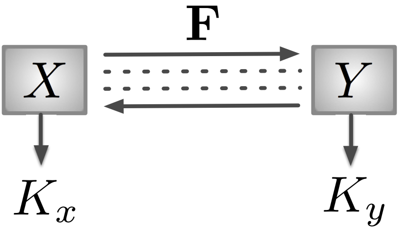
Secret Key Agreement: General Capacity and Second-Order Asymptotics

Masahito Hayashi   Himanshu Tyagi   Shun Watanabe

Two party secret key agreement

Maurer 93, Ahlswede-Csiszár 93



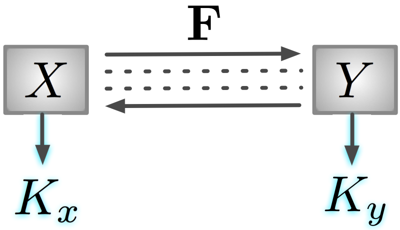
A random variable K constitutes an (ϵ, δ)-SK if:

$$\mathrm{P}\left(K_x = K_y = K\right) \geq 1 - \epsilon \quad : \text{recoverability}$$

$$\frac{1}{2}\left\|\mathrm{P}_{K\mathbf{F}} - \mathrm{P}_{\mathtt{unif}}\mathrm{P}_{\mathbf{F}}\right\| \leq \delta \quad : \text{security}$$

Two party secret key agreement

Maurer 93, Ahlswede-Csiszár 93

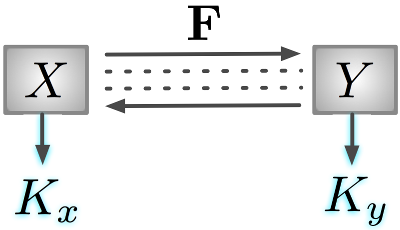


A random variable K constitutes an (ϵ, δ)-SK if:

$$
\begin{aligned}
\mathrm{P}\left(K_x = K_y = K\right) &\geq 1 - \epsilon \quad &: \text{recoverability} \\
\frac{1}{2}\left\|\mathrm{P}_{K\mathbf{F}} - \mathrm{P}_{\mathtt{unif}}\mathrm{P}_{\mathbf{F}}\right\| &\leq \delta \quad &: \text{security}
\end{aligned}
$$

Two party secret key agreement

Maurer 93, Ahlswede-Csiszár 93



A random variable K constitutes an (ϵ, δ)-SK if:

$$\mathrm{P}\left(K_x = K_y = K\right) \geq 1 - \epsilon \quad : \text{recoverability}$$

$$\frac{1}{2}\left\|\mathrm{P}_{K\mathbf{F}} - \mathrm{P}_{\mathtt{unif}}\mathrm{P}_{\mathbf{F}}\right\| \leq \delta \quad : \text{security}$$

What is the maximum length $S(X, Y)$ of a SK that can be generated?

Where do we stand?

Maurer 93, Ahlswede-Csiszár 93

$S(X^n, Y^n) = nI(X \wedge Y) + o(n)$ (Secret key capacity)

Csiszár-Narayan 04

Secret key capacity for multiple terminals

Renner-Wolf 03, 05

Single-shot bounds on $S(X, Y)$

Where do we stand?

Maurer 93, Ahlswede-Csiszár 93

$S(X^n, Y^n) = nI(X \wedge Y) + o(n)$ (Secret key capacity)

Csiszár-Narayan 04

Secret key capacity for multiple terminals

Renner-Wolf 03, 05

Single-shot bounds on $S(X, Y)$

Typical construction: X sends a compressed version of itself to Y, and the K is *extracted* from shared X using a 2-universal hash family

Where do we stand?

Maurer 93, Ahlswede-Csiszár 93

$S(X^n, Y^n) = nI(X \wedge Y) + o(n)$ (Secret key capacity)

Csiszár-Narayan 04

Secret key capacity for multiple terminals

Renner-Wolf 03, 05

Single-shot bounds on $S(X, Y)$

Typical construction: X sends a compressed version of itself to Y, and the K is *extracted* from shared X using a 2-universal hash family

Converse??

Where do we stand?

Maurer 93, Ahlswede-Csiszár 93 Fano's inequality

$S(X^n, Y^n) = nI(X \wedge Y) + o(n)$ (Secret key capacity)

Csiszár-Narayan 04 Fano's inequality

Secret key capacity for multiple terminals

Renner-Wolf 03, 05 $\sim$ *Potential function method*

Single-shot bounds on $S(X, Y)$

Typical construction: X sends a compressed version of itself to Y, and the K is *extracted* from shared X using a 2-universal hash family

Converse??

