A preliminary version of this paper appears in *Advances in Cryptology – CRYPTO ‘08*, Lecture Notes in Computer Science vol. xx, D. Wagner ed., Springer-Verlag, 2008. This is the full version.

Deterministic Encryption: Definitional Equivalences and Constructions without Random Oracles

MIHIR BELLARE[*] MARC FISCHLIN[†] ADAM O’NEILL[‡] THOMAS RISTENPART[§]

February 17, 2009

Abstract

We strengthen the foundations of deterministic public-key encryption via definitional equivalences and standard-model constructs based on general assumptions. Specifically we consider seven notions of privacy for deterministic encryption, including six forms of semantic security and an indistinguishability notion, and show them all equivalent. We then present a deterministic scheme for the secure encryption of uniformly and independently distributed messages based solely on the existence of trapdoor one-way permutations. We show a generalization of the construction that allows secure deterministic encryption of independent high-entropy messages. Finally we show relations between deterministic and standard (randomized) encryption.

[*] Dept. of Computer Science & Engineering 0404, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0404, USA. Email: `mihir@cs.ucsd.edu`. URL: `http://www.cs.ucsd.edu/~mihir`. Supported in part by NSF grants CNS 0524765 and CNS 0627779 and a gift from Intel Corporation.

[†] Dept. of Computer Science, Darmstadt University of Technology, Hochschulstrasse 10, 64289 Darmstadt, Germany. Email: `fischlin@informatik.tu-darmstadt.de`. URL: `http://www.fischlin.de`. Supported in part by the Emmy Noether Program Fi 940/2-1 of the German Research Foundation (DFG).

[‡] College of Computing, Georgia Institute of Technology, 801 Atlantic Drive, Atlanta, GA 30332, USA. Email: `amoneill@cc.gatech.edu`. URL: `http://www.cc.gatech.edu/~amoneill`. Supported in part by Alexandra Boldyreva’s NSF CAREER award 0545659.

[§] Dept. of Computer Science & Engineering 0404, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0404, USA. Email: `tristenp@cs.ucsd.edu`. URL: `http://www.cs.ucsd.edu/~tristenp`. Supported in part by the above-mentioned grants of the first author.

Contents

1 Introduction

The foundations of public-key encryption, as laid by Goldwasser and Micali [25] and their successors, involve two central threads. The first is definitional equivalences, which aim not only to increase our confidence that we have the "right" notion of privacy but also to give us definitions that are as easy to use in applications as possible. (Easy-to-use indistinguishability is equivalent to the more intuitive, but also more complex, semantic security [31, 25, 26, 23].) The second (of the two threads) is to obtain schemes achieving the definitions under assumptions as minimal as possible. In this paper we pursue these same two threads for *deterministic* encryption [3], proving definitional equivalences and providing constructions based on general assumptions.

DETERMINISTIC ENCRYPTION. A public-key encryption scheme is said to be deterministic if its encryption algorithm is deterministic. Deterministic encryption was introduced by Bellare, Boldyreva, and O'Neill [3]. The motivating application they gave is efficiently searchable encryption. Deterministic encryption permits logarithmic time search on encrypted data, while randomized encryption only allows linear time search [30, 12], meaning a search requires scanning the whole database. This difference is crucial for large outsourced databases which cannot afford to slow down search. Of course deterministic encryption cannot achieve the classical notions of security of randomized encryption, but [3] formalize a semantic security style notion PRIV that captures the "best possible" privacy achievable when encryption is deterministic, namely that an adversary provided with encryptions of plaintexts drawn from a message-space of high (super-logarithmic) min-entropy should have negligible advantage in computing any public-key independent *partial information function* of the plaintexts. The authors provide some schemes in the random-oracle (RO) model [4] meeting this definition but leave open the problem of finding standard model schemes.

The PRIV definition captures intuition well but is hard to work with. We would like to find simpler, alternative definitions of privacy for deterministic encryption —restricted forms of semantic security as well as an indistinguishablility style definition— that are equivalent to PRIV. We would also like to find schemes not only in the standard model but based on general assumptions.

NOTIONS CONSIDERED. We define seven notions of privacy for deterministic encryption inspired by the work of [20, 3]. These include a notion IND in the indistinguishability style and six notions —A-CSS, B-CSS, BB-CSS, A-SSS, B-SSS, BB-SSS— in the semantic-security style. The IND definition —adapted from [20]— asks that the adversary be unable to distinguish encryptions of plaintexts drawn from two, adversary-specified, high-entropy message spaces, and is simple and easy to use. The semantic security notions are organized along two dimensions. The first dimension is the class of partial information functions considered, and we look at three choices, namely arbitrary (A), boolean (B), or balanced boolean (BB). (A boolean function is balanced if the probabilities that it returns 0 or 1 are nearly the same.) The second dimension is whether the formalization is simulation (S) based or comparison (C) based.[1] The PRIV notion of [3] is A-CSS in our taxonomy. Low-end notions —think of BB as the lowest, then B then A and similarly C then S in the other dimension— are simpler and easier to use in applications, while high end ones are more intuitively correct. The question is whether the simplifications come at the price of power.

