Generalizing Bottleneck Problems

International Symposium on Information Theory (ISIT)

June 18, 2018

Hsiang Hsu [*], Shahab Asoodeh[†], Salman Salamatian[‡], and Flavio P. Calmon[*]

[*]Harvard University, {hsianghsu, fcalmon}@g.harvard.edu

[†]University of Chicago , shahab@uchicago.edu,

[‡]MIT, salmansa@mit.edu,

HARVARD John A. Paulson School of Engineering and Applied Sciences

MIT EECS ELECTRICAL ENGINEERING AND COMPUTER SCIENCE

THE UNIVERSITY OF CHICAGO

Outline

1. Two bottleneck problems
   - Information Bottleneck and Privacy Funnel

2. Generalizing bottleneck problems
   - Motivation and Formulation

3. Geometric properties of bottleneck problems
   - Witsenhausen and Wyner
   - How to solve generalizing bottleneck problems?

4. Applications
   - Mrs. and Mr. Gerber's Lemma
   - Arimoto's Mrs. and Mr. Gerber's Lemma
   - Estimation Bottleneck and Estimation Privacy Funnel

5. Final remarks

An extended version can be accessed at https://arxiv.org/abs/1802.05861.

Outline

1. Two bottleneck problems
   - Information Bottleneck and Privacy Funnel

2. Generalizing bottleneck problems
   - Motivation and Formulation

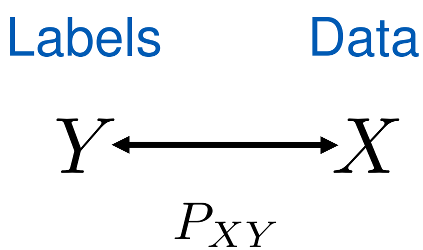
3. Geometric properties of bottleneck problems
   - Witsenhausen and Wyner
   - How to solve generalizing bottleneck problems?

4. Applications
   - Mrs. and Mr. Gerber's Lemma
   - Arimoto's Mrs. and Mr. Gerber's Lemma
   - Estimation Bottleneck and Estimation Privacy Funnel

5. Final remarks

The Information Bottleneck[1]

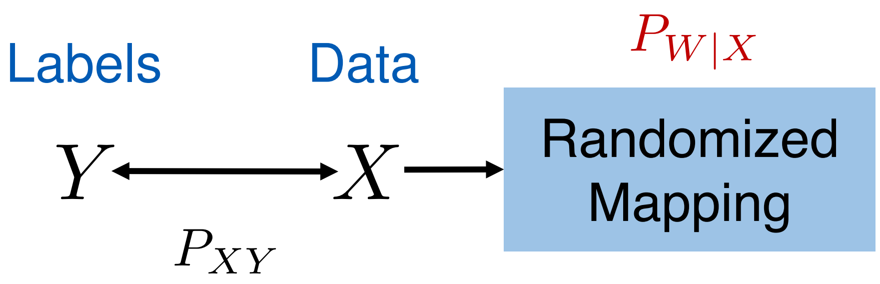
- Clustering[2], NLP[3], understanding deep learning[4,5]
- Two correlated random variables, X and Y, of finite cardinality, and P_{XY}
  - X: Noisy MNIST images, CIFAR-100 pictures
  - Y: Digits, categories



[1]N. Tishby et al.'00, [2]N. Tishby et al.'01, [3]N. Slonim et al.'00, [4]N. Tishby et al.'15, [5]R. Shwartz-Ziv et al.'17

The Information Bottleneck[1]

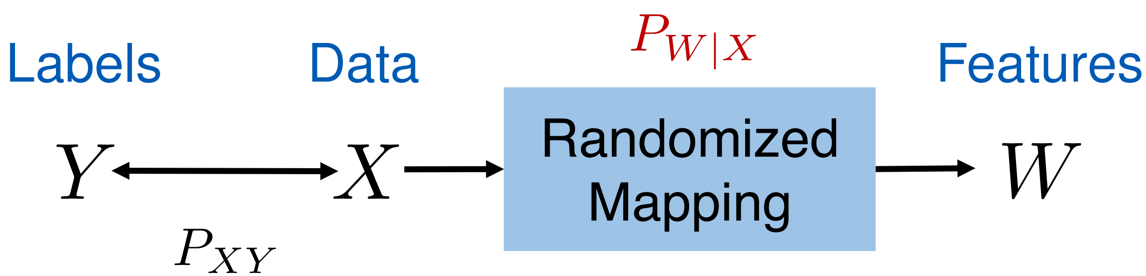
- Clustering[2], NLP[3], understanding deep learning[4,5]
- Two correlated random variables, X and Y, of finite cardinality, and P_{XY}
  - X: Noisy MNIST images, CIFAR-100 pictures
  - Y: Digits, categories



[1]N. Tishby et al.'00, [2]N. Tishby et al.'01, [3]N. Slonim et al.'00, [4]N. Tishby et al.'15, [5]R. Shwartz-Ziv et al.'17

The Information Bottleneck[1]

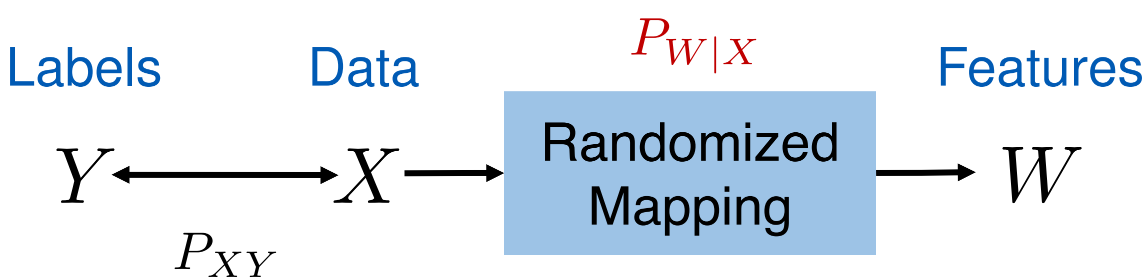
- Clustering[2], NLP[3], understanding deep learning[4,5]
- Two correlated random variables, X and Y, of finite cardinality, and P_{XY}
  - X: Noisy MNIST images, CIFAR-100 pictures
  - Y: Digits, categories



[1]N. Tishby et al.'00, [2]N. Tishby et al.'01, [3]N. Slonim et al.'00, [4]N. Tishby et al.'15, [5]R. Shwartz-Ziv et al.'17

The Information Bottleneck[1]

- Clustering[2], NLP[3], understanding deep learning[4][5]
- Two correlated random variables, X and Y, of finite cardinality, and P_{XY}
  - X: Noisy MNIST images, CIFAR-100 pictures
  - Y: Digits, categories



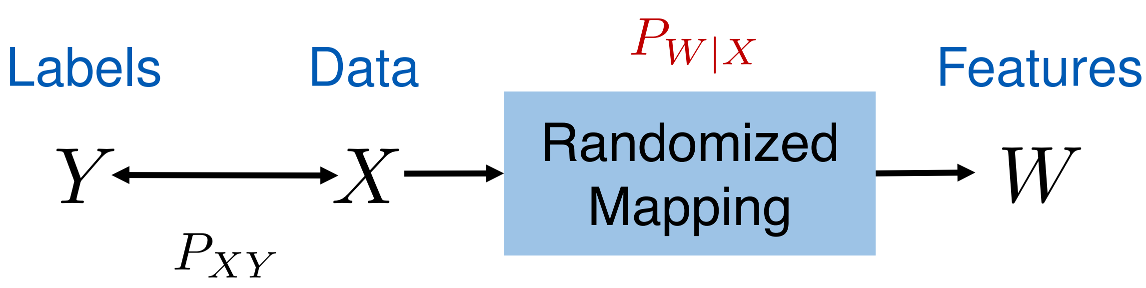
$$\max_{P_{W|X}} I(W;Y)$$

$$s.t. \ I(W;X) \leq x$$

[1]N. Tishby et al.'00, [2]N. Tishby et al.'01, [3]N. Slonim et al.'00, [4]N. Tishby et al.'15, [5]R. Shwartz-Ziv et al.'17

The Information Bottleneck[1]

- Clustering[2], NLP[3], understanding deep learning[4,5]
- Two correlated random variables, X and Y, of finite cardinality, and P_{XY}
  - X: Noisy MNIST images, CIFAR-100 pictures
  - Y: Digits, categories



Information preserved

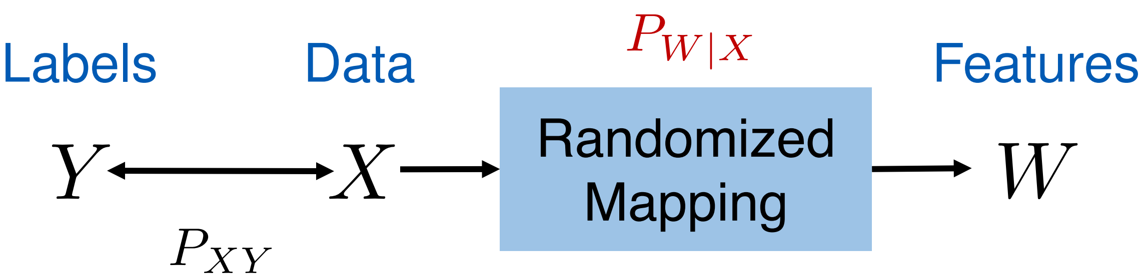
$$\max_{P_{W|X}} I(W;Y)$$

$$s.t.\ I(W;X) \leq x$$

[1]N. Tishby et al.'00, [2]N. Tishby et al.'01, [3]N. Slonim et al.'00, [4]N. Tishby et al.'15, [5]R. Shwartz-Ziv et al.'17

The Information Bottleneck[1]

- Clustering[2], NLP[3], understanding deep learning[4][5]
- Two correlated random variables, X and Y, of finite cardinality, and P_{XY}
  - X: Noisy MNIST images, CIFAR-100 pictures
  - Y: Digits, categories



Information preserved

$$\max_{P_{W|X}} I(W;Y)$$

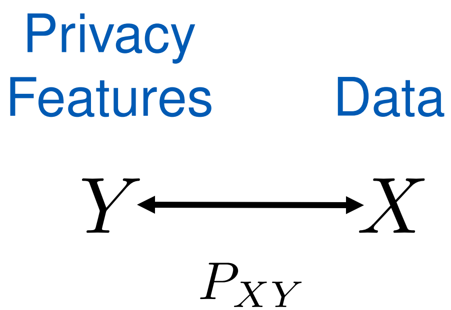
$$s.t.\ I(W;X) \leq x$$

Compression

[1]N. Tishby et al.'00, [2]N. Tishby et al.'01, [3]N. Slonim et al.'00, [4]N. Tishby et al.'15, [5]R. Shwartz-Ziv et al.'17

The Privacy Funnel[1,2]

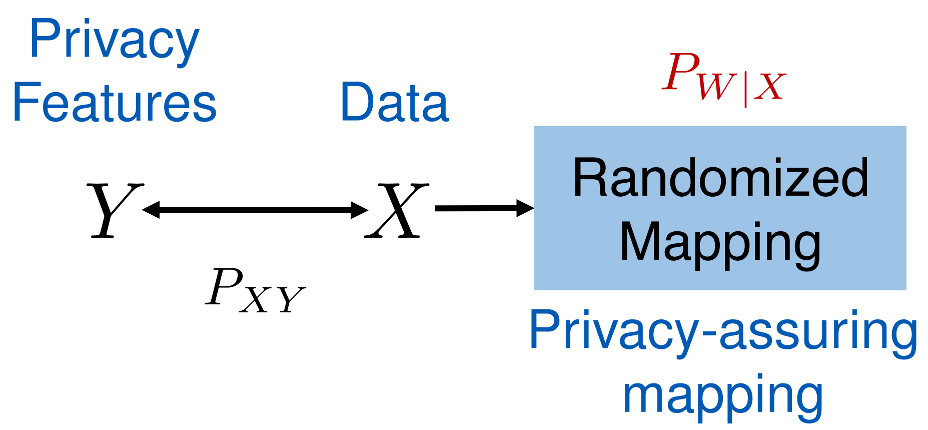
- A related optimization problem in information-theoretical privacy
  - X: Movie rating
  - Y: Political preference
  - Utility: Movie favor



[1]Ali et al.'14, [2]F. P. Calmon et al.'15.

The Privacy Funnel[1,2]

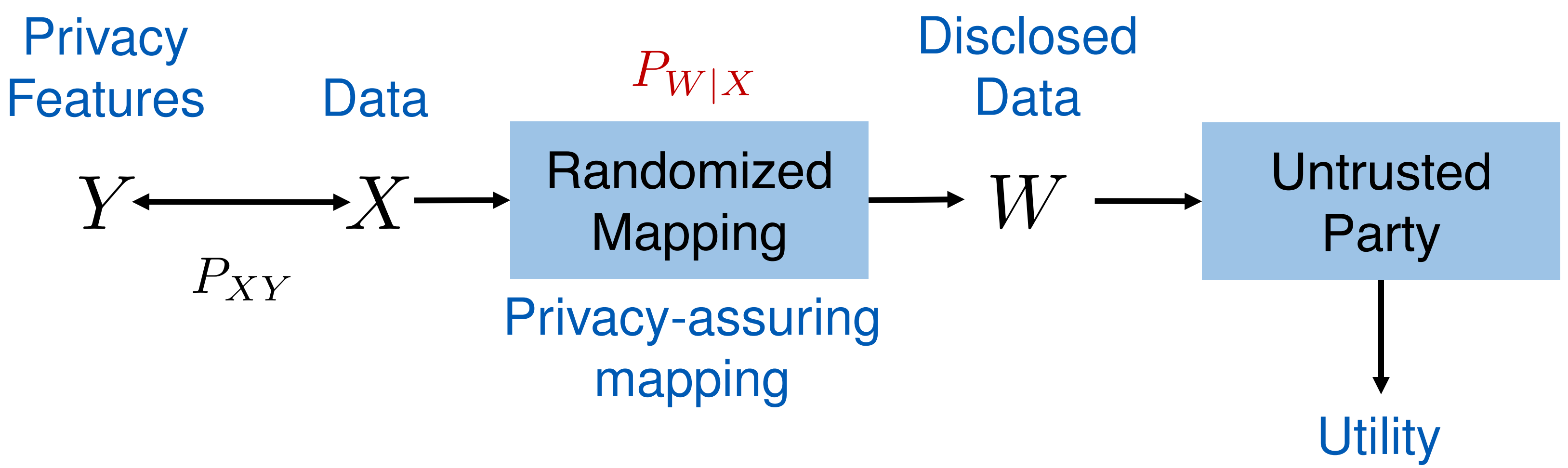
- A related optimization problem in information-theoretical privacy
  - X: Movie rating
  - Y: Political preference
  - Utility: Movie favor



[1]Ali et al.'14, [2]F. P. Calmon et al.'15.

