暗号理論にみる情報理論

定式化と安全性証明

岩本貢

電気通信大学

2024 年 3 月 7 日＠広島大学

謝辞
本研究は科研費補助金 JP23H00468, JP23H00479, JP23K17455, JP22H03590, JP23K21644
および JST CREST JPMJCR23M2 の補助を受けています

暗号理論と情報理論

- どちらも Shannon が最初に理論的論文を書いた (1945 [1949]/1948)
- 80 年代くらいまでは IEEE Trans. IT に多くの論文が載った
- それ以降は影響しあいつつ，互いの道を進んでいる
  - RSA (1977)
  - DES (1977/1979)
  - Secret Sharing (1979)
  - CRYPTO (1981) / EUROCRYPT (1982)
- 情報理論，理論計算機科学がベース
  - cf. ISIT (1950〜) , IEEE FOCS (1960〜), ACM STOC (1969〜)

本講演の目的：情報理論的暗号と計算量的暗号の狭間

情報理論的暗号理論

- 攻撃者の計算能力が無制限
- 確率論・組合せ論（情報理論）をベース

計算量的暗号理論

- 攻撃者の計算能力に制限（古典計算機・量子計算機）
- 理論計算機科学（計算量理論）がベース

本講演の目的・興味

- 両者の境界はどうなっているか，ちゃんと繋がっているか
- 特に，暗号理論における安全性の定義に情報理論がどのように関係するか

お話ししようと思っていること

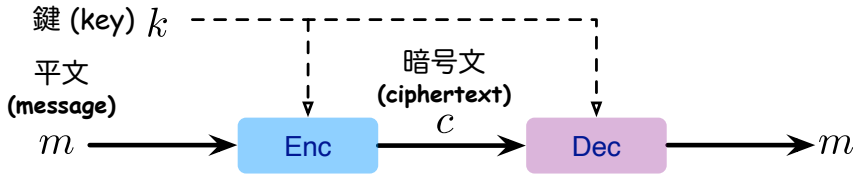
対象にする暗号とその安全性

- 共通鍵暗号
  - 識別不可能性と情報理論的安全性
  - M. Iwamoto, K. Ohta, and J. Shikata, “Security Formalizations and Their Relationships for Encryption and Key Agreement in Information-Theoretic Cryptography,” *IEEE Trans. Information Theory*, vol. 64, no. 1, pp.654–685, January 2018.
- マルチパーティ計算
  - シミュレーションベース安全性による情報理論的安全性
  - M. Iwamoto, “Information-Theoretic Perspectives for Simulation-Based Security in Multi-Party Computation,” *IEICE Trans. Fundamentals,* vol.E107–A, no.3, pp. 360–372, March 2024.

共通鍵暗号

情報理論的に安全な共通鍵暗号

共通鍵暗号



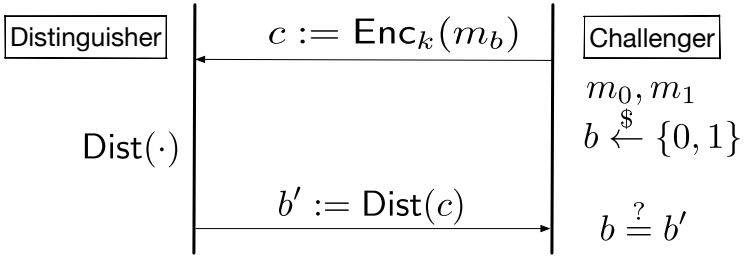
情報理論的安全性の定義の方法

- **情報漏洩**：$H(M \mid C) = H(M)$
- **確率的独立性**：$I(M;C) \leq \epsilon$

▷ 健全性（復号が正しく行われるか）は今日はあまり考慮しない

識別不可能性 (Indistinguishability)