DEFINITIONAL EQUIVALENCES. We show that all seven notions discussed above are equivalent. The results are summarized in Figure 1. These results not only show that semantic security for boolean functions (predicates) is as powerful as semantic security for arbitrary functions, but (perhaps surprisingly) that one can even restrict attention to boolean functions that are balanced, meaning semantic security for balanced

[1] In the first case, A's success in computing partial information about plaintexts from ciphertexts is measured relative to that of a simulator, while in the second it is measured relative to A's own success when it is given the encryption of plaintexts independent of the challenge ones. The terminology is from [7] who prove equivalence between simulation and comparison based notions of non-malleability.

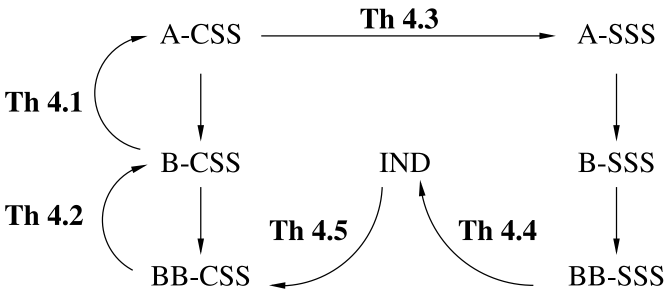

Figure 1: Notions of security for deterministic encryption schemes and implications showing that all seven notions are equivalent. An arrow $X \rightarrow Y$ means that every scheme secure under X is also secure under Y. Unlabeled implications are trivial.

boolean functions is as powerful as semantic security for arbitrary functions. We note that balance in this context originates with [20] but they only use it as a tool. We explicitly define and consider the notions BB-CSS and BB-SSS because they bear a natural and intuitive relation to IND and because we feel that the use made of balance by [20] indicates it is important. The proofs of our results rely on new techniques compared to [20, 17, 18].

DEFINITIONAL FRAMEWORK. We believe that an important and useful contribution of our paper is its definitional framework. Rather than an experiment per notion, we have a few core experiments and then use the approach of [5], capturing different notions via different adversary classes. Advantages of this approach are its easy extendability —for example we can capture the notions of [11] by simply introducing a couple of new adversary classes— and the ability to capture many definitional variants in a way that is unified, concise and yet precise.

A CONSTRUCTION FOR UNIFORM MESSAGES. Constructing a non-RO model deterministic encryption scheme meeting our strong notions of security seems like a very challenging problem. We are however able to make progress on certain special cases. We present a deterministic encryption scheme DE1 for the secure encryption of independent, uniformly distributed messages. The scheme is not only without random oracles but based on general trapdoor one-way permutations. To encrypt a random message x one iterates a trapdoor permutation f on x a number of times to get a point y. Let r denote the sequence of Goldreich-Levin [24] hardcore bits obtained in the process. Then one uses a standard IND-CPA scheme —which exists assuming trapdoor one-way permutations— to encrypt y with coins r. The interesting aspect of the scheme, and the source of the difficulty in analyzing it, is its cyclic nature, namely that the coins used for the IND-CPA encryption depend on the plaintext y that is IND-CPA encrypted. The proof manages to show that an adversary who, given y, can distinguish r from random can recover x *even though* this adversary may have partial information about the underlying seed x. The proof exploits in a crucial way that the equivalence between A-CSS and B-CSS holds even for uniformly and independently distributed messages.

ANOTHER PERSPECTIVE. A deterministic encryption scheme is (syntactically) the same thing as a family of injective trapdoor functions. Our notions can then be seen as an extension of the usual notion of one-wayness. Our construction is then a family of injective trapdoor functions which hides all (possible) partial information about its (randomly chosen) input. We believe this is a natural and useful strengthening of the usual notion of a trapdoor function that is fully achieved under standard assumptions in our work.

EFFICIENCY. The general assumption notwithstanding, our scheme admits efficient instantiations. For example with squaring as the trapdoor permutation [8] and Blum-Goldwasser [9] as the bare IND-CPA scheme, encryption and decryption come in at about double that of Blum-Goldwasser with no increase in ciphertext size. See Section 5.

A GENERALIZATION. We generalize our construction to obtain a non-RO model deterministic scheme DE2 for the encryption of independent, high min-entropy (but not necessarily uniform) plaintexts. The assumption used is that one has a trapdoor permutation that is one-way for high min-entropy distributions on its input. This increase in assumption strength is in some sense necessary, since deterministic encryption secure for some distribution trivially provides a one-way injective trapdoor function for that distribution.

FROM DETERMINISTIC TO RANDOMIZED ENCRYPTION. Another central foundational theme is relations between primitives, meaning determining which primitives imply which others. From this perspective we consider how to build IND-CPA-secure standard (randomized) encryption from PRIV-secure deterministic encryption. The obvious approach would be to use the deterministic encryption scheme as a trapdoor one-way function within some well-known general construction [24]. However, this approach leads to large ciphertexts, and we would hope to achieve better efficiency when using a primitive that provides more than one-wayness. We provide a much more efficient construction using a hybrid encryption-style approach, in which the deterministic scheme encrypts a fresh session key padded with extra randomness and the session key is used to encrypt the message. See Section 7 for the details.

CCA. Lifting our notions and equivalences to the CCA setting is straightforward; see Appendix G. Our above-mentioned construction of a randomized encryption scheme from a deterministic one works even in the CCA setting. This means, in particular, that we can generically build witness-recovering IND-CCA encryption schemes [27] from arbitrary CCA-secure deterministic schemes. (Witness-recovering encryption allows, during decryption, recovery of all randomness used to generate a ciphertext.) CCA-secure witness-recovering encryption is of use in further applications [16], and only very recently was a (not very efficient) standard-model construction produced [27]. Our construction shows that building CCA-secure deterministic schemes is at least as hard as building witness-recovering probabilistic encryption.