Converse: Conditional independence testing bound

The source of our rekindled excitement about this problem:

Theorem (Tyagi-Watanabe 2014)

Given $\epsilon, \delta \geq 0$ *with* $\epsilon + \delta < 1$ *and* $0 < \eta < 1 - \epsilon - \delta$. *It holds that*

$$S_{\epsilon,\delta}(X, Y) \leq -\log \beta_{\epsilon+\delta+\eta}\big(\mathrm{P}_{XY}, \mathrm{P}_X \mathrm{P}_Y\big) + 2\log(1/\eta)$$

Converse: Conditional independence testing bound

The source of our rekindled excitement about this problem:

Theorem (Tyagi-Watanabe 2014)

Given $\epsilon, \delta \geq 0$ with $\epsilon + \delta < 1$ and $0 < \eta < 1 - \epsilon - \delta$. It holds that

$$S_{\epsilon,\delta}\left(X, Y\right) \leq -\log \beta_{\epsilon+\delta+\eta}\big(\mathrm{P}_{XY}, \mathrm{P}_X \mathrm{P}_Y\big) + 2\log(1/\eta)$$

$$\beta_\epsilon(\mathrm{P}, \mathrm{Q}) \triangleq \inf_{\mathrm{T}\,:\,\mathrm{P}[\mathrm{T}] \geq 1-\epsilon} \mathrm{Q}[\mathrm{T}],$$

where

$$\mathrm{P}[\mathrm{T}] = \sum_v \mathrm{P}(v)\mathrm{T}(0|v) \qquad \mathrm{Q}[\mathrm{T}] = \sum_v \mathrm{Q}(v)\mathrm{T}(0|v)$$

Converse: Conditional independence testing bound

The source of our rekindled excitement about this problem:

Theorem (Tyagi-Watanabe 2014)

Given $\epsilon, \delta \geq 0$ *with* $\epsilon + \delta < 1$ *and* $0 < \eta < 1 - \epsilon - \delta$. *It holds that*

$$S_{\epsilon,\delta}(X, Y) \leq -\log \beta_{\epsilon+\delta+\eta}\big(\mathrm{P}_{XY}, \mathrm{P}_X \mathrm{P}_Y\big) + 2\log(1/\eta)$$

$$\beta_\epsilon(\mathrm{P}, \mathrm{Q}) \triangleq \inf_{\mathrm{T}\,:\,\mathrm{P}[\mathrm{T}] \geq 1-\epsilon} \mathrm{Q}[\mathrm{T}],$$

where

$$\mathrm{P}[\mathrm{T}] = \sum_v \mathrm{P}(v)\mathrm{T}(0|v) \qquad \mathrm{Q}[\mathrm{T}] = \sum_v \mathrm{Q}(v)\mathrm{T}(0|v)$$

In the spirit of *meta-converse* of Polyanskiy, Poor, and Verdu

Single-shot achievability?

Recall the two steps of SK agreement:

Step 1 (aka Information reconciliation).

Slepian-Wolf code to send X to Y

Step 2 (aka Randomness extraction or privacy amplification).

"Random function" K to extract uniform random bits from X as $K(X)$

Example. For $(X, Y) \equiv (X^n, Y^n)$

Rate of communication in step 1 $= H(X \mid Y) = H(X) - I(X \wedge Y)$

Rate of randomness extraction in step 2 $= H(X)$

The difference is the secret key capacity

Single-shot achievability?

Recall the two steps of SK agreement:

Step 1 (aka Information reconciliation).

Slepian-Wolf code to send X to Y

Step 2 (aka Randomness extraction or privacy amplification).

"Random function" K to extract uniform random bits from X as $K(X)$

Example. For $(X, Y) \equiv (X^n, Y^n)$

Rate of communication in step 1 $= H(X \mid Y) = H(X) - I(X \wedge Y)$

Rate of randomness extraction in step 2 $= H(X)$

The difference is the secret key capacity

Are we done?

Single-shot achievability?

Recall the two steps of SK agreement:

Step 1 (aka Information reconciliation).

Slepian-Wolf code to send X to Y

Step 2 (aka Randomness extraction or privacy amplification).