The Privacy Funnel[1,2]

- A related optimization problem in information-theoretical privacy
  - X: Movie rating
  - Y: Political preference
  - Utility: Movie favor

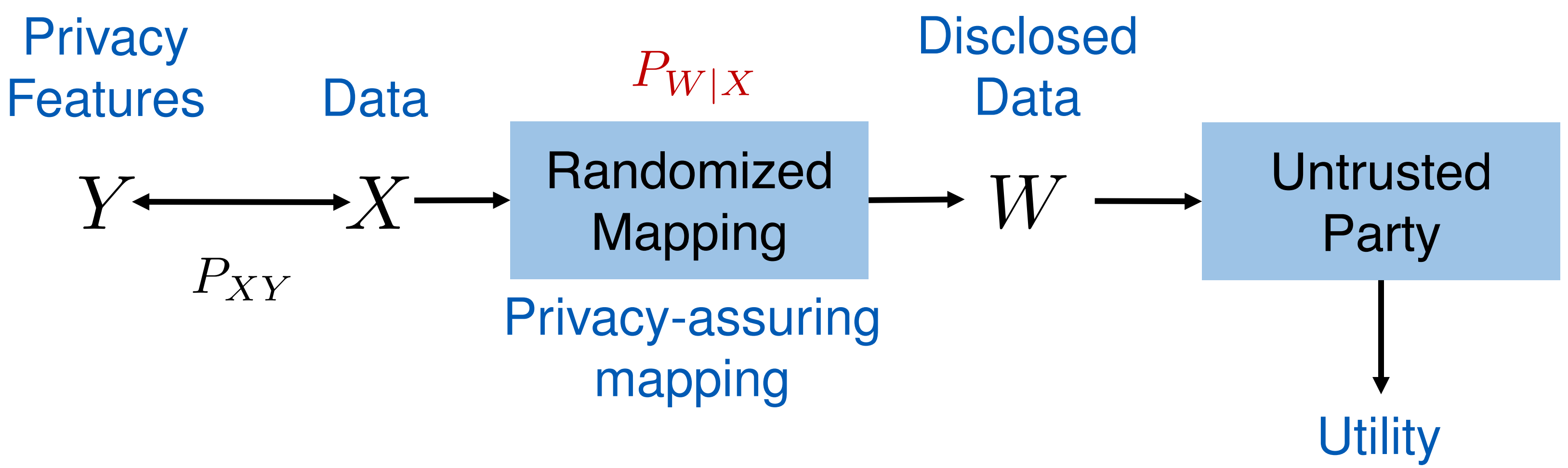


[1]Ali et al.'14, [2]F. P. Calmon et al.'15.

The Privacy Funnel[1,2]

- A related optimization problem in information-theoretical privacy
  - X: Movie rating
  - Y: Political preference
  - Utility: Movie favor

$$\min_{P_{W|X}} I(W;Y)$$
$$s.t.\ I(W;X) \geq x$$



[1]Ali et al.'14, [2]F. P. Calmon et al.'15.

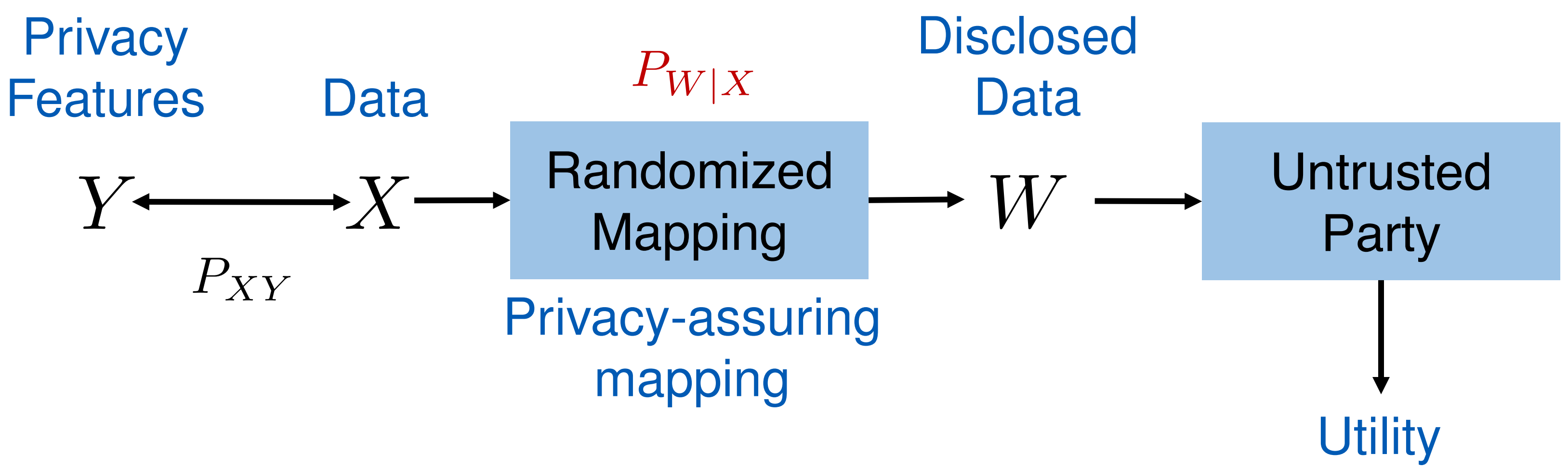
The Privacy Funnel[1,2]

- A related optimization problem in information-theoretical privacy
  - X: Movie rating
  - Y: Political preference
  - Utility: Movie favor

Privacy leakage

$$\min_{P_{W|X}} I(W;Y)$$

$$s.t.\ I(W;X) \geq x$$



[1]Ali et al.'14, [2]F. P. Calmon et al.'15.

The Privacy Funnel[1,2]

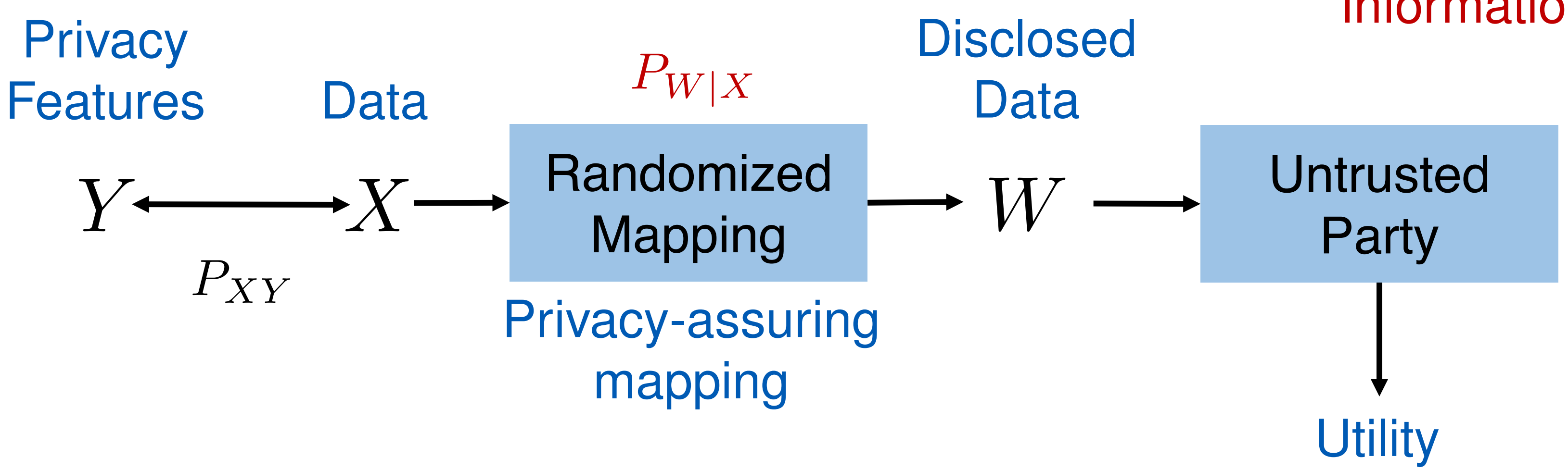
- A related optimization problem in information-theoretical privacy
  - X: Movie rating
  - Y: Political preference
  - Utility: Movie favor

Privacy leakage

$$\min_{P_{W|X}} I(W;Y)$$

$$s.t.\ I(W;X) \geq x$$

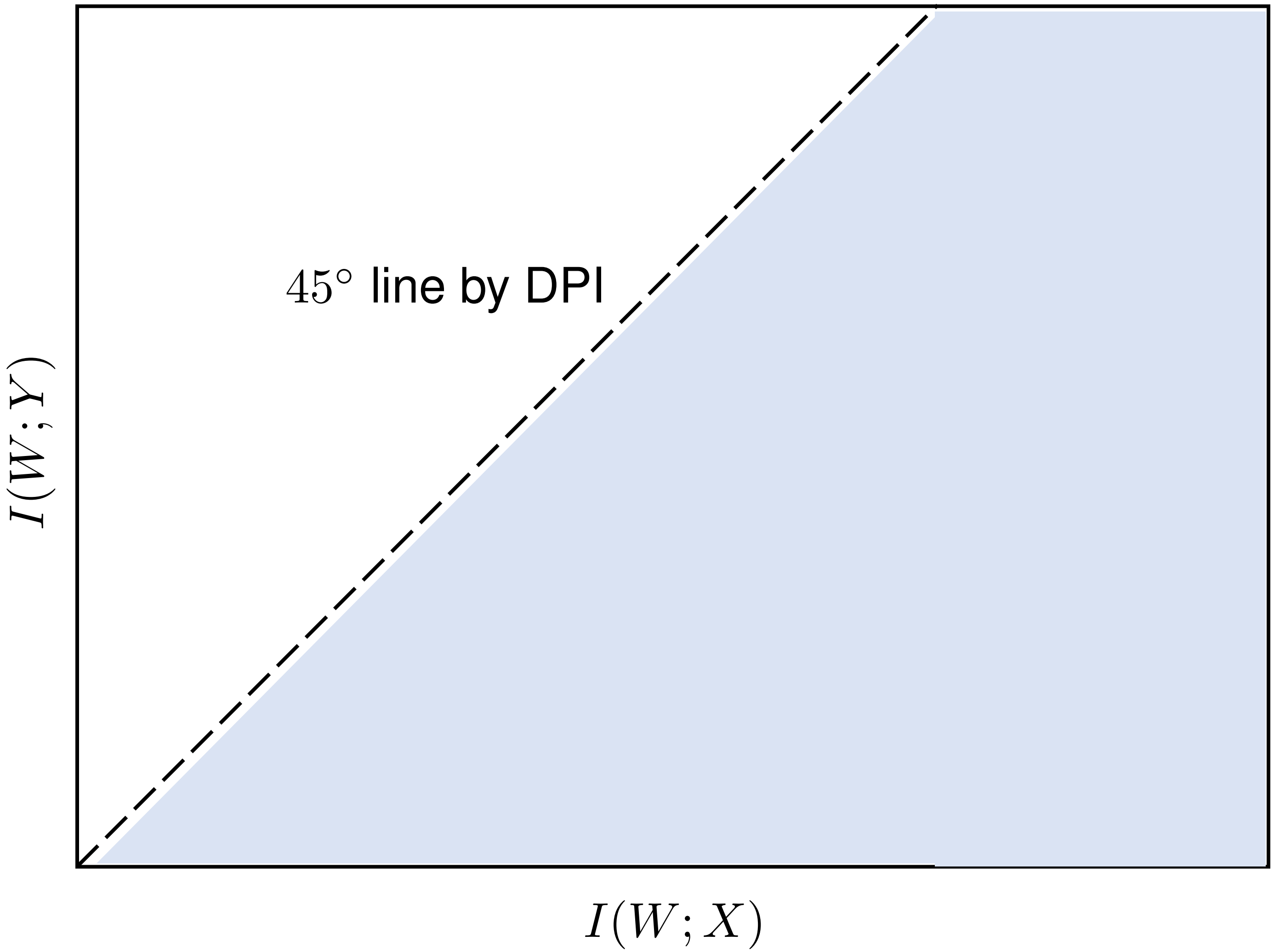
Preserve useful Information



[1]Ali et al.'14, [2]F. P. Calmon et al.'15.