RELATED WORK. Dodis and Smith's work on entropic security [20] has in common with ours the consideration of privacy for messages of high min-entropy. But there are important differences in the settings, namely that theirs is information-theoretic and symmetric while ours is computational and public-key. Dodis and Smith [20] introduce definitions that in our framework are IND, B-SSS, and BB-SSS, to complement the A-SSS-like information-theoretic notion originally proposed by Russell and Wang [28]. Also, Desrosiers [17] and Desrosiers and Dupuis [18] subsequently treat quantum entropic security, providing notions similar to our framework's B-CSS and A-CSS. These works provide some relations between the notions they define. While some of their techniques and implications lift to our setting, others do not. The salient fact that emerges is that prior work *does not* imply equivalence of all seven notions we consider. In particular, the BB-SSS and BB-CSS notions are not considered in [17, 18] and Dodis and Smith [20] only provide reductions for BB-SSS implying A-SSS that result in inefficient or restricted adversaries. See Appendix H for more information.

Another setting that deals with high min-entropy messages is that of perfectly one-way hash functions (POWHF), introduced by Canetti [13] and further studied by Canetti, Micciancio, and Reingold [14]. These are randomized hash functions that produce publically-verifiable outputs. Our definitions and equivalences can be adapted to the POWHF setting.

INDEPENDENT WORK. In concurrent and independent work, Boldyreva, Fehr, and O'Neill [11] consider a relaxation of PRIV in which message sequences need to not merely have high entropy but each message must have high entropy even given the others. They prove some relations between their notions using techniques of [20, 17, 18] but do not consider as many notions as us and in particular do not consider balance. Their schemes achieve stronger notions of security then our DE1 but at the cost of specific algebraic assumptions as opposed to our general one. Combining their results with ours shows that our DE2 achieves their notion of security while using a general (even though non-standard) assumption.

2 Preliminaries

NOTATION AND CONVENTIONS. If x is a string then $|x|$ denotes its length; if x is a number then $|x|$ denotes its absolute value; if S is a set then $|S|$ denotes its size. We denote by λ the empty string. If S is a set then $X \leftarrow\!\!\$\; S$ denotes that X is selected uniformly at random from S. We let $x[i \ldots j]$ denote bits i through j of string x, for $1 \leq i \leq j \leq |x|$. By $x_1 \parallel \cdots \parallel x_n$ we denote the concatenation of $x_1, \ldots, x_n$. Vectors are denoted in boldface, e.g. $\mathbf{x}$. If $\mathbf{x}$ is a vector then $|\mathbf{x}|$ denotes the number of components of $\mathbf{x}$ and $\mathbf{x}[i]$ denotes its i^{th} component for $1 \leq i \leq |\mathbf{x}|$. If $i \geq 1$ is an integer, we use B_i as shorthand for $\{0,1\}^i$. By $\langle a, b \rangle$ we denote the inner product modulo 2 of equal-length strings a, b.

We write $\alpha \leftarrow\!\!\$\; X(x, y, \ldots)$ to denote running X on inputs $(x, y, \ldots)$ with fresh random coins and assigning the result to α. We let $[X(x, y, \ldots)]$ denote the set of possible outputs of X when run on $x, y, \ldots \in \{0,1\}^*$. An algorithm X is *non-uniform* if its first input is 1^k and there is a collection $\{C_k\}_{k \in \mathbb{N}}$ of (randomized) circuits such that C_k computes $X(1^k, \ldots)$. The running time is the circuit size. A function f is called *negligible* if it approaches zero faster than the reciprocal of any polynomial, that is, for any polynomial p, there exists $n_p \in \mathbb{N}$ such that $f(n) \leq 1/p(n)$ for all $n \geq n_p$. "PT" stands for polynomial time. We denote by Λ the algorithm that on any inputs returns λ.

PUBLIC-KEY ENCRYPTION. A *public-key encryption (PKE)* scheme $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ is a triple of PT algorithms. The key generation algorithm $\mathcal{K}$ takes input 1^k, where $k \in \mathbb{N}$ is the security parameter, and outputs a public-key, secret-key pair (pk, sk). The encryption algorithm $\mathcal{E}$ takes inputs 1^k, pk, and plaintext $x \in \{0,1\}^*$ and outputs a ciphertext. The deterministic decryption algorithm $\mathcal{D}$ takes inputs 1^k, sk, and ciphertext y and outputs either a plaintext x or $\perp$. We say that Π is *deterministic* if $\mathcal{E}$ is deterministic. If $\mathbf{x}$ is a vector of plaintexts, then we write $\mathbf{y} \leftarrow\!\!\$\; \mathcal{E}(1^k, pk, \mathbf{x})$ to denote component-wise encryption of $\mathbf{x}$, i.e. $\mathbf{y}[i] \leftarrow\!\!\$\; \mathcal{E}(1^k, pk, \mathbf{x}[i])$ for all $1 \leq i \leq |\mathbf{x}|$.

3 Security Notions for Deterministic PKE

We first provide formal definitions and then discuss them.