識別ゲーム

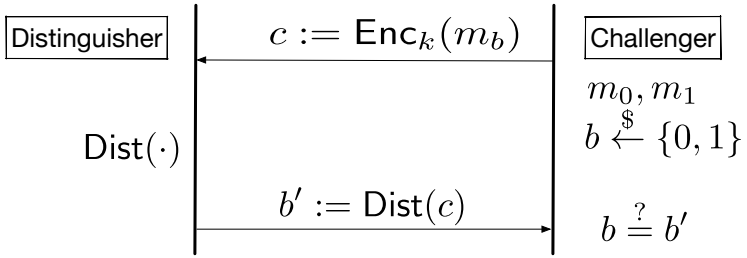


ゲームに勝つ確率と安全性

- $\forall m_0, \forall m_1, \max_{\mathsf{Dist}} \mathsf{Pr}\big(B = B'\big) = 1/2 + \epsilon$
- Dist : $\mathcal{M} \to \{0,1\}$ の選び方がポイント
  - ◊ 情報理論的安全性：任意の関数から選ぶ (Dist $\in \{0,1\}^{\mathcal{M}}$)← 候補は指数個ある
  - ◊ 計算量的安全性　：任意の PPTM (Probabilistic Polynomial-time Turing Machine) から選ぶ

情報理論的識別不可能性

識別ゲーム



識別ゲームに勝つ確率と情報理論的安全性

- $\forall m_0, \forall m_1, \max_{\mathsf{Dist} \in \{0,1\}^{\mathcal{M}}} \mathsf{Pr}\big(B = B'\big) = 1/2 + \epsilon$

  $\Leftrightarrow \forall m_0, \forall m_1,$

  $\max_{\mathsf{Dist} \in \{0,1\}^{\mathcal{M}}} |\mathsf{Pr}(\mathsf{Dist}(C) = 1 \mid B = 1) - \mathsf{Pr}(\mathsf{Dist}(C) = 1 \mid B = 0)| \leq \epsilon'$

情報理論的識別不可能性と変動（統計）距離

$$\triangleright\ \forall m_0, \forall m_1, \max_{\mathsf{Dist}\in\{0,1\}^{\mathcal{M}}} \mathsf{Pr}\big(B = B'\big) = 1/2 + \epsilon$$

$$\Leftrightarrow \forall m_0, \forall m_1,$$

$$\max_{\mathsf{Dist}\in\{0,1\}^{\mathcal{M}}} |\mathsf{Pr}(\mathsf{Dist}(C) = 1 \mid B = 1) - \mathsf{Pr}(\mathsf{Dist}(C) = 1 \mid B = 0)| \le \epsilon'$$

$$\Leftrightarrow \forall m_0, \forall m_1,$$

$$\max_{\mathsf{Dist}\in\{0,1\}^{\mathcal{M}}} |\mathsf{Pr}(\mathsf{Dist}(C_1) = 1) - \mathsf{Pr}(\mathsf{Dist}(C_0) = 1)| \le \epsilon' \ \text{ where } \ C_b := \mathsf{Enc}_{\mathsf{K}}(\mathsf{m}_\mathsf{b})$$

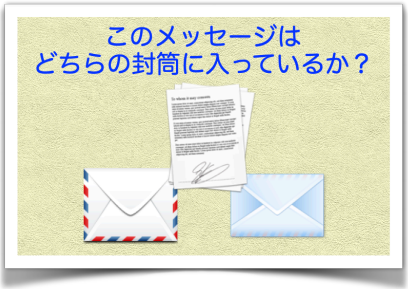
$$\Leftrightarrow \forall m_0, \forall m_1, \ \ d(C_1, C_0) \le \epsilon', \ \text{ where } \ d(X, Y) := \max_{\mathcal{B}\subseteq\mathcal{A}} |P_X(\mathcal{B}) - P_Y(\mathcal{B})|$$

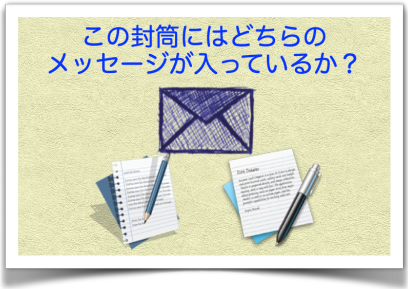
識別不可能性と統計的独立性

- 任意の平文について，それらの暗号文の分布は（ほぼ）等しい
  - ⇔ 暗号文の分布は平文に依存しない（cf. Resolvability, Han-Verdú (1993), Hayashi (2006)）
  - ⇔ 暗号文と平文は確率的に独立

