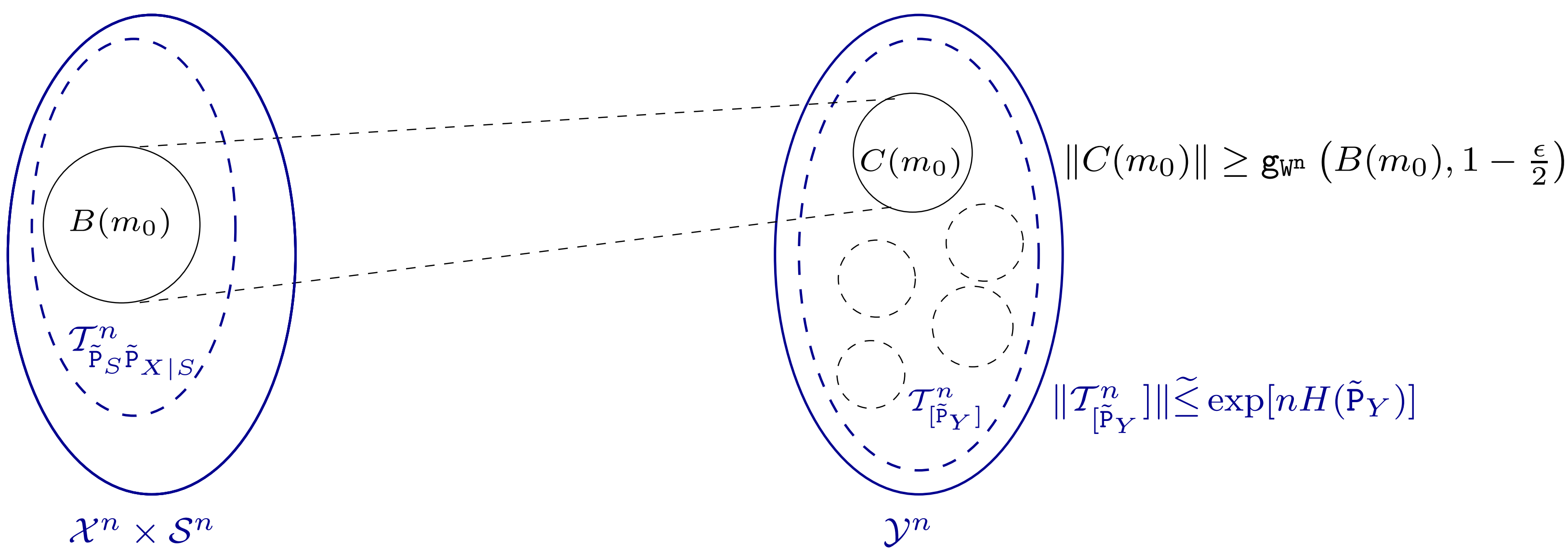
$$\frac{1}{n} \log M \leq I(U \wedge Y) - I(U \wedge S)$$

where $\mathsf{P}_{USXY} = \mathsf{P}_{U|SX} \tilde{\mathsf{P}}_S \tilde{\mathsf{P}}_{X|S} W$.

Observe: The subsets $A(m)$ of $\mathcal{S}^n$ are message-dependent.