SEMANTIC SECURITY. An *SS-adversary* $A = (A_{\mathrm{c}}, A_{\mathrm{m}}, A_{\mathrm{g}})$ is a tuple of non-uniform algorithms. A_{c} takes as input a unary encoding 1^k of the security parameter $k \in \mathbb{N}$ and returns a string st representing some state information. A_{m} takes input 1^k and st, and returns a vector of challenge messages $\mathbf{x}$ together with a test string t that represents some information about $\mathbf{x}$. A_{g} takes 1^k, a public key and the component-wise encryption of $\mathbf{x}$ under this key, and tries to compute t. The running time of A is defined as the sum of the running times of $A_{\mathrm{c}}, A_{\mathrm{m}}, A_{\mathrm{g}}$, so that A is PT if $A_{\mathrm{c}}, A_{\mathrm{m}}, A_{\mathrm{g}}$ are all PT.

Let $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ be a PKE scheme, $A = (A_{\mathrm{c}}, A_{\mathrm{m}}, A_{\mathrm{g}})$ an SS-adversary, and S a simulator (a non-uniform algorithm). Let $k \in \mathbb{N}$. Figure 2 displays the css (comparison-based semantic security) and sss (simulation-based semantic security) experiments. We define the css advantage and sss advantage of A by

$$\mathbf{Adv}^{\mathrm{css}}_{\Pi,A}(k) = 2 \cdot \Pr\left[\, \mathbf{Exp}^{\mathrm{css}}_{\Pi,A}(k) \Rightarrow \mathsf{true} \,\right] - 1 \text{ , and} \quad (1)$$

$$\mathbf{Adv}^{\mathrm{sss}}_{\Pi,A,S}(k) = 2 \cdot \Pr\left[\, \mathbf{Exp}^{\mathrm{sss}}_{\Pi,A,S}(k) \Rightarrow \mathsf{true} \,\right] - 1 \text{ .} \quad (2)$$

Our approach to defining the six notions of semantic security of interest to us is to associate to each a

```
Experiment $\mathbf{Exp}^{\mathrm{css}}_{\Pi,A}(k)$
$b \leftarrow_{\$} \{0,1\}$ ; $st \leftarrow_{\$} A_{\mathrm{c}}(1^k)$
$(\mathbf{x}_0, t_0) \leftarrow_{\$} A_{\mathrm{m}}(1^k, st)$
$(\mathbf{x}_1, t_1) \leftarrow_{\$} A_{\mathrm{m}}(1^k, st)$
$(pk, sk) \leftarrow_{\$} \mathcal{K}(1^k)$
$\mathbf{c} \leftarrow_{\$} \mathcal{E}(1^k, pk, \mathbf{x}_b)$
$g \leftarrow_{\$} A_{\mathrm{g}}(1^k, pk, \mathbf{c}, st)$
If $g = t_1$ then $b' \leftarrow 1$
Else $b' \leftarrow 0$
Ret $(b' = b)$
```

```
Experiment $\mathbf{Exp}^{\mathrm{sss}}_{\Pi,A,S}(k)$
$b \leftarrow_{\$} \{0,1\}$ ; $st \leftarrow_{\$} A_{\mathrm{c}}(1^k)$
$(\mathbf{x}, t) \leftarrow_{\$} A_{\mathrm{m}}(1^k, st)$
$(pk, sk) \leftarrow_{\$} \mathcal{K}(1^k)$
$\mathbf{c} \leftarrow_{\$} \mathcal{E}(1^k, pk, \mathbf{x})$
If $b = 1$ then
    $g \leftarrow_{\$} A_{\mathrm{g}}(1^k, pk, \mathbf{c}, st)$
Else $g \leftarrow_{\$} S(1^k, pk, st)$
If $g = t$ then $b' \leftarrow 1$
Else $b' \leftarrow 0$
Ret $(b' = b)$
```

```
Experiment $\mathbf{Exp}^{\mathrm{ind}}_{\Pi,I}(k)$
$b \leftarrow_{\$} \{0,1\}$ ; $st \leftarrow_{\$} I_{\mathrm{c}}(1^k)$
$\mathbf{x}_b \leftarrow_{\$} I_{\mathrm{m}}(1^k, b, st)$
$(pk, sk) \leftarrow_{\$} \mathcal{K}(1^k)$
$\mathbf{c} \leftarrow_{\$} \mathcal{E}(1^k, pk, \mathbf{x}_b)$
$b' \leftarrow_{\$} I_{\mathrm{g}}(1^k, pk, \mathbf{c}, st)$
Ret $(b' = b)$
```

Figure 2: Three experiments for defining security of encryption schemes.

corresponding class of adversaries and ask that the advantage of any adversary in this class be negligible. We proceed to define the relevant classes.

An SS-adversary $A = (A_{\mathrm{c}}, A_{\mathrm{m}}, A_{\mathrm{g}})$ is *legitimate* if there exists a function $v(\cdot)$, called the number of messages, and a collection $\{\mathbf{y}_k\}_{k \in \mathbb{N}}$ of *reference* message-vectors such that the following three conditions hold. First, $|\mathbf{x}| = v(k)$ for all $(\mathbf{x}, t) \in [A_{\mathrm{m}}(1^k, st)]$ and all $st \in \{0,1\}^*$. Second, $|\mathbf{x}[i]| = |\mathbf{y}_k[i]|$ for all $(\mathbf{x}, t) \in [A_{\mathrm{m}}(1^k, st)]$, all $st \in \{0,1\}^*$, and all $1 \le i \le v(k)$. Third, the function

$$\nu(k) = \Pr\Big[\, \mathrm{eq}(\mathbf{x}, \mathbf{y}_k) = 0 \;:\; st \leftarrow_{\$} A_{\mathrm{c}}(1^k)\,;\, (\mathbf{x}, t) \leftarrow_{\$} A_{\mathrm{m}}(1^k, st) \,\Big]$$

is negligible, where

$$\mathrm{eq}(\mathbf{x}, \mathbf{y}_k) = \begin{cases} 1 & \text{if } \forall i, j : \ \mathbf{x}[i] = \mathbf{x}[j] \text{ iff } \mathbf{y}_k[i] = \mathbf{y}_k[j] \\ 0 & \text{otherwise.} \end{cases} \qquad (3)$$