条件は平文であるべき？暗号文であるべき？

例え話：難しいのはどっち？





- 左が Shannon の定式化，右が識別不可能性の定式化

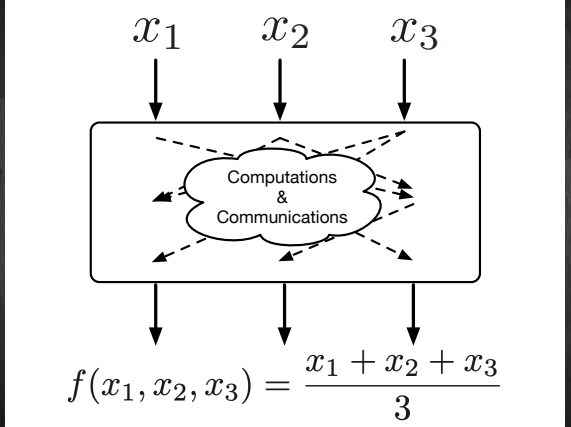
答え：$\epsilon = 0$ なら同じ，$\epsilon > 0$ なら左

- Iwamoto–Ohta–Shikata, IEEE-T-IT, 2018
- ギャップ証明には majorization を使う

マルチパーティ計算
Multi-Party Computation (MPC)

マルチパーティ計算とは?

例：試験の平均点



▷ Semi-honest な攻撃者を考える

健全性

- 平均点を計算できる

安全性

- 100 点満点で平均点は 80 点
  - ◊ よくできました！
  - ◊ …といった時点で何か漏れている
- 出力から例えば以下が分かる
  - ◊ 0 点の人はいなかった
  - ◊ 全員が 40 点以上であった
- MPC は出力から分かること以上の情報をもらさない

どうやって定義するか？

加算の MPC ($(2,3)$–しきい値法ベース)

▷ P_1, P_2 および P_3 がそれぞれ s_1, s_2, $\perp$ を入力

▷ $s := s_1 + s_2$ を計算したい

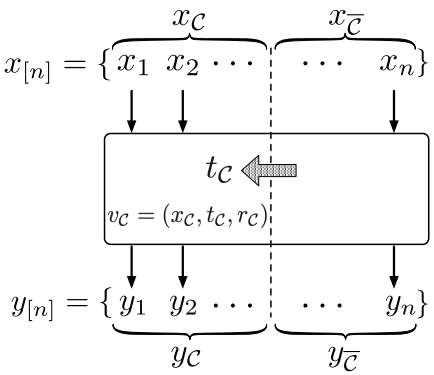
入力	P_1 (s_1 を入力)	P_2 (s_2 を入力)	P_3 (入力なし)
$f_1(x) = s_1 + r_1 x$	$f_1(1) = s_1 + r_1$	$f_1(2) = s_1 + 2r_1$	$f_1(3) = s_1 + 3r_1$
$f_2(x) = s_2 + r_2 x$	$f_2(1) = s_2 + r_2$	$f_2(2) = s_2 + 2r_2$	$f_2(3) = s_2 + 3r_2$
$f(x) = f_1(x) + f_2(x)$ $(s := s_1 + s_2, r = r_2 + r_2)$	$f(1) = s + r$	$f(2) = s + 2r$	$f(3) = s + 3r$
出力	$s = s_1 + s_2$	$s = s_1 + s_2$	$s = s_1 + s_2$

▷ $f(1), f(2), f(3)$ は $(2,3)$–しきい値法のシェア

▷ $s = s_1 + s_2$ が 3 人中任意の 2 人で復元できる（例えば P_1 と P_3）

MPC のモデルと View

プロトコルの流れ



$V_{\mathcal{C}}$：不正な参加者 $\mathcal{C}$ の View

- $\mathcal{C} \subseteq [n]$：不正な (Corrupted) 参加者
- 以下から成る確率変数の組：
  - $X_{\mathcal{C}}$：入力
  - $Y_{\mathcal{C}}$：出力
    $\Rightarrow W_{\mathcal{C}} := (X_{\mathcal{C}}, Y_{\mathcal{C}})$
  - $T_{\mathcal{C}}$: $\mathcal{C}$ の受け取る情報
  - $R_{\mathcal{C}}$: $\mathcal{C}$ が生成する一様乱数

安全性は「シミュレーション」で定義

今回の例での View

- P_i の View:
  - 入力と出力
  - P_i の受け取る情報
  - P_i が生成する一様乱数
- P_1 と P_3 の View をそれぞれ $\boxed{\text{view}}$, $\boxed{\text{view}}$ のように書く