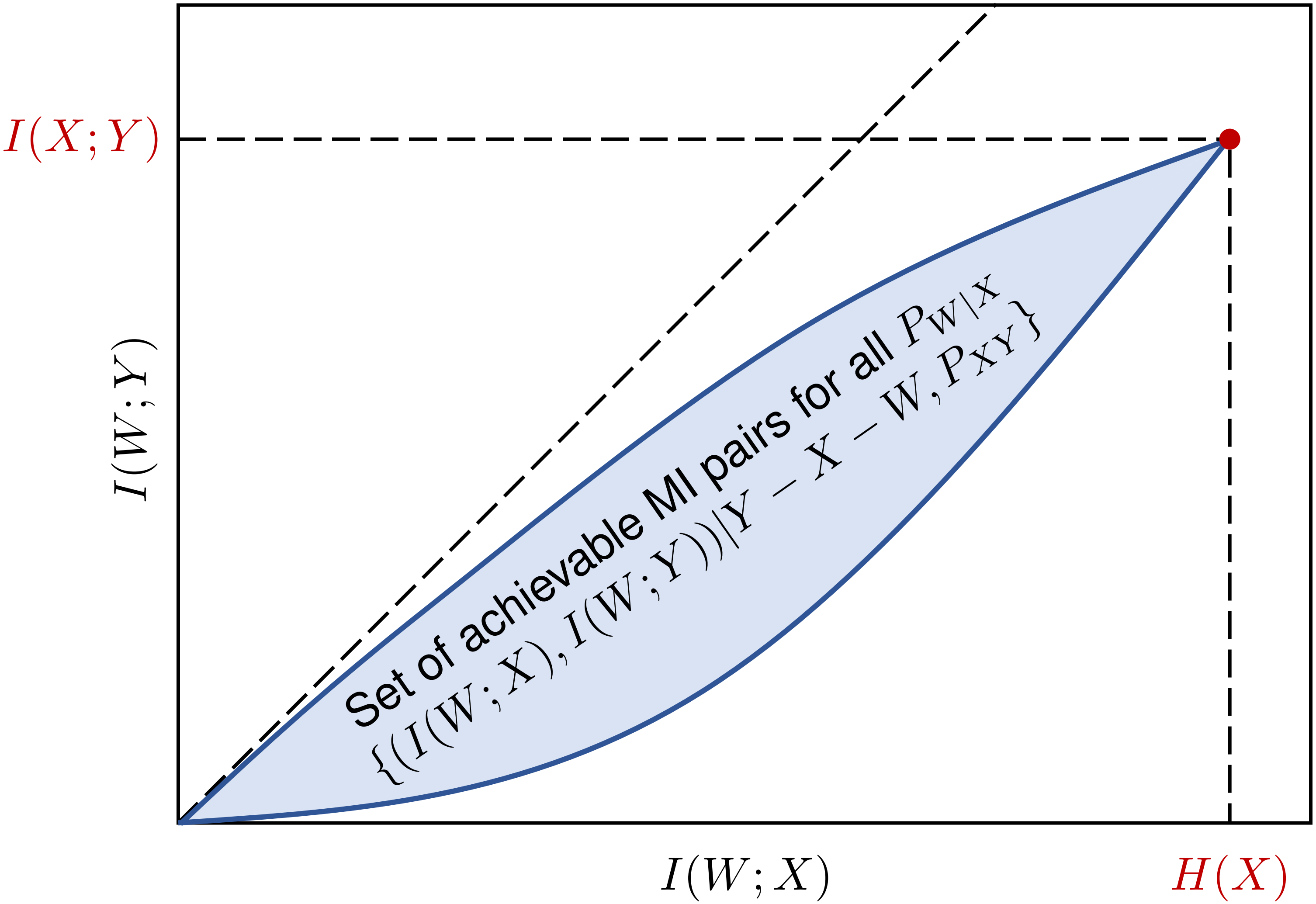
Bottleneck Problems

- Given $Y - X - W$, and a fixed P_{XY}



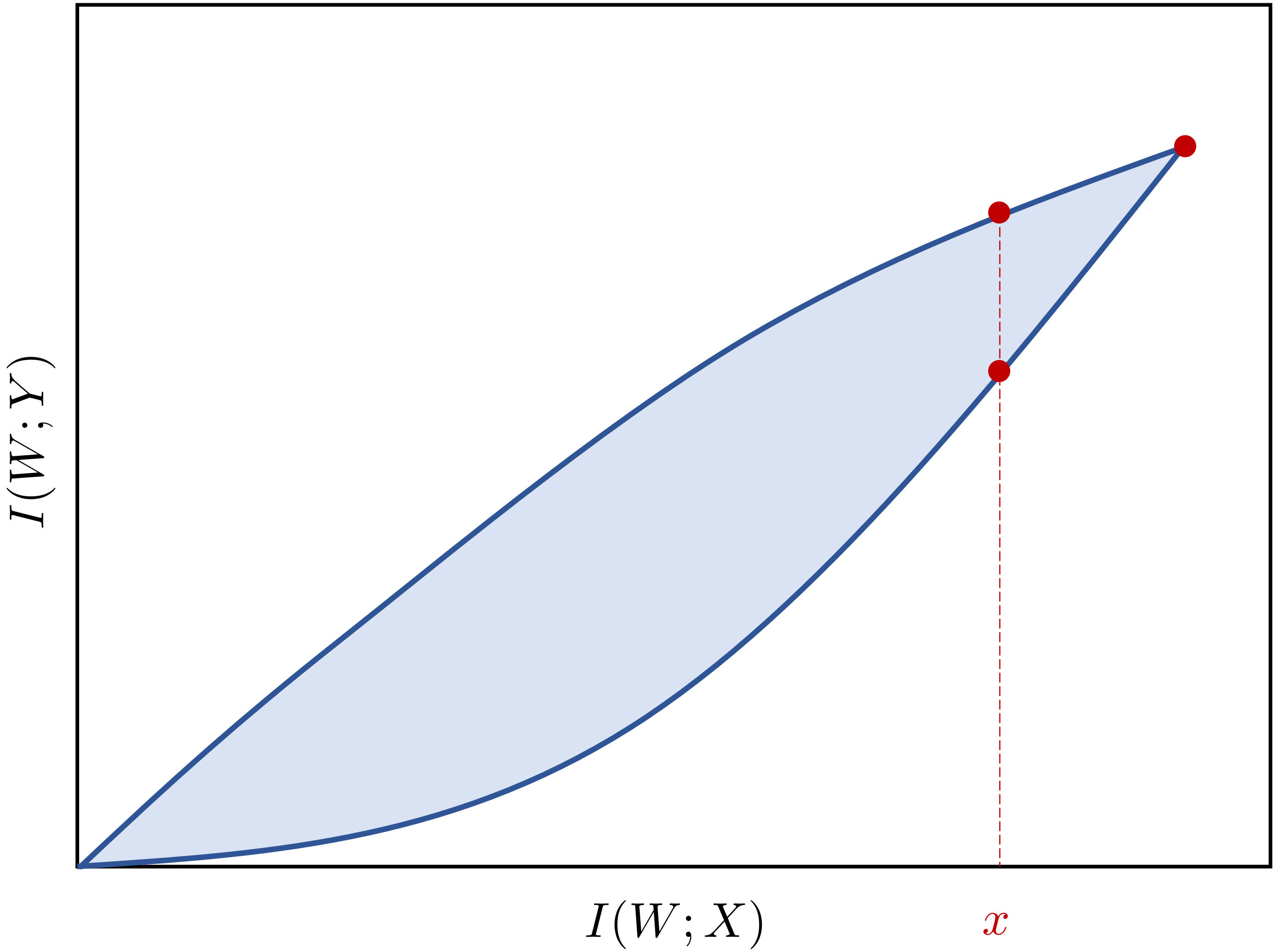
Bottleneck Problems

- Given $Y - X - W$, and a fixed P_{XY}



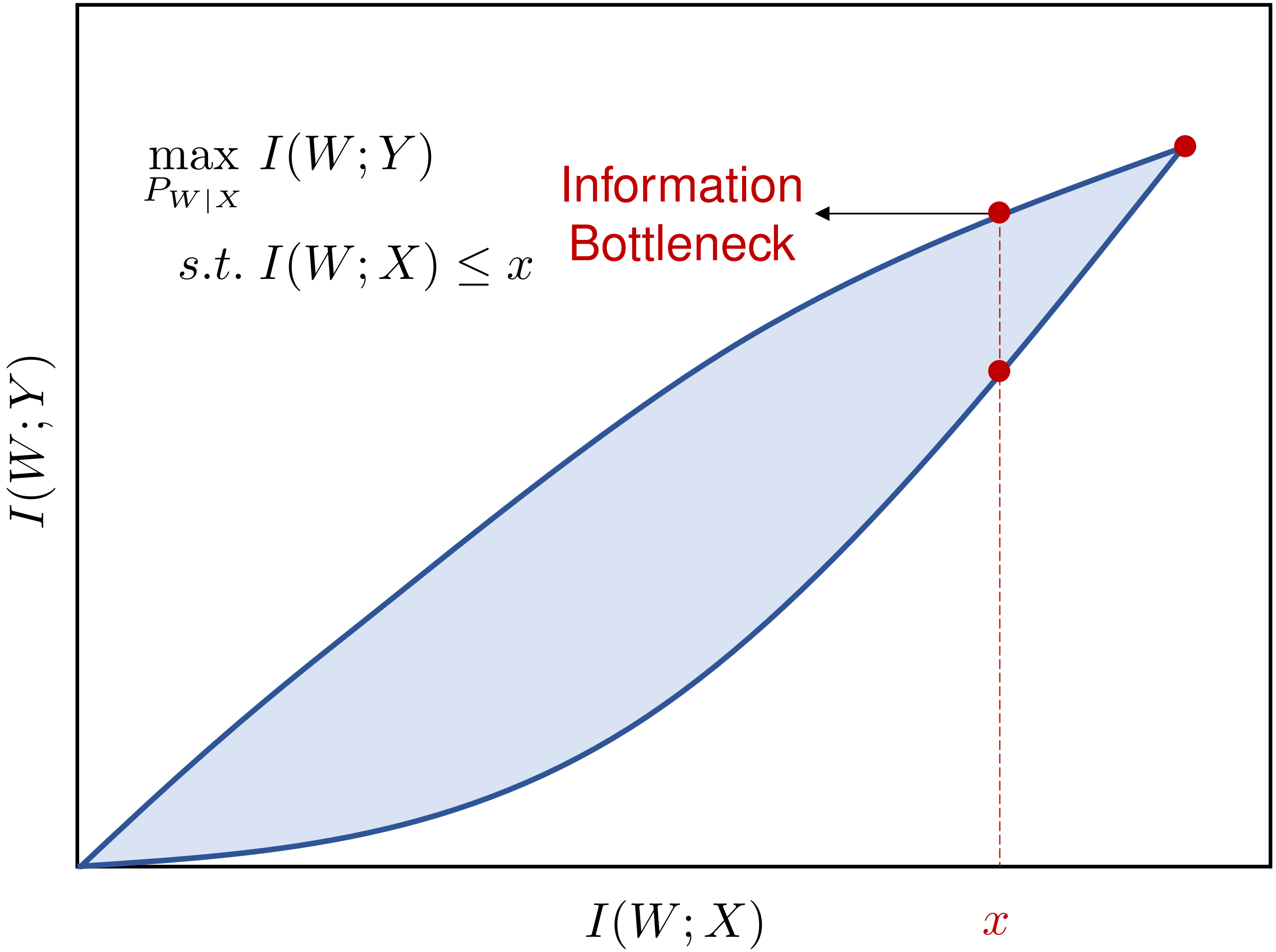
Bottleneck Problems

- Given $Y - X - W$, and a fixed P_{XY}



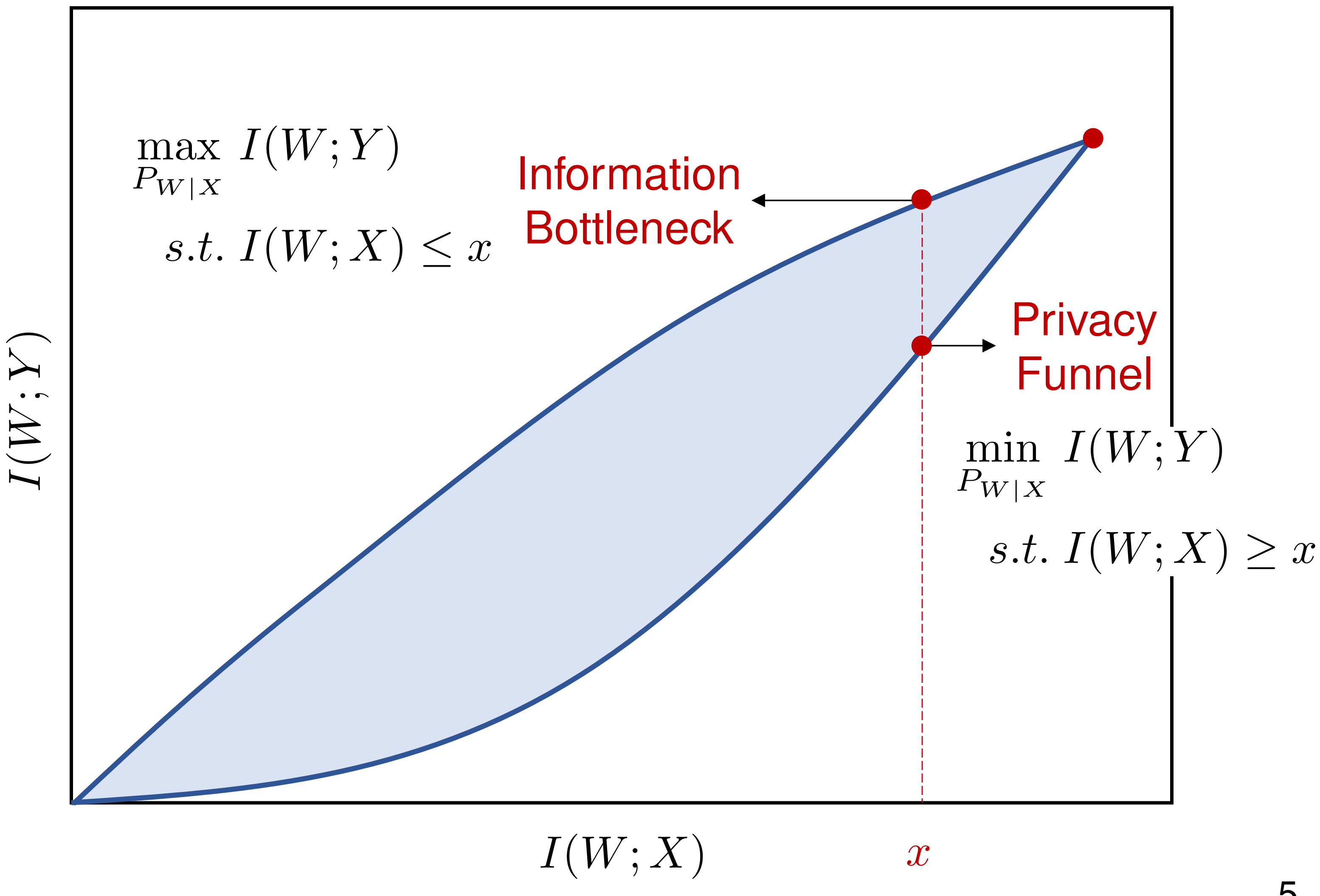
Bottleneck Problems

- Given $Y - X - W$, and a fixed P_{XY}



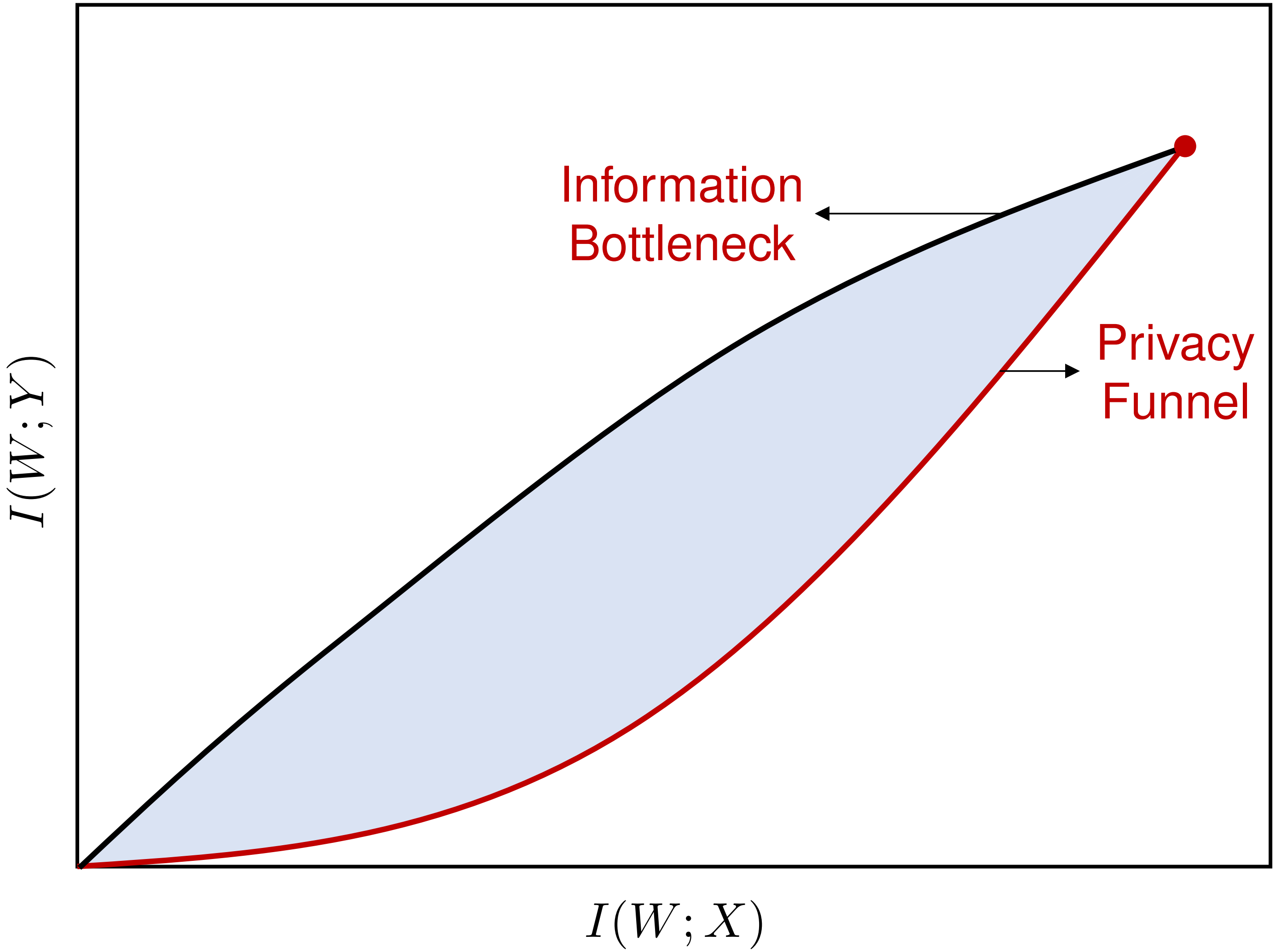
Bottleneck Problems

- Given $Y - X - W$, and a fixed P_{XY}



Bottleneck Problems

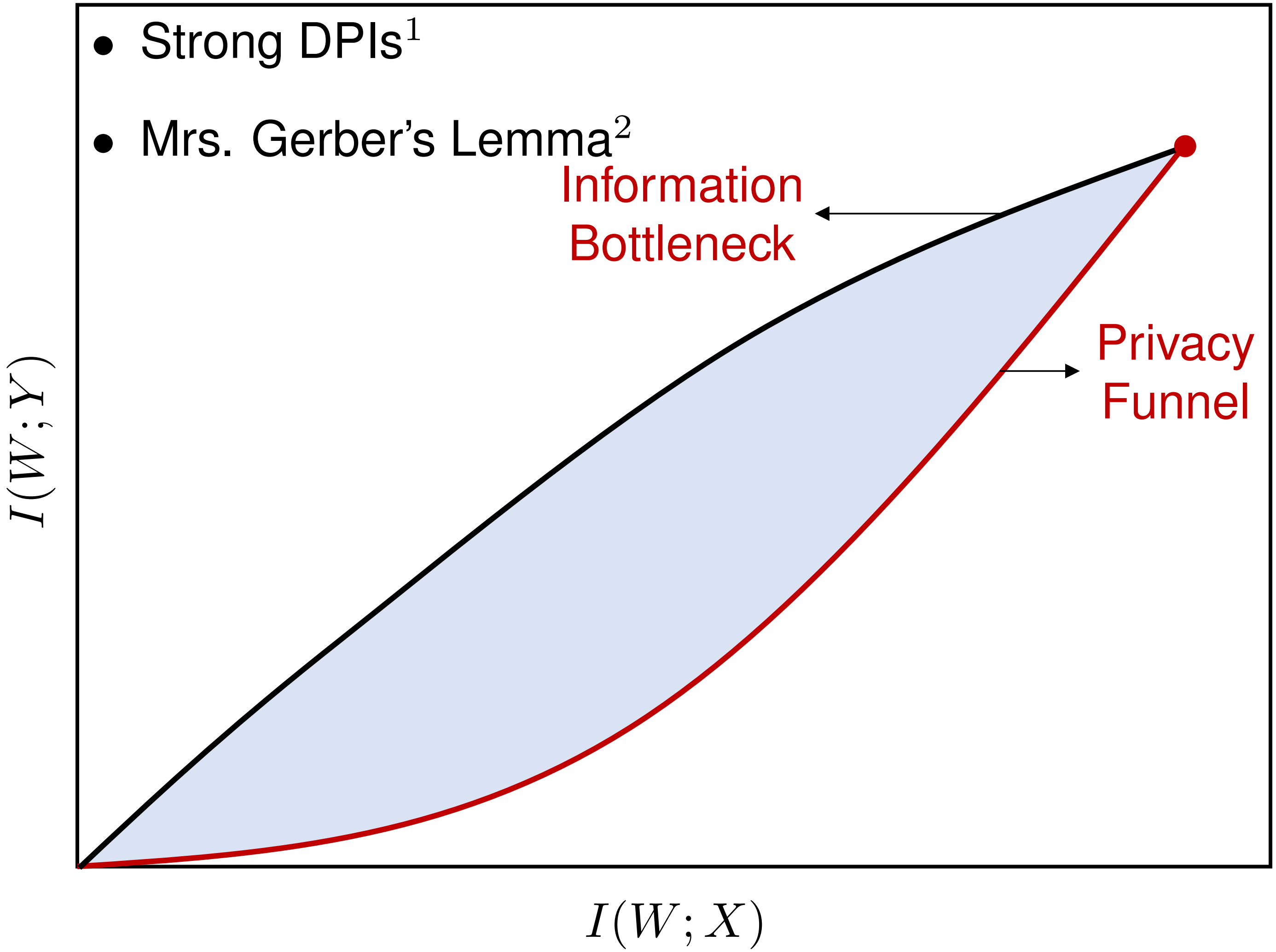
- Given $Y - X - W$, and a fixed P_{XY}



[1]Y. Polyanskiy et al.'15, [2]A. D. Wyner et al.'73

Bottleneck Problems

- Given $Y - X - W$, and a fixed P_{XY}



[1]Y. Polyanskiy et al.'15, [2]A. D. Wyner et al.'73

Outline

1. Two bottleneck problems
   - Information Bottleneck and Privacy Funnel

2. Generalizing bottleneck problems
   - Motivation and Formulation

3. Geometric properties of bottleneck problems
   - Witsenhausen and Wyner
   - How to solve generalizing bottleneck problems?

4. Applications
   - Mrs. and Mr. Gerber's Lemma
   - Arimoto's Mrs. and Mr. Gerber's Lemma
   - Estimation Bottleneck and Estimation Privacy Funnel

5. Final remarks

Why Constrain to Mutual Information?

- For an f-divergence $D_f(P\|Q)$, the f-information is given by

$$I_f(X;Y) \triangleq D_f(P_{XY}\|P_X P_Y)$$

- Other f-divergence may carry richer statistical interpretations