(The third condition reflects that every deterministic scheme leaks plaintext equality.) Let $\mathcal{A}_{\mathrm{SS}}$ be the set of all legitimate, PT SS-adversaries. We say that A has *trivial state function* if $A_{\mathrm{c}} = \Lambda$. Let $\mathcal{A}_{\lambda}$ be the set of all SS-adversaries with trivial state functions.

Without loss of generality (through suitable padding) we can assume there is a function $\ell(\cdot)$ such that the output of $A_{\mathrm{g}}(1^k, \cdot, \cdot)$ and any test string t output by $A_{\mathrm{m}}(1^k, \cdot)$ always have length $\ell(k)$. We call ℓ the *information length* of A. An SS-adversary $A = (A_{\mathrm{c}}, A_{\mathrm{m}}, A_{\mathrm{g}}) \in \mathcal{A}_{\mathrm{SS}}$ is *boolean* if it has information length $\ell(\cdot) = 1$. Let $\mathcal{A}_{\mathrm{B}} \subseteq \mathcal{A}_{\mathrm{SS}}$ be the class of all boolean SS-adversaries. A boolean SS-adversary $A = (A_{\mathrm{c}}, A_{\mathrm{m}}, A_{\mathrm{g}}) \in \mathcal{A}_{\mathrm{B}}$ is δ-balanced if for every st we have

$$\left| \Pr\Big[\, t = 0 \;:\; (\mathbf{x}, t) \leftarrow_{\$} A_{\mathrm{m}}(1^k, st) \,\Big] - \frac{1}{2} \right| \le \delta \,. \qquad (4)$$

When $\delta = 0$ we say that A is *perfectly balanced*. We say that A is *balanced* if it is δ-balanced for some $\delta < 1/2$. Let $\mathcal{A}^{\delta}_{\mathrm{BB}} \subseteq \mathcal{A}_{\mathrm{B}}$ be the class of all δ-balanced boolean SS-adversaries. An SS-adversary $A = (A_{\mathrm{c}}, A_{\mathrm{m}}, A_{\mathrm{g}}) \in \mathcal{A}_{\mathrm{SS}}$ has min-entropy μ if

$$\Pr\Big[\, \mathbf{x}[i] = x \;:\; (\mathbf{x}, t) \leftarrow_{\$} A_{\mathrm{m}}(1^k, st) \,\Big] \;\le\; 2^{-\mu(k)}$$

for all $k \in \mathbb{N}$, all $1 \le i \le v(k)$, all $x \in \{0,1\}^*$, and all $st \in \{0,1\}^*$. Let $\mathcal{A}^{\mu}_{\mathrm{ME}} \subseteq \mathcal{A}_{\mathrm{SS}}$ be the class of all SS-adversaries with min-entropy μ. We say that A has *high min-entropy* if it is in $\mathcal{A}^{\mu}_{\mathrm{ME}}$ for some $\mu(k) \in \omega(\log k)$. Let $\mathcal{A}_{\mathrm{HE}} \subseteq \mathcal{A}_{\mathrm{SS}}$ be the class of all SS-adversaries that have high min-entropy.

Let Π be a PKE scheme. We say that Π is A-CSS secure if $\mathbf{Adv}^{\mathrm{css}}_{\Pi,A}(\cdot)$ is negligible for all $A \in$

$\mathcal{A}_{\rm HE} \cap \mathcal{A}_\lambda$; Π is B-CSS-secure if $\mathbf{Adv}^{\rm css}_{\Pi,A}(\cdot)$ is negligible for all $A \in \mathcal{A}_{\rm HE} \cap \mathcal{A}_\lambda \cap \mathcal{A}_{\rm B}$; and Π is BB-CSS-secure if there exists $\delta < 1/2$ such that $\mathbf{Adv}^{\rm css}_{\Pi,A}(\cdot)$ is negligible for all $A \in \mathcal{A}_{\rm HE} \cap \mathcal{A}_\lambda \cap \mathcal{A}^{\delta}_{\rm BB}$.

Similarly, we say that Π is A-SSS-secure if for all $A \in \mathcal{A}_{\rm HE} \cap \mathcal{A}_\lambda$ there exists a PT simulator S such that $\mathbf{Adv}^{\rm sss}_{\Pi,A,S}(\cdot)$ is negligible; Π is B-SSS-secure if for all $A \in \mathcal{A}_{\rm HE} \cap \mathcal{A}_\lambda \cap \mathcal{A}_{\rm B}$ there exists a PT simulator S such that $\mathbf{Adv}^{\rm sss}_{\Pi,A,S}(\cdot)$ is negligible; and Π is BB-SSS-secure if there exists $\delta < 1/2$ such that for all $A \in \mathcal{A}_{\rm HE} \cap \mathcal{A}_\lambda \cap \mathcal{A}^{\delta}_{\rm BB}$ there exists a PT simulator S such that $\mathbf{Adv}^{\rm sss}_{\Pi,A,S}(\cdot)$ is negligible.