入力	$\boxed{P_1}$ (s_1 を入力)	P_2 (s_2 を入力)	$\boxed{P_3}$ (入力なし)
$f_1(x) = s_1 + r_1 x$	$f_1(1) = \boxed{s_1} + \boxed{r_1}$	$f_1(2) = s_1 + 2r_1$	$f_1(3) = \boxed{s_1 + 3r_1}$
$f_2(x) = s_2 + r_2 x$	$f_2(1) = \boxed{s_2 + r_2}$	$f_2(2) = s_2 + 2r_2$	$f_2(3) = \boxed{s_2 + 3r_2}$
$f(x) = f_1(x) + f_2(x)$ $(s \coloneqq s_1 + s_2, r = r_2 + r_2)$	$f(1) = \boxed{s + r}$	$f(2) = s + 2r$	$f(3) = \boxed{s + 3r}$
出力	$s = \boxed{s_1 + s_2}$	$s = s_1 + s_2$	$s = \boxed{s_1 + s_2}$

シミュレーションベース安全性

シミュレーションベース安全性

- $V_{\mathcal{C}}(w_{[n]})$: $w_{[n]} \coloneqq (x_{[n]}, y_{[n]})$ を入出力としたときの $\mathcal{C}$ の View
- $S(w_{\mathcal{C}})$:
  $w_{\mathcal{C}} \coloneqq (x_{\mathcal{C}}, y_{\mathcal{C}})$ を用いて $\mathcal{C}$ の View をシミュレート（模倣）する確率変数

$$S(w_{\mathcal{C}}) \overset{\text{perf}}{\equiv} V_{\mathcal{C}}(w_{[n]})$$

直観的意味

- $w_{[n]}$ から生成された $\mathcal{C}$ の view（確率変数 $V_{\mathcal{C}}$）は
  - $V_{\mathcal{C}} = (W_{\mathcal{C}}, T_{\mathcal{C}}, R_{\mathcal{C}})$ は $\overline{\mathcal{C}}$ から受け取る情報 $T_{\mathcal{C}}$ を含む
- $w_{\mathcal{C}} \coloneqq (x_{\mathcal{C}}, y_{\mathcal{C}})$ を入力としてシミュレート（模倣）できる
  - $w_{\overline{\mathcal{C}}}$ を使わなくてもシミュレートできる

$\Rightarrow$ $V_{\mathcal{C}}$ から $w_{\overline{\mathcal{C}}}$ についての情報が得られない

条件付確率によるシミュレーションベース安全性の表現

シミュレーションベース安全性

- $V_{\mathcal{C}}(w_{[n]})$ は $w_{[n]} \coloneqq (x_{[n]}, y_{[n]})$ を入力とする
  確率的アルゴリズム（計算量無制限）でシミュレートできる
- $S(w_{\mathcal{C}})$：シミュレータの出力を表す確率変数

$$S(w_{\mathcal{C}}) \overset{\text{perf}}{\equiv} V_{\mathcal{C}}(w_{[n]})$$

条件付確率による表現

$$\forall v_{\mathcal{C}}, \forall w_{[n]},\ P_{V_{\mathcal{C}}|W_{\mathcal{C}}}(v_{\mathcal{C}}|w_{\mathcal{C}}) = P_{V_{\mathcal{C}}|W_{[n]}}(v_{\mathcal{C}}|w_{[n]})$$

直観的意味

- $w_{[n]} = (w_{\mathcal{C}}, w_{\overline{\mathcal{C}}})$ だから，$P_{V_{\mathcal{C}}|W_{\mathcal{C}}}(v_{\mathcal{C}}|w_{\mathcal{C}}) = P_{V_{\mathcal{C}}|W_{\mathcal{C}}W_{\overline{\mathcal{C}}}}(v_{\mathcal{C}}|w_{\mathcal{C}}w_{\overline{\mathcal{C}}})$ とも書ける
  - ◊ $w_{\mathcal{C}}$ を条件としたときに，$W_{\overline{\mathcal{C}}}$ と $V_{\mathcal{C}}$ は（条件付）独立
  - ⇔ $\mathcal{C}$ が $w_{\mathcal{C}}$ を知っているとき，$W_{\overline{\mathcal{C}}}$ に関する（それ以上の）情報は $V_{\mathcal{C}}$ から漏れない

Chapter 6

How to Simulate It – A Tutorial on the Simulation Proof Technique

Yehuda Lindell

Abstract One of the most fundamental notions of cryptography is that of *simulation*. It stands behind the concepts of semantic security, zero knowledge, and security for multiparty computation. However, writing a simulator and proving security via the use of simulation is a nontrivial task, and one that many newcomers to the field often find difficult. In this tutorial, we provide a guide to how to write simulators and prove security via the simulation paradigm. Although we have tried to make this tutorial as stand-alone as possible, we assume some familiarity with the notions of secure encryption, zero-knowledge, and secure computation.