f-Divergence	$f(t)$	Usage
KL-Divergence	$t\log t$	Large Deviation Theory
χ^2-divergence	t^2-1	Mean Square Error
Total Variation	$\frac{1}{2}\lvert t-1\rvert$	Hypothesis Testing
Hellinger Distance	$(\sqrt{t}-1)^2$	Classification Problems

Why Constrain to Mutual Information?

- For an f-divergence $D_f(P\|Q)$, the f-information is given by

$$I_f(X;Y) \triangleq D_f(P_{XY}\|P_X P_Y)$$

- Other f-divergence may carry richer statistical interpretations

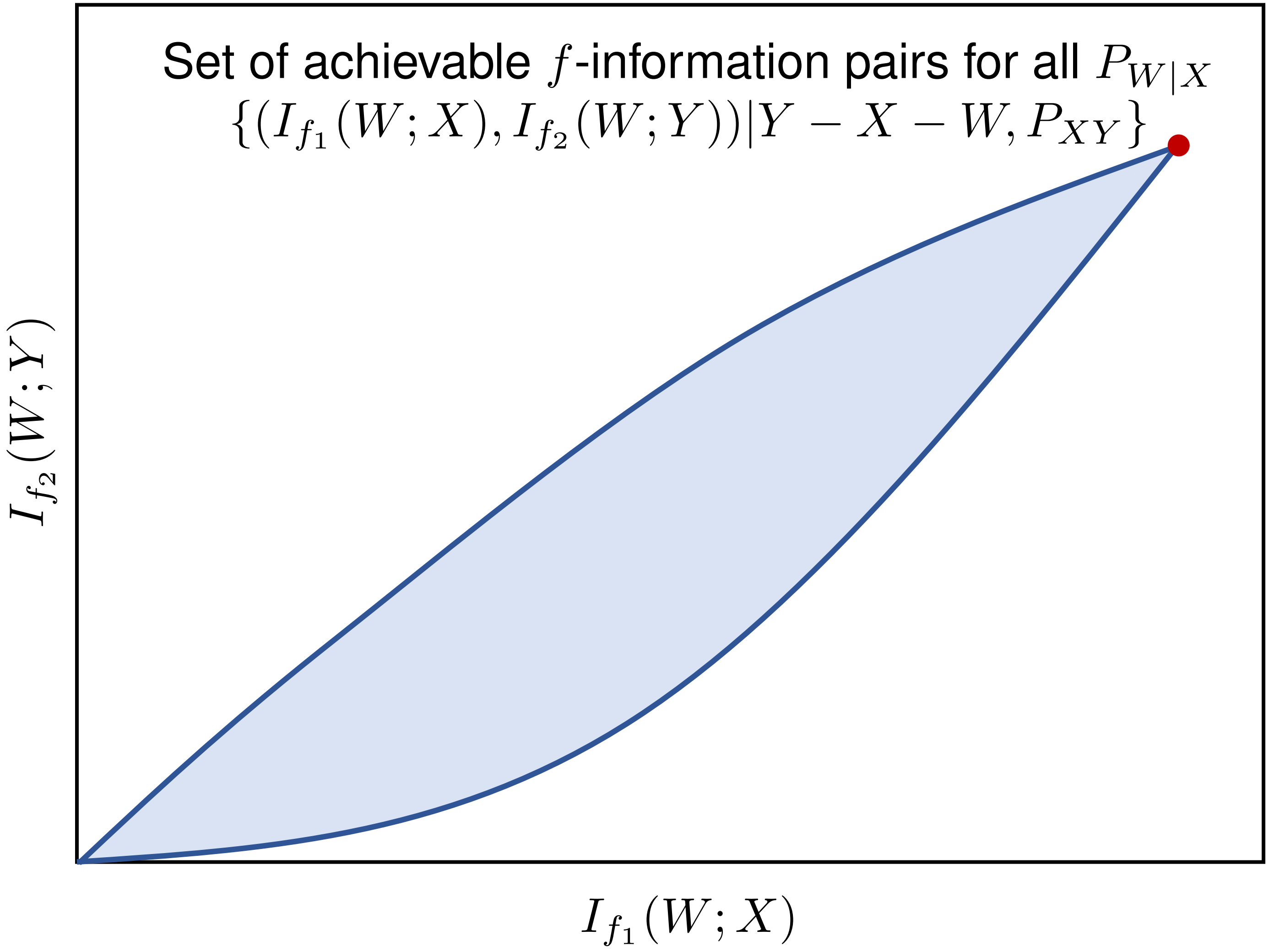
f-Divergence	$f(t)$	Usage
KL-Divergence	$t\log t$	Large Deviation Theory
χ^2-divergence	t^2-1	Mean Square Error
Total Variation	$\frac{1}{2}\|t-1\|$	Hypothesis Testing
Hellinger Distance	$(\sqrt{t}-1)^2$	Classification Problems

- Generalizing Bottleneck Problems: Given two convex functions f_1 and f_2, we are interested in the upper and lower boundaries of the set

$$\left\{\Big(I_{f_1}(W;X), I_{f_2}(W;Y)\Big)|Y-X-W, P_{XY}\right\}$$

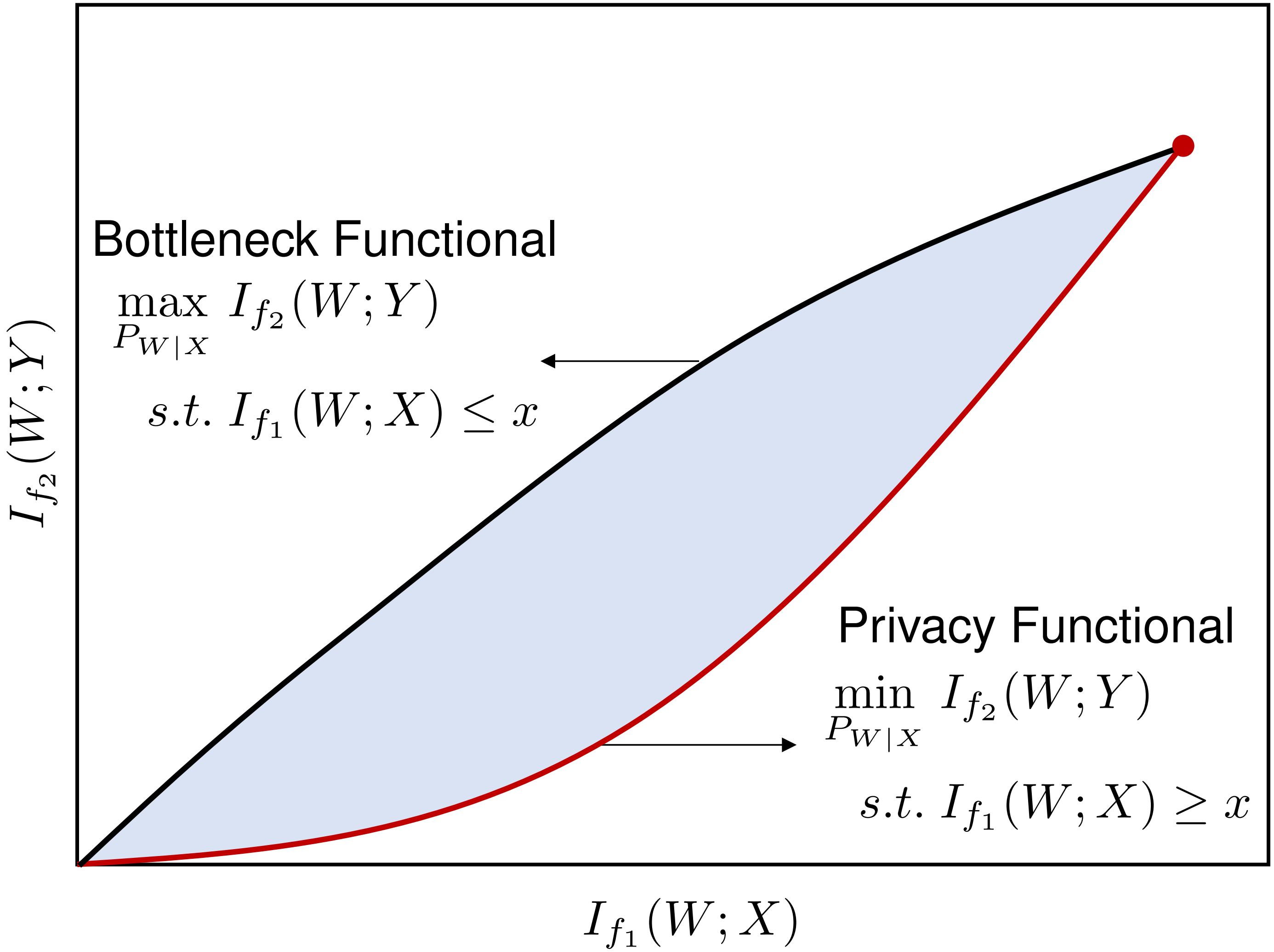
Generalizing Bottleneck Problems

- Given $Y - X - W$, and a fixed P_{XY}



Generalizing Bottleneck Problems

- Given $Y - X - W$, and a fixed P_{XY}



Outline

1. Two bottleneck problems
   - Information Bottleneck and Privacy Funnel

2. Generalizing bottleneck problems
   - Motivation and Formulation

3. Geometric properties of bottleneck problems
   - Witsenhausen and Wyner
   - How to solve generalizing bottleneck problems?