The *message space* of an SS-adversary $A = (A_{\rm c}, A_{\rm m}, A_{\rm g})$ is the algorithm $A_{\rm d}$ that on input $1^k, st$ lets $(\mathbf{x}, t) \leftarrow_{\$} A_{\rm m}(1^k, st)$ and returns $\mathbf{x}$. An SS-adversary is said to produce independent messages if the coordinates of $\mathbf{x}$ are independently distributed when $\mathbf{x} \leftarrow_{\$} A_{\rm d}(1^k, st)$ for all k, st. Let $\mathcal{A}_\times$ be the class of all SS-adversaries which produce independent messages.

For each $d \in \{0,1\}$, we let $\mathbf{Exp}^{\text{css-d}}_{\Pi,A}(k)$ be the same as $\mathbf{Exp}^{\rm css}_{\Pi,A}(k)$ except that the first line sets $b \leftarrow d$ rather than picking b at random. We similarly define $\mathbf{Exp}^{\text{sss-d}}_{\Pi,A,S}(k)$. A standard argument gives

$$\mathbf{Adv}^{\rm css}_{\Pi,A}(k) \;=\; \Pr\left[\,\mathbf{Exp}^{\text{css-1}}_{\Pi,A}(k) \Rightarrow \mathsf{true}\,\right] - \Pr\left[\,\mathbf{Exp}^{\text{css-0}}_{\Pi,A}(k) \Rightarrow \mathsf{false}\,\right] \quad \text{and} \tag{5}$$

$$\mathbf{Adv}^{\rm sss}_{\Pi,A,S}(k) \;=\; \Pr\left[\,\mathbf{Exp}^{\text{sss-1}}_{\Pi,A,S}(k) \Rightarrow \mathsf{true}\,\right] - \Pr\left[\,\mathbf{Exp}^{\text{sss-0}}_{\Pi,A,S}(k) \Rightarrow \mathsf{false}\,\right]\ . \tag{6}$$

INDISTINGUISHABILITY. An *IND-adversary* $I = (I_{\rm c}, I_{\rm m}, I_{\rm g})$ is a tuple of non-uniform algorithms. $I_{\rm c}$ takes as input 1^k and returns a string st representing some state information. $I_{\rm m}$ takes input 1^k, a bit b, and st, and returns a vector of messages $\mathbf{x}$. $I_{\rm g}$ takes 1^k, a public key, the component-wise encryption of $\mathbf{x}$ under this key, and st and tries to compute the bit b. The running time of I is defined as the sum of the running times of $I_{\rm c}, I_{\rm m}, I_{\rm g}$, so that I is PT if $I_{\rm c}, I_{\rm m}, I_{\rm g}$ are all PT.

Let $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ be a PKE scheme, $I = (I_{\rm c}, I_{\rm m}, I_{\rm g})$ an IND-adversary and $k \in \mathbb{N}$. Figure 2 displays the ind experiment. We define the ind advantage of I by

$$\mathbf{Adv}^{\rm ind}_{\Pi,I}(k) \;=\; 2 \cdot \Pr\left[\,\mathbf{Exp}^{\rm ind}_{\Pi,I}(k) \Rightarrow \mathsf{true}\,\right] - 1\ . \tag{7}$$

We next define classes of IND-adversaries. An IND-adversary $I = (I_{\rm c}, I_{\rm m}, I_{\rm g})$ is *legitimate* if there exists a function $v(\cdot)$, called the number of messages, and a collection $\{\mathbf{y}_k\}_{k \in \mathbb{N}}$ of *reference* message-vectors such that the following three conditions hold. First, $|\mathbf{x}| = v(k)$ for all $\mathbf{x} \in [I_{\rm m}(1^k, b, st)]$, all $b \in \{0,1\}$, and all $st \in \{0,1\}^*$. Second, $|\mathbf{x}[i]| = |\mathbf{y}_k[i]|$ for all $\mathbf{x} \in [I_{\rm m}(1^k, b, st)]$, all $b \in \{0,1\}$, all $st \in \{0,1\}^*$, and all $1 \le i \le v(k)$. Third, the function

$$\nu(k) = \Pr\left[\,\mathrm{eq}(\mathbf{x}, \mathbf{y}_k) = 0\ :\ st \leftarrow_{\$} I_{\rm c}(1^k)\ ;\ b \leftarrow_{\$} \{0,1\}\ ;\ \mathbf{x} \leftarrow_{\$} I_{\rm m}(1^k, b, st)\,\right]$$

is negligible, where $\mathrm{eq}(\mathbf{x}, \mathbf{y}_k)$ was defined by (3). Let $\mathcal{I}$ be the set of all legitimate, polynomial time IND-adversaries. We say that I has *trivial state function* if $I_{\rm c} = \Lambda$. Let $\mathcal{I}_\lambda \subseteq \mathcal{I}$ be the set of all IND-adversaries with trivial state functions. An IND-adversary $I = (I_{\rm c}, I_{\rm m}, I_{\rm g}) \in \mathcal{I}$ has *min-entropy* μ if