以下，MPC のシミュレーションベース安全性について５つの等価な定義を示す

MPC における幾つかの安全性定義

Markov 連鎖による定義

条件付確率による定義（再掲）

$$\forall v_{\mathcal{C}}, \forall w_{[n]},\ P_{V_{\mathcal{C}}|W_{\mathcal{C}}}(v_{\mathcal{C}}|w_{\mathcal{C}}) = P_{V_{\mathcal{C}}|W_{[n]}}(v_{\mathcal{C}}|w_{[n]})$$

- ▷ これが $P_{V_{\mathcal{C}}|W_{\mathcal{C}}}(v_{\mathcal{C}}|w_{\mathcal{C}}) = P_{V_{\mathcal{C}}|W_{\mathcal{C}}W_{\overline{\mathcal{C}}}}(v_{\mathcal{C}}|w_{\mathcal{C}}w_{\overline{\mathcal{C}}})$ と書けることに注意すれば，上の関係は次と同値：

$$W_{\overline{\mathcal{C}}} \longleftrightarrow W_{\mathcal{C}} \longleftrightarrow V_{\mathcal{C}}$$

MPC の等価な安全性定義

任意の不正者集合 $\mathcal{C} \subseteq [n]$ を固定したとき，以下の (i)–(iii) は等価

(i) シミュレーションベース安全性

(ii) $\forall v_{\mathcal{C}}, \forall w_{[n]},\ P_{V_{\mathcal{C}}|W_{[n]}}(v_{\mathcal{C}}|w_{[n]}) = P_{V_{\mathcal{C}}|W_{\mathcal{C}}}(v_{\mathcal{C}}|w_{\mathcal{C}})$

(iii) $W_{\overline{\mathcal{C}}} \longleftrightarrow W_{\mathcal{C}} \longleftrightarrow V_{\mathcal{C}}$

情報漏洩量による表現

MPC の等価な安全性定義

任意の不正者集合 $\mathcal{C} \subseteq [n]$ を固定したとき，以下の (i)–(iv) は等価

(i) シミュレーションベース安全性

(ii) $\forall v_{\mathcal{C}}, \forall w_{[n]},\ P_{V_{\mathcal{C}}|W_{[n]}}(v_{\mathcal{C}}|w_{[n]}) = P_{V_{\mathcal{C}}|W_{\mathcal{C}}}(v_{\mathcal{C}}|w_{\mathcal{C}})$

(iii) $W_{\overline{\mathcal{C}}} \longleftrightarrow W_{\mathcal{C}} \longleftrightarrow V_{\mathcal{C}}$

(iv) $I(W_{\overline{\mathcal{C}}}; V_{\mathcal{C}}|W_{\mathcal{C}}) = 0$

(iv) の直観的意味

- $\mathcal{C}$ が $W_{\mathcal{C}}$ をもっているとき，
  $V_{\mathcal{C}}$ から $W_{\overline{\mathcal{C}}}$ について（$W_{\mathcal{C}}$ 以上に）漏れる情報は 0 ビット

シミュレーションベース安全性の５つの表現

MPC の安全性の等価な表現

任意の不正者集合 $\mathcal{C} \subseteq [n]$ を固定したとき，以下の (i)–(v) は等価

(i) シミュレーションベース安全性

(ii) $\forall v_{\mathcal{C}}, \forall w_{[n]},\ P_{V_{\mathcal{C}}|W_{[n]}}(v_{\mathcal{C}}|w_{[n]}) = P_{V_{\mathcal{C}}|W_{\mathcal{C}}}(v_{\mathcal{C}}|w_{\mathcal{C}})$

(iii) $W_{\overline{\mathcal{C}}} \longleftrightarrow W_{\mathcal{C}} \longleftrightarrow V_{\mathcal{C}}$