4. Applications
   - Mrs. and Mr. Gerber's Lemma
   - Arimoto's Mrs. and Mr. Gerber's Lemma
   - Estimation Bottleneck and Estimation Privacy Funnel

5. Final remarks

Geometric Properties of Bottleneck Problems

- Assume $|\mathcal{X}| = m$, $|\mathcal{Y}| = n$
- For a $P_X = \mathbf{q}$ and a channel $\mathbf{T} = P_{Y|X}$, we have a set of achievable f-information pairs

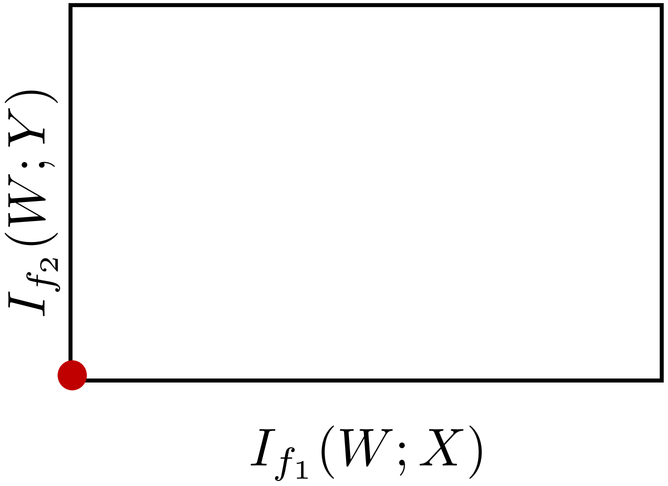
$$\{(I_{f_1}(W;X), I_{f_2}(W;Y))|Y - X - W, P_{XY}\}$$

Geometric Properties of Bottleneck Problems

- Assume $|\mathcal{X}| = m$, $|\mathcal{Y}| = n$
- For a $P_X = \mathbf{q}$ and a channel $\mathbf{T} = P_{Y|X}$, we have a set of achievable f-information pairs

$$\{(I_{f_1}(W;X), I_{f_2}(W;Y))|Y - X - W, P_{XY}\}$$

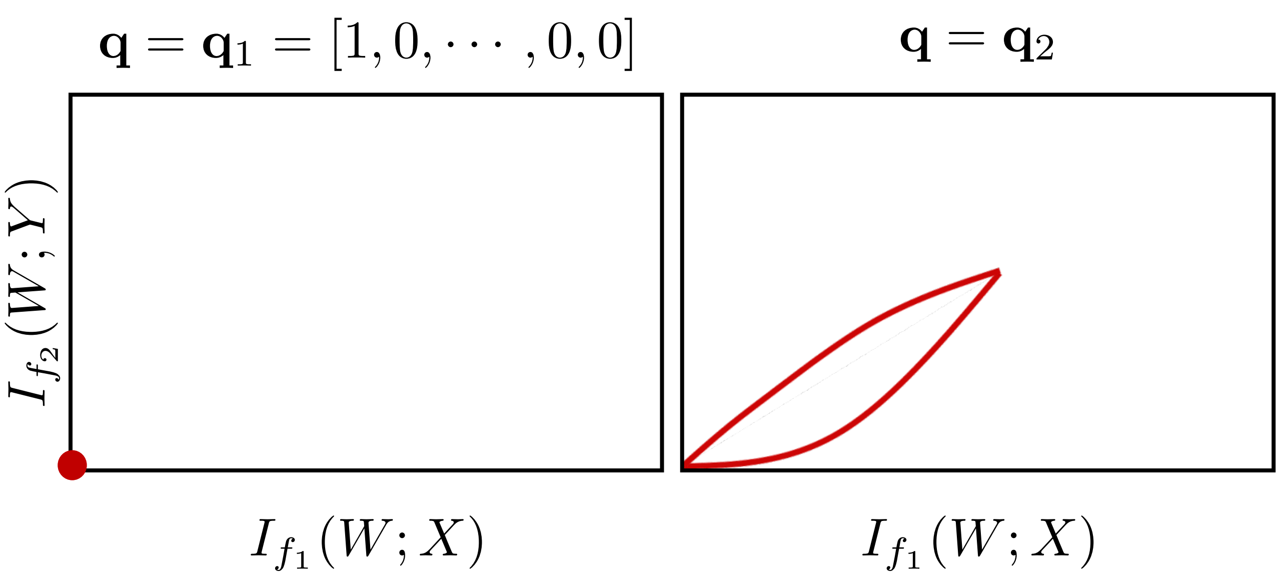
$$\mathbf{q} = \mathbf{q}_1 = [1, 0, \cdots, 0, 0]$$



Geometric Properties of Bottleneck Problems

- Assume $|\mathcal{X}| = m$, $|\mathcal{Y}| = n$
- For a $P_X = \mathbf{q}$ and a channel $\mathbf{T} = P_{Y|X}$, we have a set of achievable f-information pairs

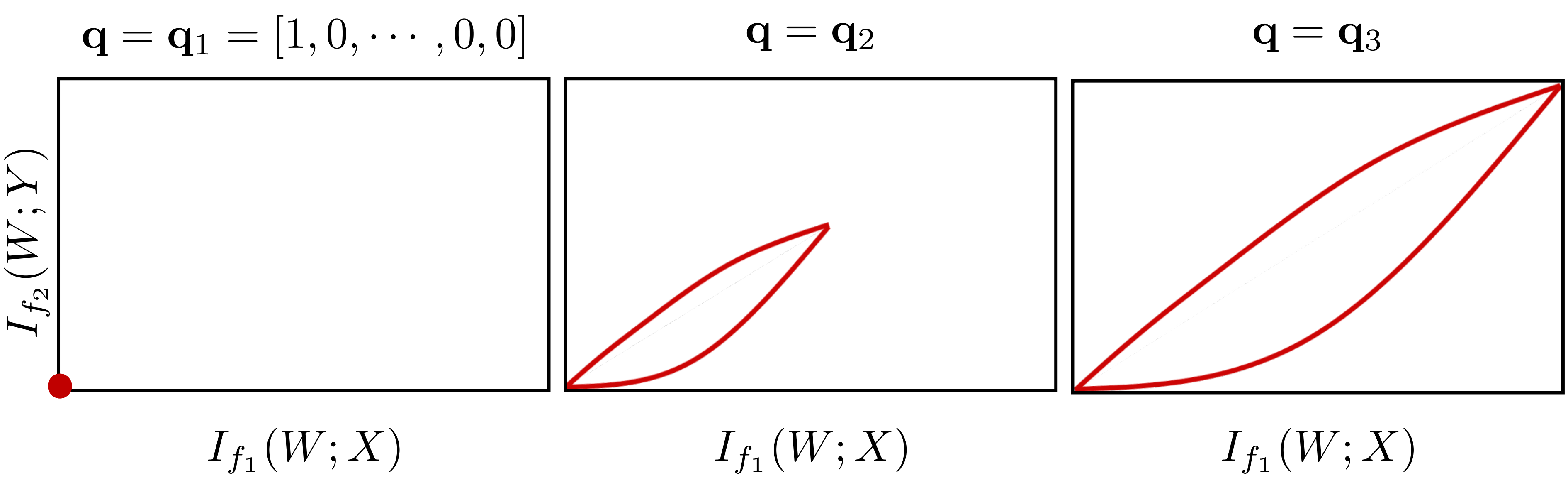
$$\{(I_{f_1}(W;X), I_{f_2}(W;Y))|Y - X - W, P_{XY}\}$$



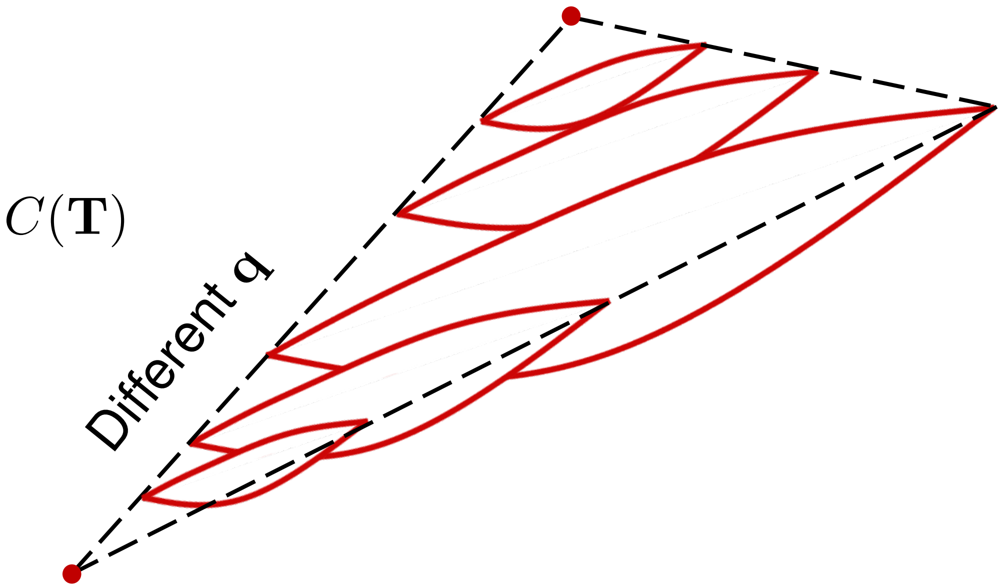
Geometric Properties of Bottleneck Problems

- Assume $|\mathcal{X}| = m$, $|\mathcal{Y}| = n$
- For a $P_X = \mathbf{q}$ and a channel $\mathbf{T} = P_{Y|X}$, we have a set of achievable f-information pairs

$$\{(I_{f_1}(W;X), I_{f_2}(W;Y))|Y - X - W, P_{XY}\}$$



Geometric Properties of Bottleneck Problems



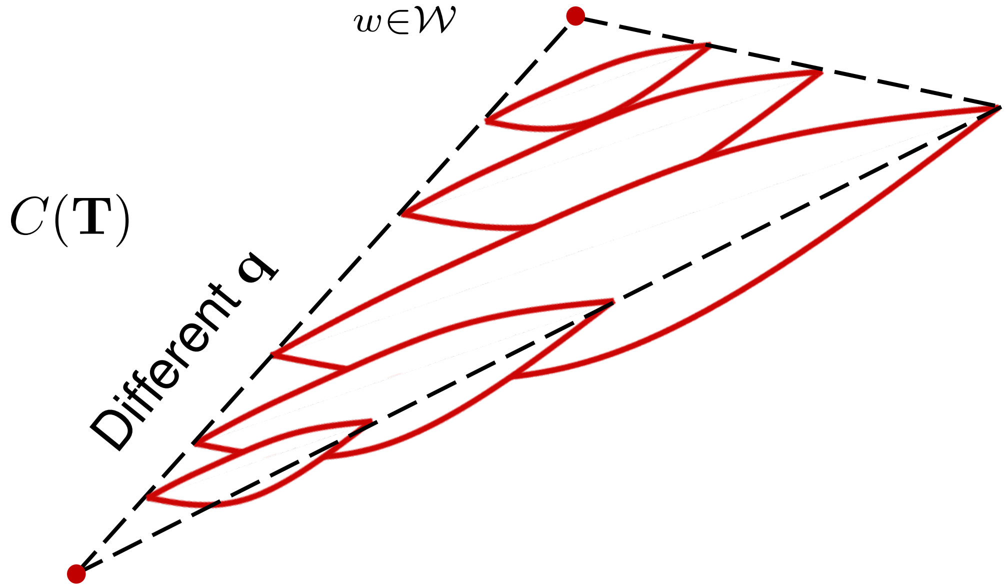
Geometric Properties of Bottleneck Problems

- Consider two continuous and bounded mappings f and g

$$f(P_{X|W}) \triangleq D_{f_1}(P_{X|W} \| P_X), \; g(\mathbf{T}P_{X|W}) \triangleq D_{f_2}(\mathbf{T}P_{X|W} \| P_Y)$$

- $I_{f_1}(W;X) = \mathbb{E}_{P_W}[f(P_{X|W})]$, $I_{f_2}(W;Y) = \mathbb{E}_{P_W}[g(\mathbf{T}P_{X|W})]$
- Collect all the achievable f-information pairs

$$C(\mathbf{T}) \triangleq \Big\{ \big(\mathbf{q}, \mathbb{E}_{P_W}[f(P_{X|W})], \mathbb{E}_{P_W}[g(\mathbf{T}P_{X|W})]\big) \, |P_{X|W=w} \in \Delta_m, \sum_{w \in \mathcal{W}} P_W(w) P_{X|W}(\cdot|W = w) = \mathbf{q}, \sum_{w \in \mathcal{W}} P_W(w) = 1 \Big\}$$



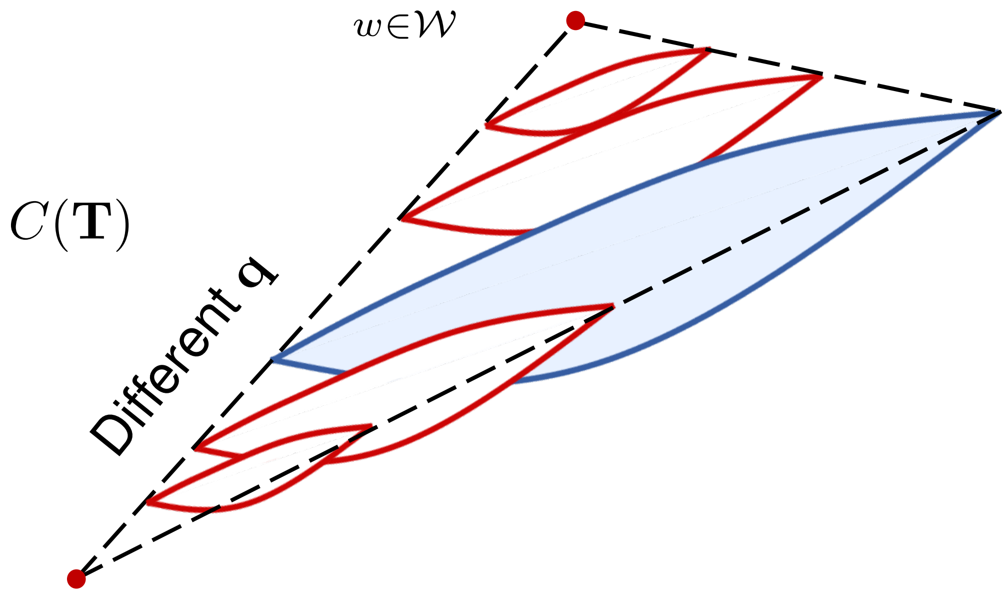
Geometric Properties of Bottleneck Problems

- Consider two continuous and bounded mappings f and g

$$f(P_{X|W}) \triangleq D_{f_1}(P_{X|W}\|P_X),\ g(\mathbf{T}P_{X|W}) \triangleq D_{f_2}(\mathbf{T}P_{X|W}\|P_Y)$$

- $I_{f_1}(W;X) = \mathbb{E}_{P_W}[f(P_{X|W})]$, $I_{f_2}(W;Y) = \mathbb{E}_{P_W}[g(\mathbf{T}P_{X|W})]$
- Collect all the achievable f-information pairs