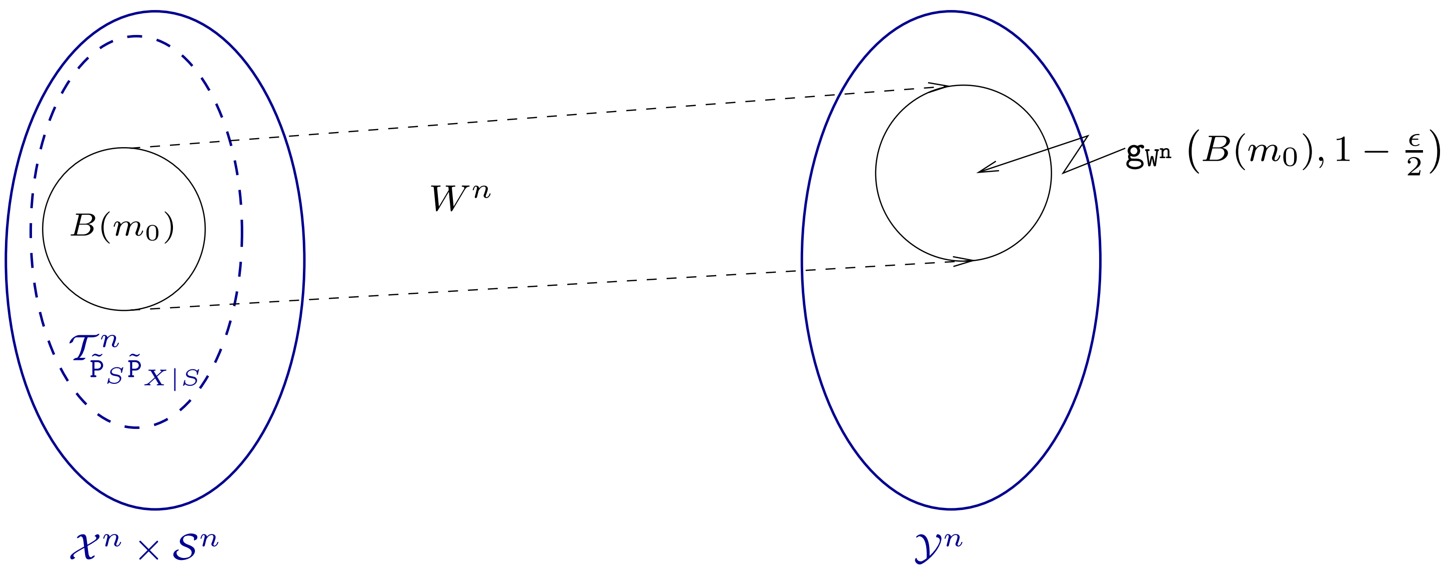
Outline of Proof of Technical Lemma

- $W^n(\phi^{-1}(m) \mid f(m,\mathbf{s}),\mathbf{s}) \geq 1-\epsilon, \qquad \mathbf{s} \in A(m)$
- $B(m) \triangleq \{(f(m,\mathbf{s}),\mathbf{s}) \in \mathcal{X}^n \times \mathcal{S}^n : \mathbf{s} \in A(m)\}, \qquad m \in \mathcal{M}$
- $C(m) \triangleq \phi^{-1}(m) \cap \mathcal{T}^n_{[\tilde{\mathtt{P}}_Y]} \Rightarrow W^n(C(m) \mid f(m,\mathbf{s}),\mathbf{s}) \geq 1-\frac{\epsilon}{2}, \qquad \mathbf{s} \in A(m)$
- $m_0 = \mathtt{arg}\min_m \|C(m)\|$.