"Random function" K to extract uniform random bits from X as $K(X)$

Example. For $(X, Y) \equiv (X^n, Y^n)$

Rate of communication in step 1 $= H(X \mid Y) = H(X) - I(X \wedge Y)$

Rate of randomness extraction in step 2 $= H(X)$

The difference is the secret key capacity

Are we done? Not quite. Let's take a careful look

Step 1: Slepian-Wolf theorem

Miyake Kanaya 95, Han 03

Lemma (Slepian-Wolf coding)

There exists a code (e, d) *of size* M *with encoder* $e : \mathcal{X} \to \{1, ..., M\}$, *and a decoder* $d : \{1, ..., M\} \times \mathcal{Y} \to \mathcal{X}$, *such that*

$$\begin{aligned} &\mathrm{P}_{XY}\left(\{(x, y) \mid x \neq d(e(x), y)\}\right) \\ &\leq \mathrm{P}_{XY}\left(\{(x, y) \mid -\log \mathrm{P}_{X|Y}(x \mid y) \geq \log M - \gamma\}\right) + 2^{-\gamma}. \end{aligned}$$

Step 1: Slepian-Wolf theorem

Miyake Kanaya 95, Han 03

Lemma (Slepian-Wolf coding)

There exists a code (e, d) *of size* M *with encoder* $e : \mathcal{X} \to \{1, ..., M\}$, *and a decoder* $d : \{1, ..., M\} \times \mathcal{Y} \to \mathcal{X}$, *such that*

$$\begin{aligned} &\mathrm{P}_{XY}\left(\{(x, y) \mid x \neq d(e(x), y)\}\right) \\ &\leq \mathrm{P}_{XY}\left(\{(x, y) \mid -\log \mathrm{P}_{X|Y}(x \mid y) \geq \log M - \gamma\}\right) + 2^{-\gamma}. \end{aligned}$$

$$-\log \mathrm{P}_{X|Y} = -\log \mathrm{P}_X - \log(\mathrm{P}_{Y|X}/\mathrm{P}_Y)$$

Compare with

$$H(X|Y) = H(X) - I(X \wedge Y)$$

The second term is a proxy for the mutual information

Step 1: Slepian-Wolf theorem

Miyake Kanaya 95, Han 03

Lemma (Slepian-Wolf coding)

There exists a code (e, d) *of size* M *with encoder* $e : \mathcal{X} \to \{1, ..., M\}$, *and a decoder* $d : \{1, ..., M\} \times \mathcal{Y} \to \mathcal{X}$, *such that*

$$\begin{aligned} &\mathrm{P}_{XY}\left(\{(x, y) \mid x \neq d(e(x), y)\}\right) \\ &\leq \mathrm{P}_{XY}\left(\{(x, y) \mid \geq \log M - \gamma\}\right) + 2^{-\gamma}. \end{aligned}$$

$$-\log \mathrm{P}_{X|Y} = -\log \mathrm{P}_X - \log(\mathrm{P}_{Y|X}/\mathrm{P}_Y)$$

Compare with

$$H(X|Y) = H(X) - I(X \wedge Y)$$

The second term is a proxy for the mutual information

Communication rate needed is approximately equal to

(large probability upper bound on $-\log \mathrm{P}_X$) $- \log(\mathrm{P}_{Y|X}/\mathrm{P}_Y)$

Step 2: Leftover hash lemma

Lesson from the step 1: Communication rate is approximately

$$(\text{large probability upper bound on } -\log \mathrm{P}_X) - \log(\mathrm{P}_{Y|X}/\mathrm{P}_Y)$$

Recall that the *min entropy* of X is given by

$$H_{\min}(\mathrm{P}_X) = -\log \max_x \mathrm{P}_X(x)$$

Impagliazzo et. al. 89, Bennett et. al. 95, Renner-Wolf 05

Lemma (Leftover hash)

There exists a function K of X taking values in $\mathcal{K}$ such that

$$\|\mathrm{P}_{KZ} - \mathrm{P}_{\mathtt{unif}}\mathrm{P}_Z\| \leq \sqrt{|\mathcal{K}||\mathcal{Z}|2^{-H_{\min}(\mathrm{P}_X)}}$$

Step 2: Leftover hash lemma

Lesson from the step 1: Communication rate is approximately

(large probability upper bound on $-\log \mathrm{P}_X$) $- \log(\mathrm{P}_{Y|X}/\mathrm{P}_Y)$

Recall that the *min entropy* of X is given by

$$H_{\min}(\mathrm{P}_X) = -\log \max_x \mathrm{P}_X(x)$$

Impagliazzo et. al. 89, Bennett et. al. 95, Renner-Wolf 05

Lemma (Leftover hash)