(iv) $I(W_{\overline{\mathcal{C}}}; V_{\mathcal{C}}|W_{\mathcal{C}}) = 0$

(v) $W_{\mathcal{C}} = \tau(V_{\mathcal{C}})$ を $V_{\mathcal{C}}$ の統計量と見なすと，$W_{\mathcal{C}}$ は $w_{\overline{\mathcal{C}}}$ に関する十分統計量

MPC における十分統計量

(v) 十分統計量による表現

- $W_{\mathcal{C}} = \tau(V_{\mathcal{C}})$ を $V_{\mathcal{C}}$ の統計量と見なすと，$W_{\mathcal{C}}$ は $w_{\overline{\mathcal{C}}}$ に関する十分統計量

▷ Markov 連鎖による特徴付け

◇ プロトコルに現れる確率変数の順番：
$W_{\overline{\mathcal{C}}} \longleftrightarrow V_{\mathcal{C}} \longrightarrow W_{\mathcal{C}}$，$(\because)\ I(W_{\mathcal{C}}; W_{\overline{\mathcal{C}}}|V_{\mathcal{C}}) = H(W_{\mathcal{C}}|V_{\mathcal{C}}) - H(W_{\mathcal{C}}|V_{\mathcal{C}}W_{\overline{\mathcal{C}}}) = 0$

◇ シミュレータの存在：$W_{\overline{\mathcal{C}}} \longleftrightarrow W_{\mathcal{C}} \longleftrightarrow V_{\mathcal{C}}$

(v) の直観的意味

- 被害者の入出力 $w_{\overline{\mathcal{C}}}$ を推定するには $V_{\mathcal{C}}$ の代わりに $W_{\mathcal{C}}$ を知っていれば十分

まとめ：何が嬉しいの？

【安全性の理解】

- ▷ 安全性は抽象概念なので，幾つもの定式化で納得できる方が良い

【安全性証明など】

- ▷ 数学的に違うテクニックで証明できる
  - ◇ 情報量：MPC の安全性を秘密分散と同様のテクニックで示せる
  - ◇ 十分統計量：Fisher の分解定理が使える
- ▷ 限界証明
  - ◇ 乱数長
  - ◇ シェアサイズ

Fisher の分解定理に基づく安全性証明

Fisher の分解定理

$\tau(\vec{X})$ は θ に関して十分 $\Longleftrightarrow$

- $g_\theta(\cdot)$: θ に依存する 関数
- $h(\cdot)$: θ に依存しない 関数

が存在し，$P_{\vec{X}|\Theta}(\vec{x}|\theta) = g_\theta(\tau(\vec{x}))h(\vec{x})$ が成立

MPC における分解定理

$W_{\mathcal{C}}$ は $w_{\overline{\mathcal{C}}}$ に関して十分（情報理論的に安全）$\Longleftrightarrow$

- $g_{w_{\overline{\mathcal{C}}}}(\cdot)$: $w_{\overline{\mathcal{C}}}$ に依存する 関数
- $h(\cdot)$: $w_{\overline{\mathcal{C}}}$ に依存しない 関数

が存在し $P_{V_{\mathcal{C}}|W_{\overline{\mathcal{C}}}}(v_{\mathcal{C}}|w_{\overline{\mathcal{C}}}) = g_{w_{\overline{\mathcal{C}}}}(w_{\mathcal{C}})h(v_{\mathcal{C}})$ が成立

References

1. R. Cramer, I. B. Damgård, J. B. Nielsen, “Secure Multiparty Computation and Secret Sharing,” Cambridge U. Press, 2015.
2. G. Asharov and Y. Lindell, “A Full Proof of the BGW Protocol for Perfectly Secure Multiparty Computation,” *J. of Cryptology*, vol. 30, issue 1, pp. 58–101, 2017.
3. Y. Lindell, “How to Simulate It — A Tutorial on the Simulation Proof Technique,” Tutorials on the Foundations of Cryptography pp. 277–346, 2017, Springer.
4. E. D. Karnin, J. W. Greene, and M. E. Hellman, “On secret sharing systems,” *IEEE Trans. Inform. Theory*, vol. 29, no. 1, pp. 35–41, 1983.
5. T. S. Han and S. Verdú, “Approximation Theory of Output Statistics,” *IEEE Trans. Inform. Theory*, vol. 39, no. 3, pp. 752–772, 1993.
6. M. Hayashi, “General Nonasymptotic and Asymptotic Formulas in Channel Resolvability and Identification Capacity and Their Application to the Wiretap Channel,” *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1562–1575, 2006.