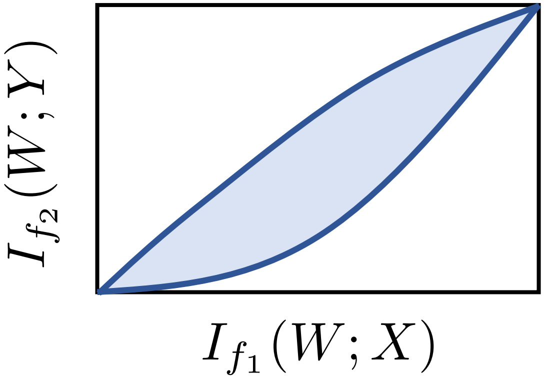
$$C(\mathbf{T}) \triangleq \Big\{ \big(\mathbf{q}, \mathbb{E}_{P_W}[f(P_{X|W})], \mathbb{E}_{P_W}[g(\mathbf{T}P_{X|W})]\big)\,|P_{X|W=w} \in \Delta_m, \sum_{w\in\mathcal{W}} P_W(w)P_{X|W}(\cdot|W=w) = \mathbf{q}, \sum_{w\in\mathcal{W}} P_W(w) = 1\Big\}$$

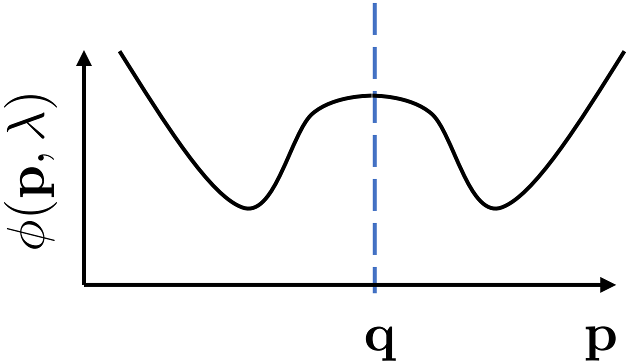


Geometric Properties of Bottleneck Problems

- Dual formulation:

$$\phi(\mathbf{p}, \lambda) \triangleq g(\mathbf{T}\mathbf{p}) - \lambda f(\mathbf{p}) = D_{f_2}(\mathbf{T}\mathbf{p} \| P_Y) - \lambda D_{f_1}(\mathbf{p} \| P_X)$$

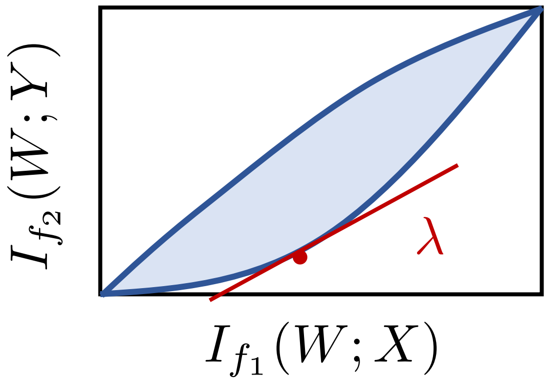


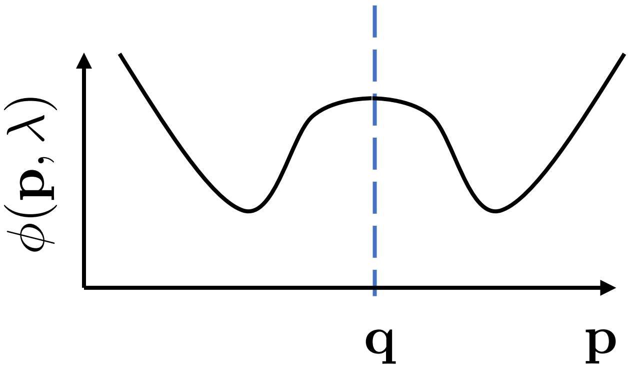


Geometric Properties of Bottleneck Problems

- Dual formulation:

$$\phi(\mathbf{p}, \lambda) \triangleq g(\mathbf{T}\mathbf{p}) - \lambda f(\mathbf{p}) = D_{f_2}(\mathbf{T}\mathbf{p} \| P_Y) - \lambda D_{f_1}(\mathbf{p} \| P_X)$$





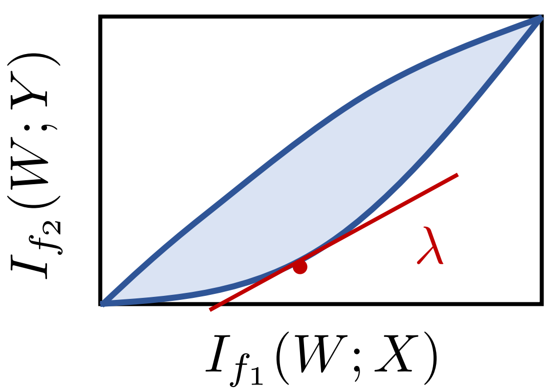
Geometric Properties of Bottleneck Problems

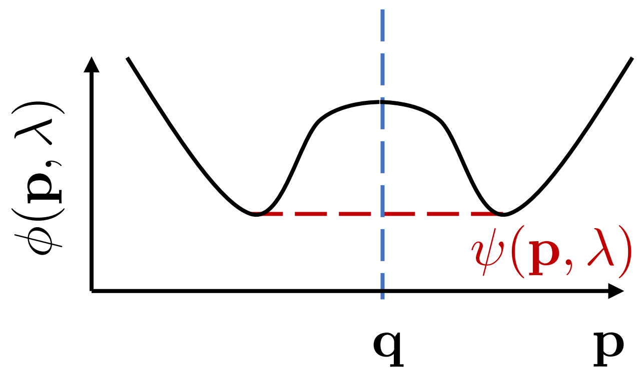
- Dual formulation:

$$\phi(\mathbf{p}, \lambda) \triangleq g(\mathbf{Tp}) - \lambda f(\mathbf{p}) = D_{f_2}(\mathbf{Tp} \| P_Y) - \lambda D_{f_1}(\mathbf{p} \| P_X)$$

- The lower convex envelope of $\phi(\mathbf{p}, \lambda)$ at $\mathbf{q} = P_X$

$$\begin{aligned} \psi(\mathbf{q}, \lambda) &= \min I_{f_2}(W;Y) - \lambda I_{f_1}(W;X) \\ &\Rightarrow (I_{f_1}(W;X),\ \min I_{f_2}(W;Y)) \text{ corresponds to slope } \lambda \end{aligned}$$





Geometric Properties of Bottleneck Problems

- Dual formulation:

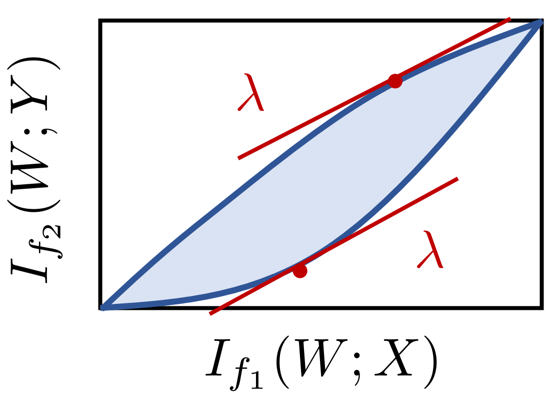
$$\phi(\mathbf{p}, \lambda) \triangleq g(\mathbf{Tp}) - \lambda f(\mathbf{p}) = D_{f_2}(\mathbf{Tp} \| P_Y) - \lambda D_{f_1}(\mathbf{p} \| P_X)$$

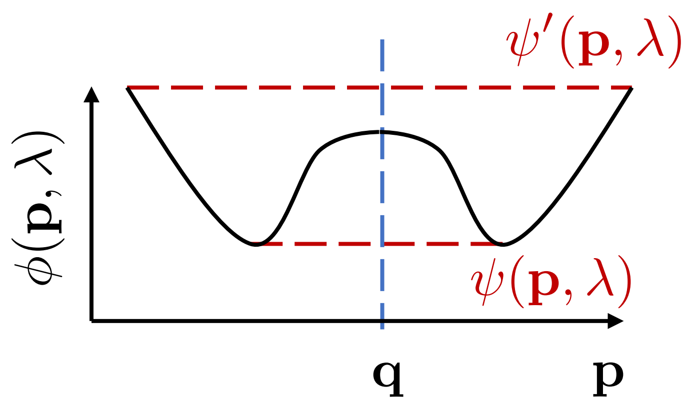
- The lower convex envelope of $\phi(\mathbf{p}, \lambda)$ at $\mathbf{q} = P_X$

$$\begin{aligned} \psi(\mathbf{q}, \lambda) &= \min I_{f_2}(W;Y) - \lambda I_{f_1}(W;X) \\ &\Rightarrow (I_{f_1}(W;X),\ \min I_{f_2}(W;Y)) \text{ corresponds to slope } \lambda \end{aligned}$$

- Similarly, the upper convex envelope $\psi'(\mathbf{q}, \lambda)$ gives

$$(I_{f_1}(W;X),\ \max I_{f_2}(W;Y)) \text{ corresponds to slope } \lambda$$





Outline

1. Two bottleneck problems
   - Information Bottleneck and Privacy Funnel

2. Generalizing bottleneck problems
   - Motivation and Formulation

3. Geometric properties of bottleneck problems
   - Witsenhausen and Wyner
   - How to solve generalizing bottleneck problems?

4. Applications
   - Mrs. and Mr. Gerber's Lemma
   - Arimoto's Mrs. and Mr. Gerber's Lemma
   - Estimation Bottleneck and Estimation Privacy Funnel

5. Final remarks

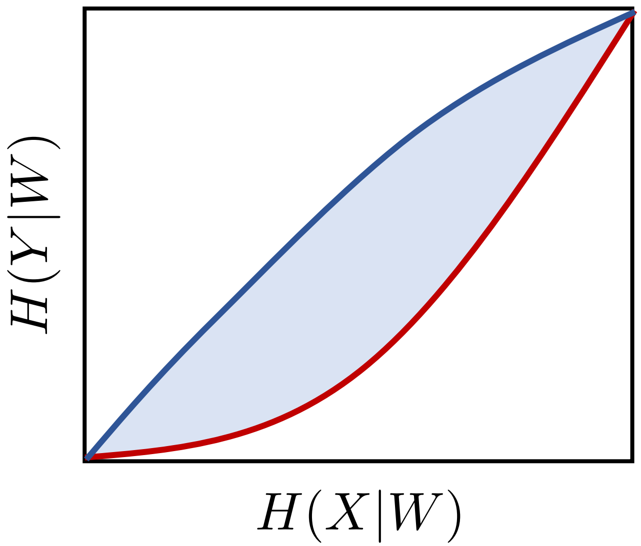
Mrs. and Mr. Gerber's Lemma

- $\Pr(X = 1) = q$, $\mathbf{T} = P_{Y|X} =$ BSC with crossover probability δ
- $f_1 = f_2 = h_b$, the binary entropy function

Lemma 1 (Mrs. Gerber's Lemma)**.** *Given* $0 \le x \le H(X)$

$$\inf_{\substack{P_{W|X} \\ H(X|W) \ge x}} H(Y|W) = h_b(\delta \star h_b^{-1}(x)),$$

where $h_b^{-1} : [0,1] \to [0, \frac{1}{2}]$ *is the inverse function of* $h_b(\cdot)$*, and* $a \star b \triangleq (1-a)\,b + (1-b)\,a$*, for* $a, b \in [0,1]$*.*



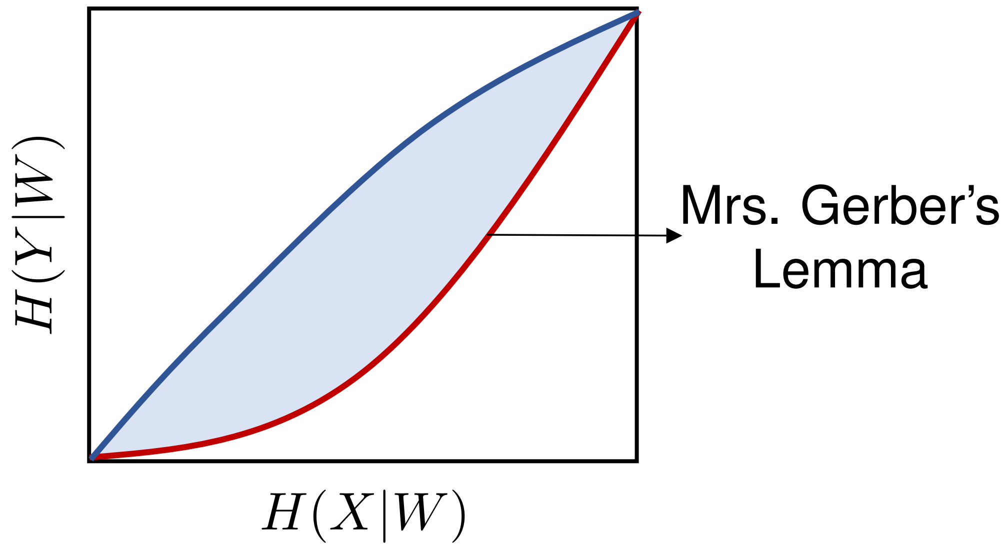
Mrs. and Mr. Gerber's Lemma

- $\Pr(X = 1) = q$, $\mathbf{T} = P_{Y|X}$ = BSC with crossover probability δ
- $f_1 = f_2 = h_b$, the binary entropy function

Lemma 1 (Mrs. Gerber's Lemma)**.** *Given* $0 \leq x \leq H(X)$

$$\inf_{\substack{P_{W|X} \\ H(X|W) \geq x}} H(Y|W) = h_b(\delta \star h_b^{-1}(x)),$$

where $h_b^{-1} : [0,1] \to [0, \frac{1}{2}]$ *is the inverse function of* $h_b(\cdot)$*, and* $a \star b \triangleq (1-a)b + (1-b)a$*, for* $a, b \in [0,1]$*.*



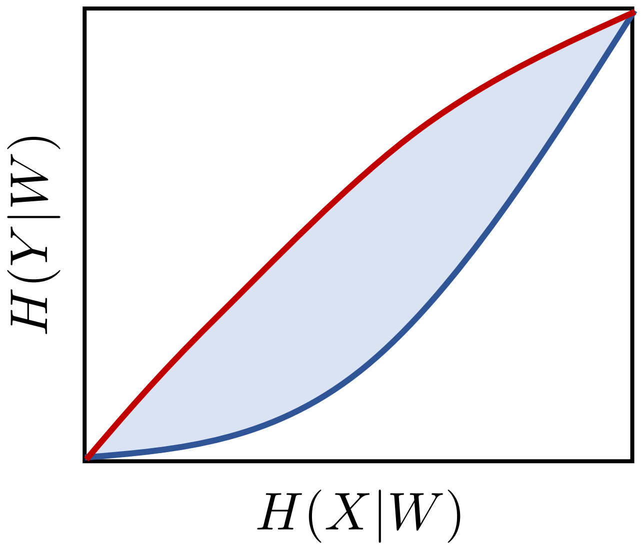
Mrs. and Mr. Gerber's Lemma

- $\Pr(X=1)=q$, $\mathbf{T}=P_{Y|X}=$ BSC with crossover probability δ
- $f_1=f_2=h_b$, the binary entropy function

Theorem 1 (Mr. Gerber's Lemma)**.** *Given* $0\leq x\leq H(X)$

$$\sup_{\substack{P_{W|X}\\ H(X|W)\leq x}} H(Y|W)=\alpha h_b\left(\delta\star\frac{q}{z}\right)+\bar{\alpha}h_b\left(\delta\right),$$

where $x=\alpha h_b\left(\frac{q}{z}\right)$ *and* $z=\max\left(\alpha,2q\right)$, $\alpha\in[0,1]$*.*