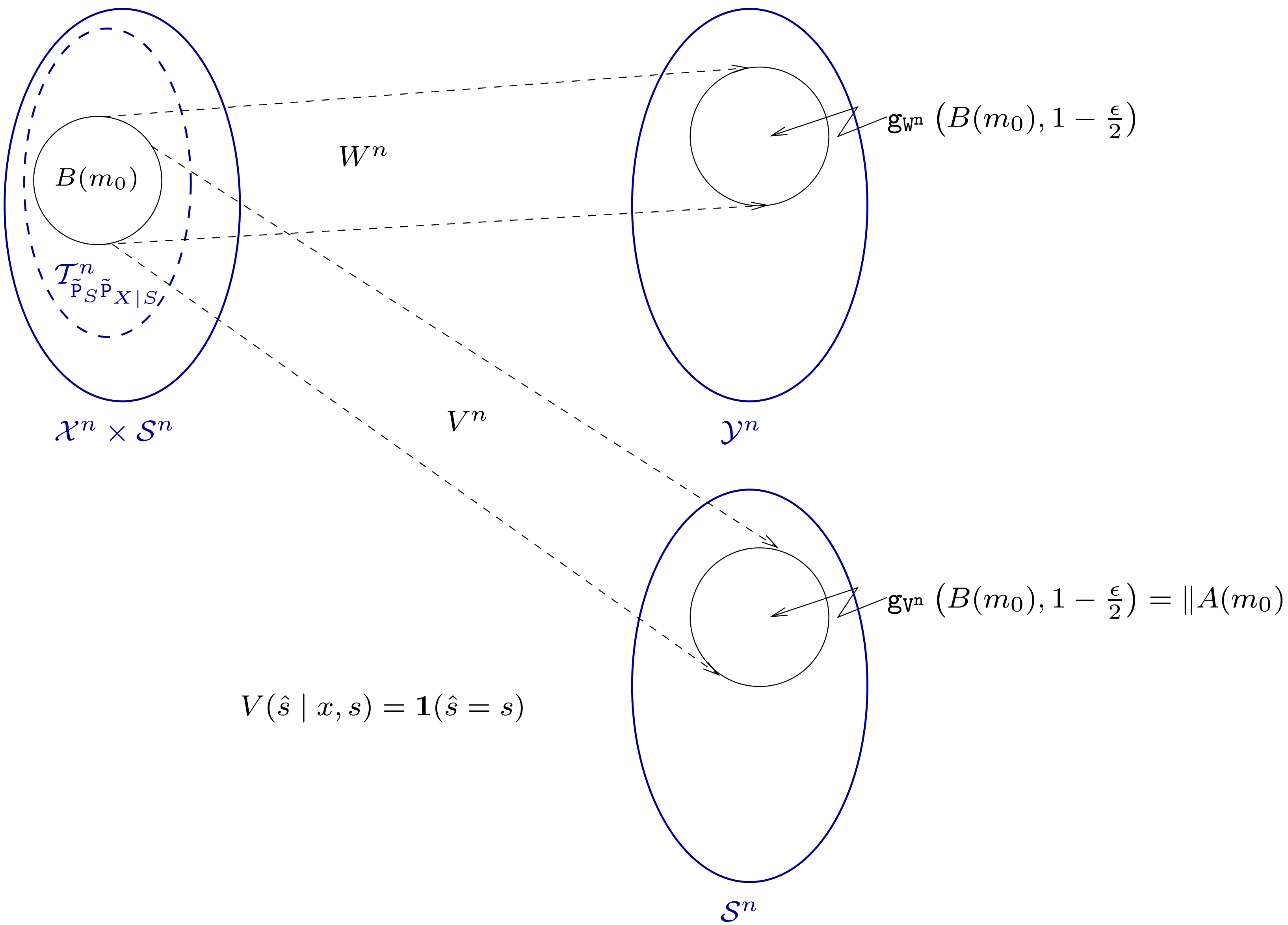


$$\Rightarrow \frac{1}{n}\log M \;\; \widetilde{\leq} \;\; H(\tilde{\mathtt{P}}_Y) - \frac{1}{n}\log \mathtt{g}_{\mathtt{W}^n}\left(B(m_0), 1-\frac{\epsilon}{2}\right)$$