There exists a function K of X taking values in $\mathcal{K}$ such that

$$\|\mathrm{P}_{KZ} - \mathrm{P}_{\mathtt{unif}}\mathrm{P}_Z\| \leq \sqrt{|\mathcal{K}||\mathcal{Z}|2^{-H_{\min}(\mathrm{P}_X)}}$$

Randomness can be extracted at a rate approximately equal to

(large probability lower bound on $-\log \mathrm{P}_X$)

Step 2: Leftover hash lemma

Lesson from the step 1: Communication rate is approximately

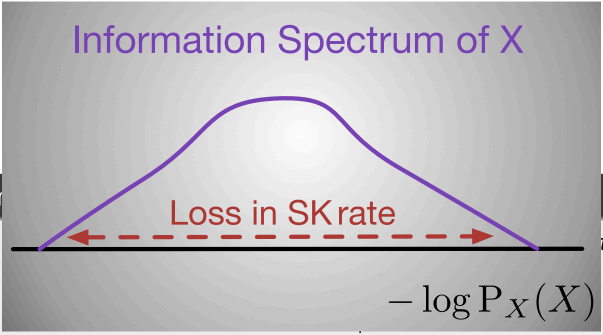
(large probability upper bound on $-\log \mathrm{P}_X$) $- \log(\mathrm{P}_{Y|X}/\mathrm{P}_Y)$

Recall

Impag

Lemm

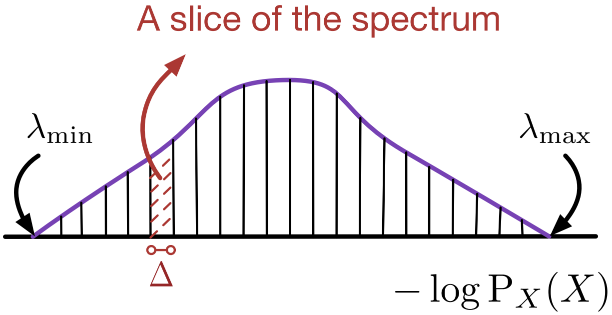
There *t*



Randomness can be extracted at a rate approximately equal to

(large probability lower bound on $-\log \mathrm{P}_X$)

Spectrum slicing



Slice the spectrum of X into L bins of length Δ and send the bin number to Y

Single-shot achievability

Theorem

For every $\gamma > 0$ and $0 \leq \lambda \leq \lambda_{\min}$, there exists an (ϵ, δ)-SK K taking values in $\mathcal{K}$ with

$$\epsilon \leq \mathrm{P}\left(\log \frac{\mathrm{P}_{XY}(X,Y)}{\mathrm{P}_X(X)\,\mathrm{P}_Y(Y)} \leq \lambda + \gamma + \Delta\right) + \mathrm{P}\left(-\log \mathrm{P}_X(X) \notin (\lambda_{\min}, \lambda_{\max})\right) + \frac{1}{L}$$

$$\delta \leq \frac{1}{2}\sqrt{|\mathcal{K}| 2^{-(\lambda - 2\log L)}}$$

Secret key capacity for general sources

Consider a sequence of sources (X_n, Y_n)

The SK capacity C is defined as

$$C \triangleq \sup_{\epsilon_n, \delta_n} \liminf_{n \to \infty} \frac{1}{n} S_{\epsilon_n, \delta_n}(X_n, Y_n)$$

where the $\sup$ is over all $\epsilon_n, \delta_n \geq 0$ such that

$$\lim_{n \to \infty} \epsilon_n + \delta_n = 0$$

Secret key capacity for general sources

Consider a sequence of sources (X_n, Y_n)

The SK capacity C is defined as

$$C \triangleq \sup_{\epsilon_n, \delta_n} \liminf_{n \to \infty} \frac{1}{n} S_{\epsilon_n, \delta_n}(X_n, Y_n)$$

where the $\sup$ is over all $\epsilon_n, \delta_n \geq 0$ such that

$$\lim_{n \to \infty} \epsilon_n + \delta_n = 0$$

The inf-mutual information rate $\underline{I}(\mathbf{X} \wedge \mathbf{Y})$ is defined as

$$\underline{I}(\mathbf{X} \wedge \mathbf{Y}) \triangleq \sup \left\{ \alpha \mid \quad \lim_{n \to \infty} \mathrm{P}(Z_n < \alpha) = 0 \right\}$$

where

$$Z_n = \frac{1}{n} \log \frac{\mathrm{P}_{X_n Y_n}(X_n, Y_n)}{\mathrm{P}_{X_n}(X_n)\, \mathrm{P}_{Y_n}(Y_n)}$$