$$\Pr\left[\,\mathbf{x}[i] = x\ :\ \mathbf{x} \leftarrow_{\$} I_{\rm m}(1^k, b, st)\,\right] \;\le\; 2^{-\mu(k)}$$

for all $k \in \mathbb{N}$, all $b \in \{0,1\}$, all $1 \le i \le v(k)$, all $x \in \{0,1\}^*$, and all $st \in \{0,1\}^*$. Let $\mathcal{I}^{\mu}_{\rm ME} \subseteq \mathcal{I}$ be the class of all IND-adversaries with min-entropy μ. We say I has *high min-entropy* if it is in $\mathcal{I}^{\mu}_{\rm ME}$ for some $\mu(k) \in \omega(\log k)$. Let $\mathcal{I}_{\rm HE}$ be the class of all IND-adversaries that have high min-entropy. We say that Π is IND-secure if $\mathbf{Adv}^{\rm ind}_{\Pi,I}(\cdot)$ is negligible for all $I \in \mathcal{I}_{\rm HE} \cap \mathcal{I}_\lambda$.

For each $d \in \{0,1\}$, we let $\mathbf{Exp}^{\text{ind-d}}_{\Pi,I}(k)$ be the same as $\mathbf{Exp}^{\rm ind}_{\Pi,I}(k)$ except that the first line sets $b \leftarrow d$ rather than picking b at random. A standard argument gives

$$\mathbf{Adv}^{\rm ind}_{\Pi,A}(k) \;=\; \Pr\left[\,\mathbf{Exp}^{\text{ind-1}}_{\Pi,I}(k) \Rightarrow \mathsf{true}\,\right] - \Pr\left[\,\mathbf{Exp}^{\text{ind-0}}_{\Pi,I}(k) \Rightarrow \mathsf{false}\,\right]\ . \tag{8}$$

DISCUSSION. A-CSS is exactly the PRIV definition of [3]. As discussed in [3], it is important that $A_{\rm m}$ does not take input the public key, and this carries over to $I_{\rm m}$. In the classical setting a standard hybrid argument [2] shows that the security of encrypting one message implies the security of encrypting multiple messages. In the deterministic encryption setting this is not true in general, which is why $A_{\mathrm{m}}, I_{\mathrm{m}}$ output vectors of messages.

Following [3], message spaces are not explicit but rather implicitly defined by their PT sampling algorithms A_{m} and I_{m}. As a consequence, message spaces are PT sampleable.

Following [3], the partial information function is not explicit. Think of t as its value on $\mathbf{x}$. This is more general because t is allowed to depend on coins underlying the generation of $\mathbf{x}$ rather than merely on $\mathbf{x}$ itself. (This is stronger than merely allowing the function to be randomized, which is standard.) It allows us in particular to capture "history." However, we show in Appendix A that this formulation is equivalent to one where the partial information is computed as a function of the message. Note that the (implicit or explicit) partial information functions are PT.

Our security definitions quantify only over adversaries with trivial state functions. We do this for compatibility with [3, 20]. So why introduce the common state function at all? The reason is that it is useful in proofs. Indeed, [20] use such a function implicitly in many places. We believe making it explicit increases clarity. In the end we can always hardwire a "best" state and thereby end up with an adversary in $\mathcal{A}_\lambda$.

4 Relating the Security Notions

In this section and its supporting appendices we justify the implications summarized by Figure 1. The implications given by the unlabeled arrows are trivial and can be justified by the fact that $X \rightarrow Y$ whenever the adversary class corresponding to Y is a subset of the one corresponding to X. We focus on the implications: A-CSS ⇒ A-SSS; BB-SSS ⇒ IND; IND ⇒ BB-CSS; BB-CSS ⇒ B-CSS; and B-CSS ⇒ A-CSS.

Theorem 4.1 [B-CSS ⇒ A-CSS] Let $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ be a PKE scheme. Let $A = (A_{\mathrm{c}}, A_{\mathrm{m}}, A_{\mathrm{g}}) \in \mathcal{A}^{\mu}_{\mathrm{ME}} \cap \mathcal{A}_\lambda$ be an SS-adversary having information length $\ell(\cdot)$. Then there exists a boolean SS-adversary $A' = (A'_{\mathrm{c}}, A'_{\mathrm{m}}, A'_{\mathrm{g}}) \in \mathcal{A}^{\mu}_{\mathrm{ME}} \cap \mathcal{A}_\lambda \cap \mathcal{A}_{\mathrm{B}}$ such that for all $k \in \mathbb{N}$

$$\mathbf{Adv}^{\mathrm{css}}_{\Pi,A}(k) \leq 2 \cdot \mathbf{Adv}^{\mathrm{css}}_{\Pi,A'}(k) \; . \tag{9}$$

A' has the same message space as A and its running time is that of A plus $\mathcal{O}(\ell)$. □

Proof: The proof is from [20] and repeated here in order to provide intuition for Theorem 4.2. Below we write ℓ for $\ell(k)$. Then let

```
algorithm $A^*_{\mathrm{c}}(1^k)$:
    $r \leftarrow\$ \{0,1\}^\ell$
    $s \leftarrow\$ \{0,1\}$
    Ret $(r, s)$

algorithm $A^*_{\mathrm{m}}(1^k, (r, s))$:
    $(\mathbf{x}, t) \leftarrow\$ A_{\mathrm{m}}(1^k, \lambda)$
    Ret $(\mathbf{x}, \langle r, t\rangle \oplus s))$

algorithm $A^*_{\mathrm{g}}(1^k, pk, \mathbf{c}, (r, s))$:
    $g \leftarrow\$ A_{\mathrm{g}}(1^k, pk, \mathbf{c}, \lambda)$
    Ret $\langle r, g\rangle \oplus s$
```