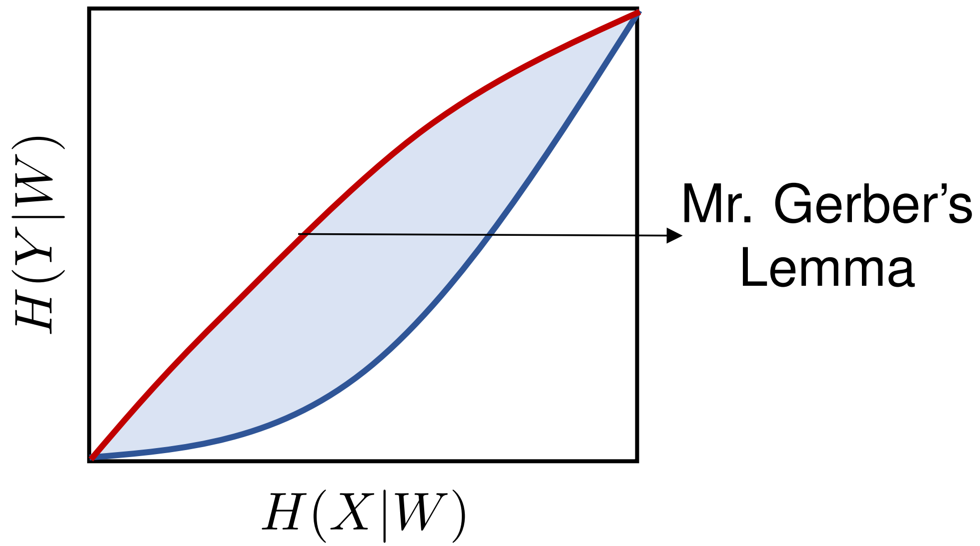
Mrs. and Mr. Gerber's Lemma

- $\Pr(X=1) = q$, $\mathbf{T} = P_{Y|X} =$ BSC with crossover probability δ
- $f_1 = f_2 = h_b$, the binary entropy function

Theorem 1 (Mr. Gerber's Lemma)**.** *Given* $0 \leq x \leq H(X)$

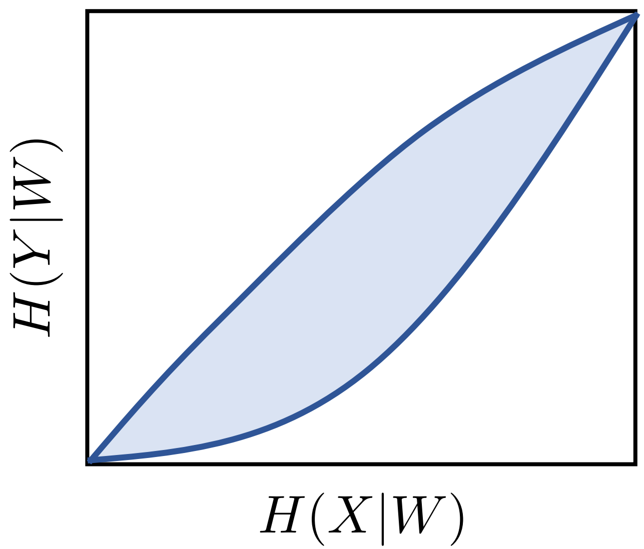
$$\sup_{\substack{P_{W|X} \\ H(X|W) \leq x}} H(Y|W) = \alpha h_b\left(\delta \star \frac{q}{z}\right) + \bar{\alpha} h_b(\delta),$$

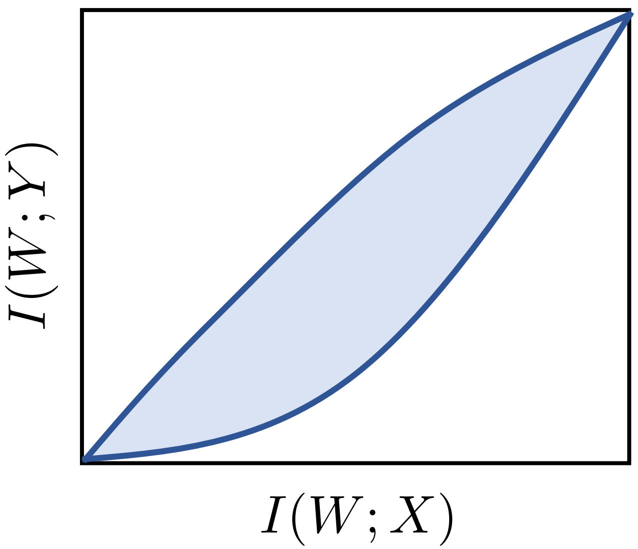
where $x = \alpha h_b\left(\frac{q}{z}\right)$ *and* $z = \max(\alpha, 2q)$, $\alpha \in [0,1]$*.*



Mrs. and Mr. Gerber's Lemma

- $\Pr(X = 1) = q$, $\mathbf{T} = P_{Y|X} =$ BSC with crossover probability δ
- $f_1 = f_2 = h_b$, the binary entropy function

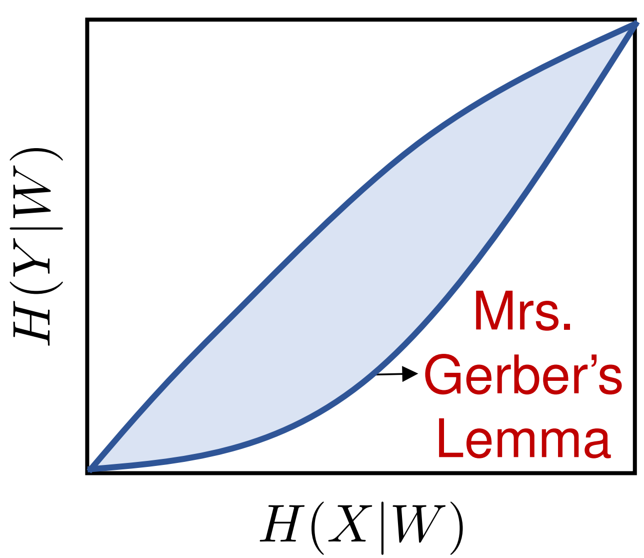


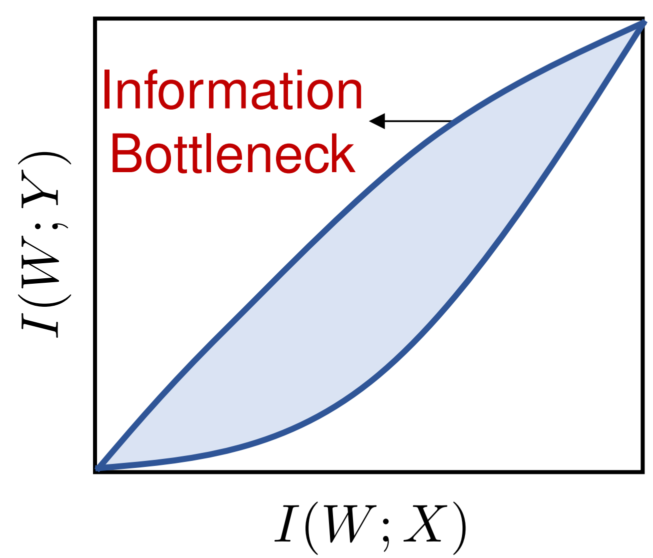


Mrs. and Mr. Gerber's Lemma

- $\Pr(X = 1) = q$, $\mathbf{T} = P_{Y|X} =$ BSC with crossover probability δ
- $f_1 = f_2 = h_b$, the binary entropy function

Mrs. Gerber's Lemma corresponds to Information Bottleneck for BSC



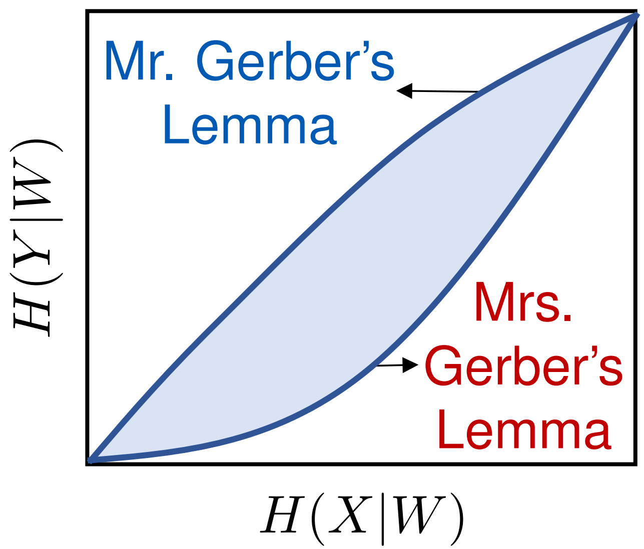


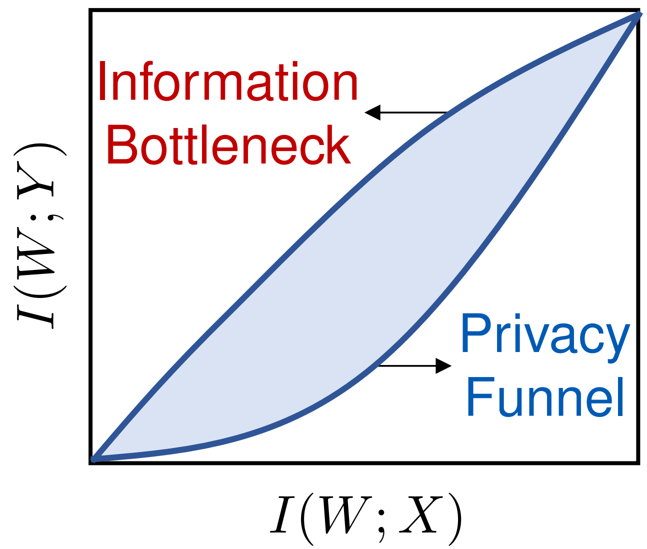
Mrs. and Mr. Gerber's Lemma

- $\Pr(X=1)=q$, $\mathbf{T}=P_{Y|X}=$ BSC with crossover probability δ
- $f_1=f_2=h_b$, the binary entropy function

Mrs. Gerber's Lemma corresponds to Information Bottleneck for BSC

Mr. Gerber's Lemma corresponds to Privacy Funnel for BSC





Outline

1. Two bottleneck problems
   - Information Bottleneck and Privacy Funnel

2. Generalizing bottleneck problems
   - Motivation and Formulation

3. Geometric properties of bottleneck problems
   - Witsenhausen and Wyner
   - How to solve generalizing bottleneck problems?

4. Applications
   - Mrs. and Mr. Gerber's Lemma
   - Arimoto's Mrs. and Mr. Gerber's Lemma
   - Estimation Bottleneck and Estimation Privacy Funnel

5. Final remarks

Arimoto's Mrs. and Mr. Gerber's Lemma

- Use ℓ^β-norm for f and g with $\beta \geq 2$
- Arimoto's conditional entropy:

$$H_\beta(X|W) \triangleq \frac{\beta}{1-\beta} \log \mathbb{E}\left[\|P_{X|W}(\cdot|W)\|_\beta\right]$$

- Rényi entropy of order β:

$$H_\beta(X) = \log \|P_X\|_\beta$$

- Arimoto's mutual information: $I_\beta(X;W) = H_\beta(X) - H_\beta(X|W)$
- Useful in statistics and hypothesis testing[1]

[1] I. Sason et al.'17

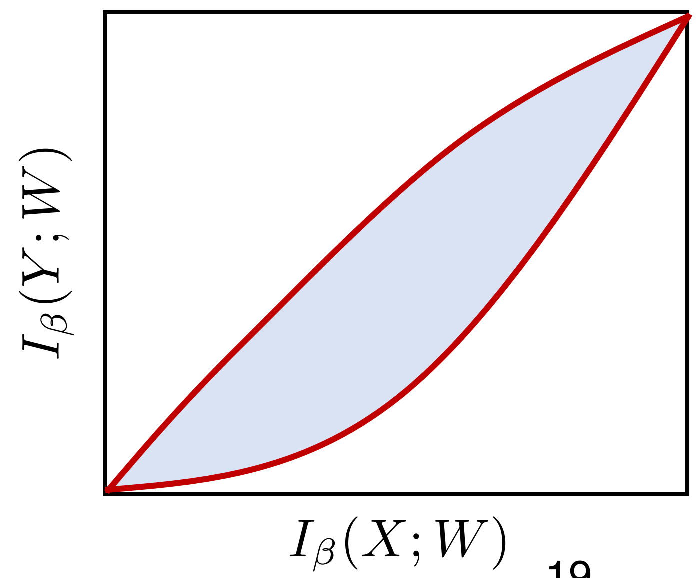
Arimoto's Mrs. and Mr. Gerber's Lemma

- $\mathbf{T} = \text{BSC}$
- Arimoto's Mrs. Gerber's Lemma solves

$$\inf_{\substack{P_{W|X} \\ H_\beta(X|W) \geq x}} H_\beta(Y|W)$$

- Arimoto's Mr. Gerber's Lemma solves

$$\sup_{\substack{P_{W|X} \\ H_\beta(X|W) \leq x}} H_\beta(Y|W)$$



Arimoto's Mrs. and Mr. Gerber's Lemma

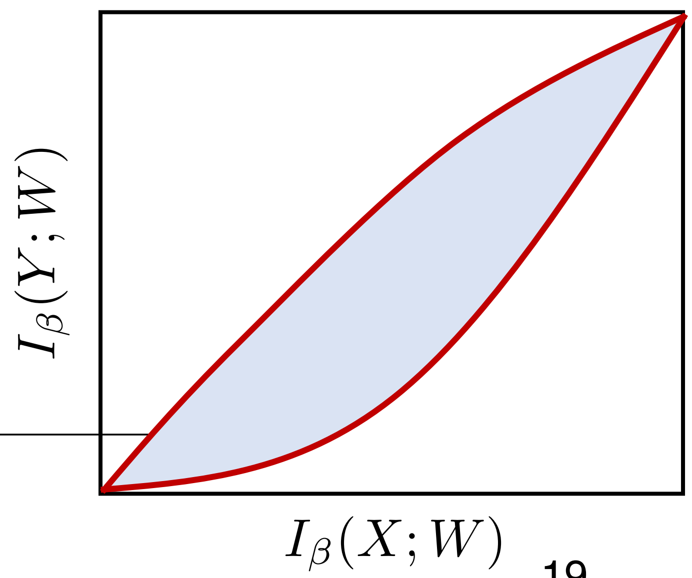
- $\mathbf{T} = \text{BSC}$
- Arimoto's Mrs. Gerber's Lemma solves