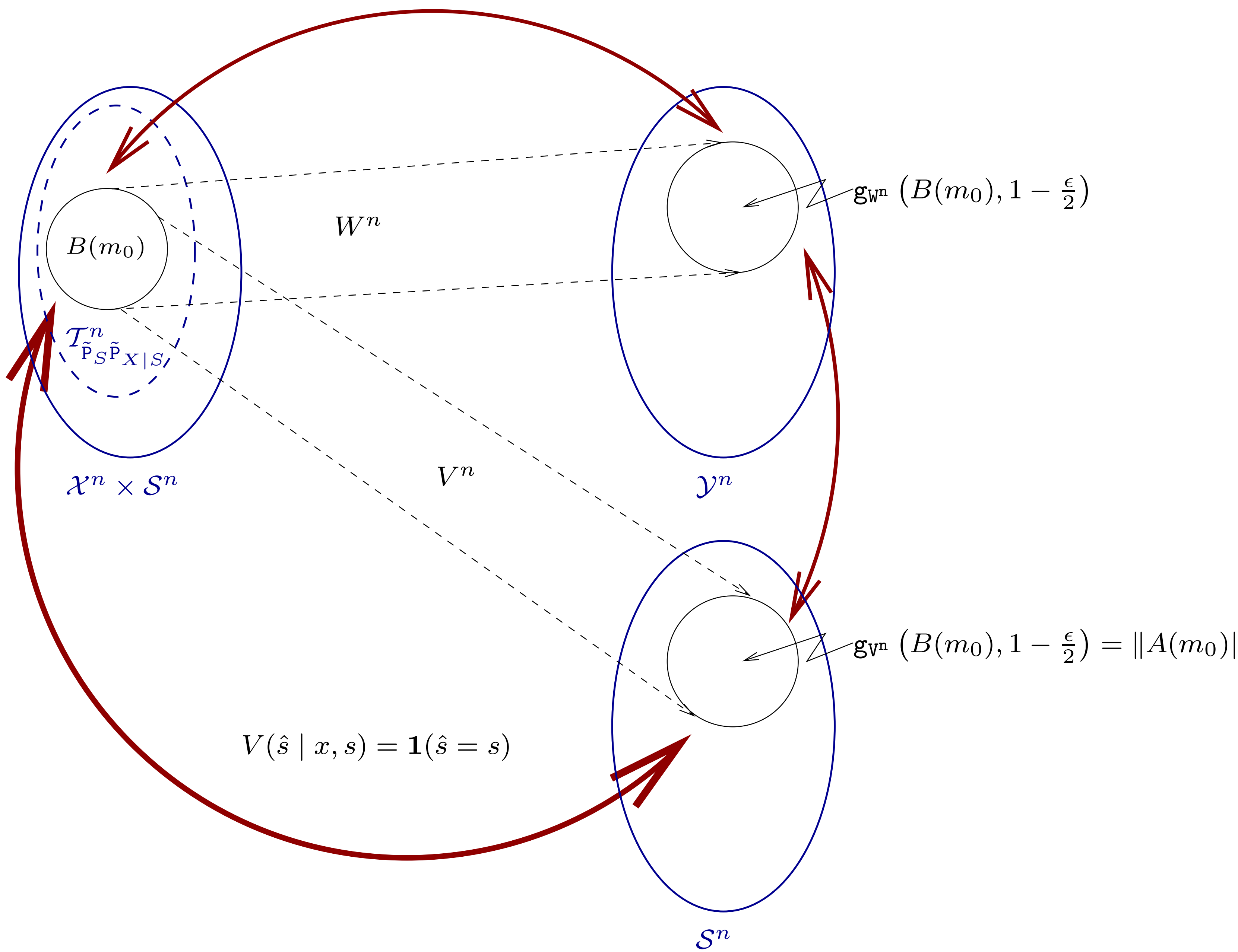
Outline of Proof of Technical Lemma



Outline of Proof of Technical Lemma



Outline of Proof of Technical Lemma



Outline of Proof of Technical Lemma

Image-size characterization

- $\frac{1}{n} \log \mathtt{g}_{\mathtt{V}^n}(B(m_0), 1 - \epsilon/2) \cong H(S|U) + t$
- $\frac{1}{n} \log \mathtt{g}_{\mathtt{W}^n}(B(m_0), 1 - \epsilon/2) \cong H(Y|U) + t$

where $0 \leq t \leq \min\{I(U \wedge Y),\, I(U \wedge S)\}$.

$$\Rightarrow \frac{1}{n} \log M \;\; \tilde{\leq} \;\; I(U \wedge Y) - I(U \wedge S).$$

Outline of Proof of Strong Converse

- Fix $0 < \epsilon < 1$.
- Given a (M, n)-code (f, ϕ) with $e(f, \phi) \leq \epsilon$.
- *Extraction of subsets of $\mathcal{T}^n_{[\mathtt{P}_S]}$ with "good code behavior:"*

$$\mathtt{P}_S\Bigg(\underbrace{\left\{\mathbf{s} \in \mathcal{T}^n_{[\mathtt{P}_S]} : W^n(\phi^{-1}(m) \mid f(m, \mathbf{s}), \mathbf{s}) > \frac{1-\epsilon}{2}\right\}}_{\hat{A}(m)}\Bigg) \geq \frac{1-\epsilon}{3}.$$

- *Extraction of sets $A(m)$ from $\hat{A}(m)$*:
  - Partition $\hat{A}(m)$ into (polynomially many) conditional types of $f(m, \mathbf{s})$ given $\mathbf{s}$; take the largest cell to be $A(m)$.
  - $A(m)$ satisfies all the conditions of the Lemma.
- By the Lemma,

$$\frac{1}{n} \log M \;\; \tilde{\leq} \;\; I(U \wedge Y) - I(U \wedge S).$$