General capacity

Theorem (Secret key capacity)

The SK capacity C for a sequence of sources $\{X_n, Y_n\}_{n=1}^{\infty}$ is given by

$$C = \underline{I}(\mathbf{X} \wedge \mathbf{Y})$$

General capacity

Theorem (Secret key capacity)

The SK capacity C for a sequence of sources $\{X_n, Y_n\}_{n=1}^{\infty}$ is given by

$$C = \underline{I}(\mathbf{X} \wedge \mathbf{Y})$$

Converse. Follows from our *conditional independence testing bound* with:

Lemma (Verdú)

For every ϵ_n such that

$$\lim_{n\to\infty} \epsilon_n = 0$$

it holds that

$$\liminf_{n} -\frac{1}{n} \log \beta_{\epsilon_n}\left(\mathrm{P}_{X_n Y_n}, \mathrm{P}_{X_n}\mathrm{P}_{Y_n}\right) \leq \underline{I}(\mathbf{X} \wedge \mathbf{Y})$$

General capacity

Theorem (Secret key capacity)

The SK capacity C for a sequence of sources $\{X_n, Y_n\}_{n=1}^{\infty}$ is given by

$$C = \underline{I}(\mathbf{X} \wedge \mathbf{Y})$$

Achievability. Use the single-shot construction with

$$\lambda_{\max} = n\left(\overline{H}(\mathbf{X}) + \Delta\right)$$

$$\lambda_{\min} = n\left(\underline{H}(\mathbf{X}) - \Delta\right)$$

$$\lambda = n\left(\underline{I}\left(\mathbf{X} \wedge \mathbf{Y}\right) - \Delta\right)$$

Towards characterizing finite-blocklength performance

We identify the second term in the asymptotic expansion of $S(X^n, Y^n)$:

Theorem (Second order asymptotics)

For every $0 < \epsilon < 1$ and IID RVs X^n, Y^n, we have

$$S_\epsilon\left(X^n, Y^n\right) = nI(X \wedge Y) - \sqrt{nV}Q^{-1}(\epsilon) + o(\sqrt{n})$$

The quantity V is given by

$$V = \mathrm{Var}\left[\log \frac{\mathrm{P}_{XY}\left(X, Y\right)}{\mathrm{P}_X\left(X\right)\mathrm{P}_Y\left(Y\right)}\right]$$

Towards characterizing finite-blocklength performance

We identify the second term in the asymptotic expansion of $S(X^n, Y^n)$:

Theorem (Second order asymptotics)

For every $0 < \epsilon < 1$ and IID RVs X^n, Y^n, we have

$$S_\epsilon\left(X^n, Y^n\right) = nI(X \wedge Y) - \sqrt{nV}Q^{-1}(\epsilon) + o(\sqrt{n})$$

The quantity V is given by

$$V = \mathrm{Var}\left[\log \frac{\mathrm{P}_{XY}\left(X, Y\right)}{\mathrm{P}_X\left(X\right)\mathrm{P}_Y\left(Y\right)}\right]$$

Proof relies on the use of Berry-Esseen theorem as in Polyanskiy-Poor-Verdu 10

Towards characterizing finite-blocklength performance

We identify the second term in the asymptotic expansion of $S(X^n, Y^n)$:

Theorem (Second order asymptotics)

For every $0 < \epsilon < 1$ and IID RVs X^n, Y^n, we have

$$S_\epsilon\left(X^n, Y^n\right) = nI(X \wedge Y) - \sqrt{nV}Q^{-1}(\epsilon) + o(\sqrt{n})$$

The quantity V is given by

$$V = \mathrm{Var}\left[\log \frac{\mathrm{P}_{XY}\left(X, Y\right)}{\mathrm{P}_X\left(X\right)\mathrm{P}_Y\left(Y\right)}\right]$$

Proof relies on the use of Berry-Esseen theorem as in Polyanskiy-Poor-Verdu 10

What about $S_{\epsilon,\delta}(X^n, Y^n)$?

Looking ahead ...

What if the eavesdropper has side information Z?

Best known converse bound on SK capacity due to Gohari-Ananthram 08

Recently we obtained a one-shot version of this bound

Tyagi and Watanabe, *Converses for Secret Key Agreement and Secure Computing*, preprint arXiv:1404.5715, 2014 - arxiv.org

Also, we have a single-shot achievability scheme that is asymptotically tight when X, Y, Z form a Markov chain