$$\inf_{\substack{P_{W|X} \\ H_\beta(X|W) \geq x}} H_\beta(Y|W)$$

- Arimoto's Mr. Gerber's Lemma solves

$$\sup_{\substack{P_{W|X} \\ H_\beta(X|W) \leq x}} H_\beta(Y|W)$$

Arimoto's
Mrs. Gerber's Lemma



Arimoto's Mrs. and Mr. Gerber's Lemma

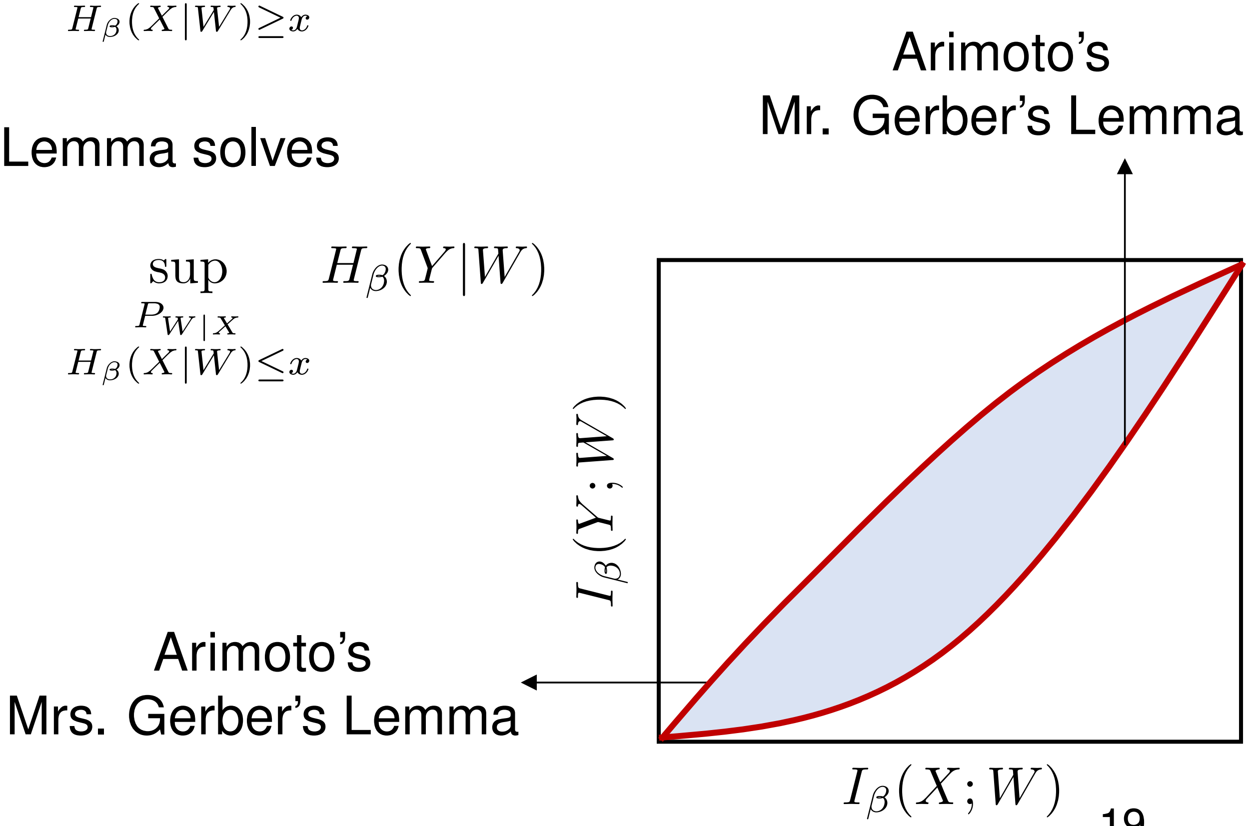
- $\mathbf{T} = \text{BSC}$

- Arimoto's Mrs. Gerber's Lemma solves

$$\inf_{\substack{P_{W|X} \\ H_\beta(X|W) \geq x}} H_\beta(Y|W)$$

- Arimoto's Mr. Gerber's Lemma solves

$$\sup_{\substack{P_{W|X} \\ H_\beta(X|W) \leq x}} H_\beta(Y|W)$$



Outline

Estimation Bottleneck and Privacy Funnel

- $f_1(t) = f_2(t) = t^2 - 1$
- χ^2-information: $\chi^2(W;X) = \mathbb{E}\left[\frac{P_{WX}(W,X)}{P_W(W)P_X(X)}\right] - 1$

Estimation Bottleneck

$$\max_{P_{W|X}} \chi^2(W;Y)$$
$$s.t.\ \chi^2(W;X) \leq x$$

Estimation Privacy Funnel

$$\min_{P_{W|X}} \chi^2(W;Y)$$
$$s.t.\ \chi^2(W;X) \geq x$$

[1]F. P. Calmon et al.'17, [2]A. Makur et al.'15

Estimation Bottleneck and Privacy Funnel

- $f_1(t) = f_2(t) = t^2 - 1$
- χ^2-information: $\chi^2(W;X) = \mathbb{E}\left[\frac{P_{WX}(W,X)}{P_W(W)P_X(X)}\right] - 1$

Estimation Bottleneck

$$\max_{P_{W|X}} \chi^2(W;Y)$$
$$s.t.\ \chi^2(W;X) \le x$$

Estimation Privacy Funnel

$$\min_{P_{W|X}} \chi^2(W;Y)$$
$$s.t.\ \chi^2(W;X) \ge x$$

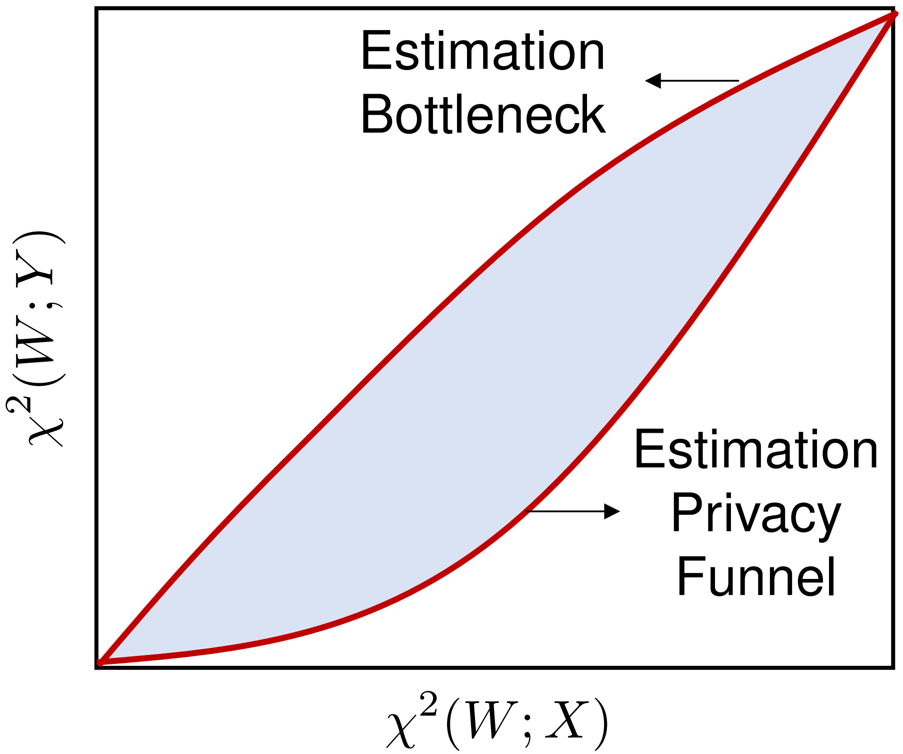
Why χ^2-information?

- $\chi^2(W;Y) = \sum_{i=1}^{d} \lambda_i(W;Y)$, where $d = \min(|\mathcal{W}|, |\mathcal{Y}|) - 1$ and $\lambda_i(W;Y)$ is the i$^{\text{th}}$ principal inertia component (PIC)[1] of W and Y (larger PICs give smaller MMSE)
- χ^2-information bounds f-information[2]

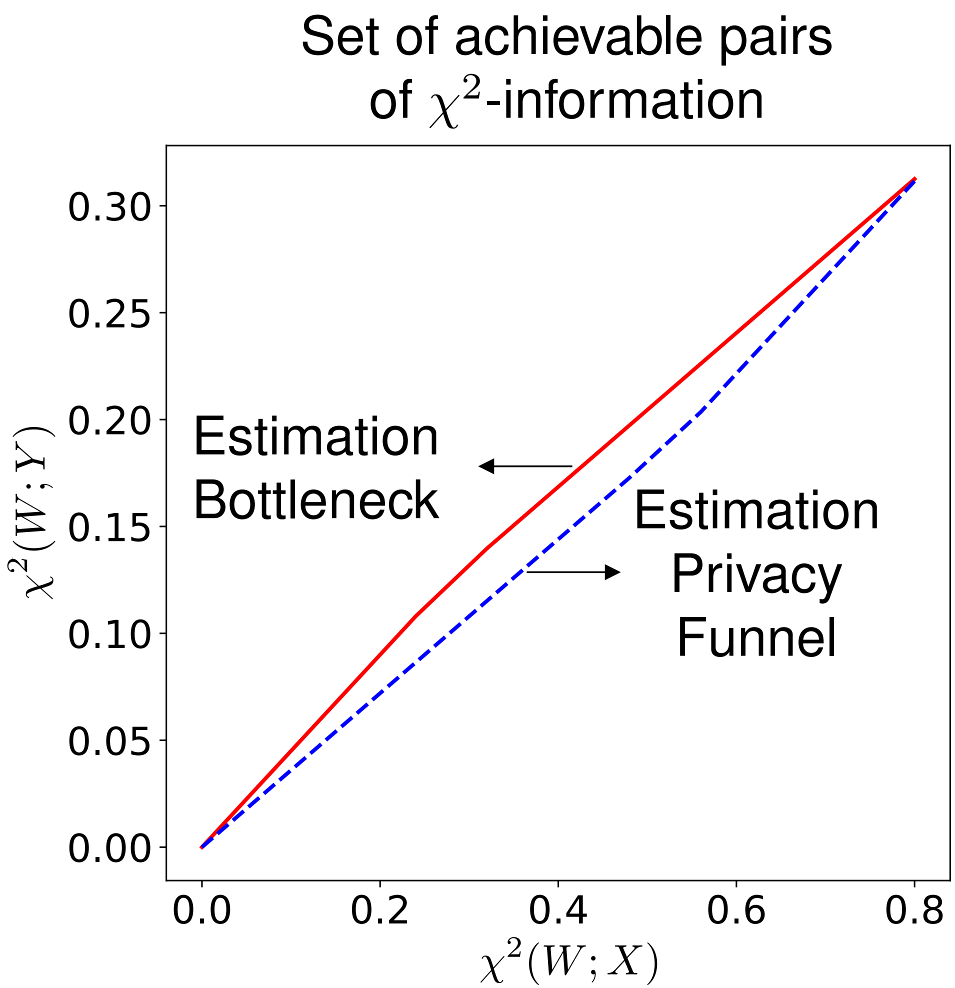
[1]F. P. Calmon et al.'17, [2]A. Makur et al.'15

Estimation Bottleneck and Privacy Funnel

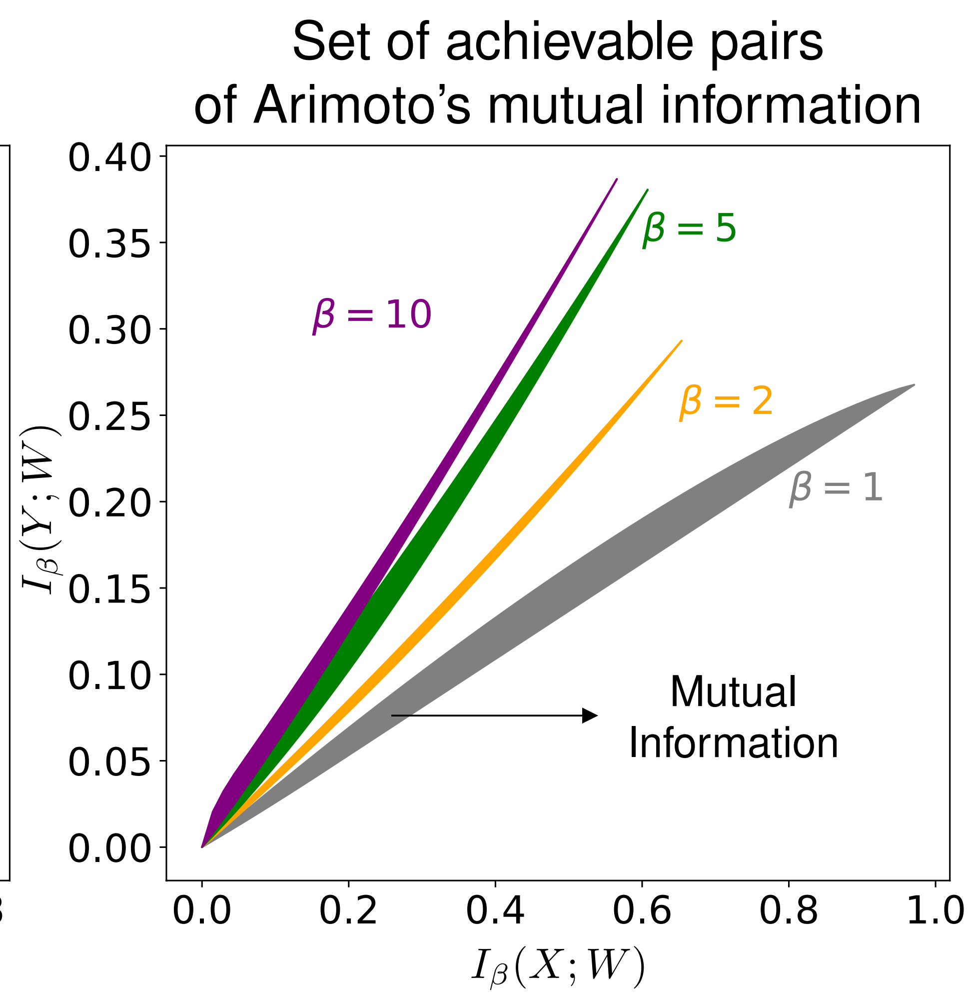
- $f_1(t) = f_2(t) = t^2 - 1$
- χ^2-information: $\chi^2(W;X) = \mathbb{E}\left[\frac{P_{WX}(W,X)}{P_W(W)P_X(X)}\right] - 1$



Numerical Results for BSC



$\delta = 0.1,\ q = 0.1$



$\delta = 0.2,\ q = 0.4$

Outline

1. Two bottleneck problems
   - Information Bottleneck and Privacy Funnel

2. Generalizing bottleneck problems
   - Motivation and Formulation

3. Geometric properties of bottleneck problems
   - Witsenhausen and Wyner
   - How to solve generalizing bottleneck problems?

4. Applications
   - Mrs. and Mr. Gerber's Lemma
   - Arimoto's Mrs. and Mr. Gerber's Lemma
   - Estimation Bottleneck and Estimation Privacy Funnel

5. Final remarks

Final Remarks

- Revisit the geometry of generalized bottleneck problems

$$\{(I_{f_1}(W;X), I_{f_2}(W;Y))\,|Y - X - W, P_{XY}\}$$

  and provide a systematic framework to find out the upper and lower boundaries

- Generalize the information bottleneck and privacy funnel
  - Binary entropy function: Mrs. And Mr. Gerber's Lemmas
  - ℓ^β-norm: Arimoto's Mrs. And Mr. Gerber's Lemmas
  - χ^2-divergence: Estimation bottleneck and privacy funnel

- These results can be potentially useful for new applications of information theory in machine learning

Thank You for Listening!