Then $A^* = (A^*_{\mathrm{c}}, A^*_{\mathrm{m}}, A^*_{\mathrm{g}})$ is certainly boolean, and

$$\begin{aligned} P_{A^*}(k) &= P_A(k) + \frac{1}{2}\left[1 - P_A(k)\right] \\ Q_{A^*}(k) &= Q_A(k) + \frac{1}{2}\left[1 - Q_A(k)\right] \end{aligned}$$

where $P_X(k) = \Pr\left[\, \mathbf{Exp}^{\mathrm{css}\text{-}1}_{\Pi,X}(k) \Rightarrow \mathsf{true} \,\right]$ and $Q_X(k) = \Pr\left[\, \mathbf{Exp}^{\mathrm{css}\text{-}0}_{\Pi,X}(k) \Rightarrow \mathsf{false} \,\right]$. Subtracting, we get $\mathbf{Adv}^{\mathrm{css}}_{\Pi,A^*}(k) = \frac{1}{2} \cdot \mathbf{Adv}^{\mathrm{css}}_{\Pi,A}(k)$. We are not done yet because A^* does not have trivial state function. Let A' be obtained from A^* by hardwiring in a "best" choice of r, s and we are done. ▮

Now we wish to show that BB-CSS $\Rightarrow$ B-CSS. Note that if the adversary A' constructed in the proof of Theorem 4.1 were balanced, we would be done. But, A' need not be balanced. Dodis and Smith [20] give a partial solution to this problem, showing that it is in fact possible to find an r that, when hardwired into A^*, results in a balanced adversary, as long as $p \leq \epsilon^2/4$, where p is the maximum probability of any t being output by A_{m} and $\epsilon = \mathbf{Adv}^{\mathrm{css}}_{\Pi,A}(\cdot)$.

We will remove this restriction by proceeding as follows. Let A be a given SS-adversary, which from Theorem 4.1 we can assume is boolean (but not balanced). We again construct an A^* with non-trivial state, but this will consist of n independently chosen keys $\mathbf{K}[1], \ldots, \mathbf{K}[n]$ for a family of pair-wise independent hash functions H. Then $A^*_{\mathrm{m}}(1^k, \mathbf{K})$ first runs $(\mathbf{x}, t) \leftarrow\$\, A_{\mathrm{m}}(1^k, \lambda)$ and then returns $(\mathbf{x}, H(\mathbf{K}[i], t))$ for random $i \in \{1, \ldots, n\}$, while $A^*_{\mathrm{g}}(1^k, pk, \mathbf{c}, \mathbf{K})$ picks its own independent random j and returns $H(\mathbf{K}[j], A_{\mathrm{g}}(1^k, pk, \mathbf{c}, \lambda))$. Our analysis will show that for a suitable choice of n there exists a choice of the vector $\mathbf{K}$ which, when hardwired into A^*, yields an adversary A' having all the claimed properties. The theorem is below and the proof is in Appendix B.

Theorem 4.2 **[BB-CSS $\Rightarrow$ B-CSS]** Let $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ be a PKE scheme. Let $A = (A_{\mathrm{c}}, A_{\mathrm{m}}, A_{\mathrm{g}}) \in \mathcal{A}^{\mu}_{\mathrm{ME}} \cap \mathcal{A}_{\lambda} \cap \mathcal{A}_{\mathrm{B}}$ be a boolean SS-adversary. Let $\epsilon(\cdot) = \mathbf{Adv}^{\mathrm{css}}_{\Pi,A}(\cdot) > 0$ and let $\delta = 1/4$. Then there exists an SS-adversary $A' = (A'_{\mathrm{c}}, A'_{\mathrm{m}}, A'_{\mathrm{g}}) \in \mathcal{A}^{\mu}_{\mathrm{ME}} \cap \mathcal{A}_{\lambda} \cap \mathcal{A}^{\delta}_{\mathrm{BB}}$ such that for all $k \in \mathbb{N}$

$$\mathbf{Adv}^{\mathrm{css}}_{\Pi,A}(k) \leq 4n(k) \cdot \mathbf{Adv}^{\mathrm{css}}_{\Pi,A'}(k) \ ,$$

where $n(k) = \max\{485 \ , \ \lceil 64 \cdot \ln(1/\epsilon(k)) + 64 \ln 4 \rceil\}$. A' has the same message space as A and its running time is that of A plus $\mathcal{O}(\log(1/\epsilon(k)) + k)$. $\square$

Below are theorem statements for the other three implications. The proofs are found in Appendices C, D, and E, respectively.

Theorem 4.3 **[A-CSS $\Rightarrow$ A-SSS]** Let $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ be a PKE scheme. Let $A = (A_{\mathrm{c}}, A_{\mathrm{m}}, A_{\mathrm{g}}) \in \mathcal{A}^{\mu}_{\mathrm{ME}} \cap \mathcal{A}_{\lambda}$ be an SS-adversary outputting at most v messages. Then there exists a simulator S such that for all $k \in \mathbb{N}$

$$\mathbf{Adv}^{\mathrm{sss}}_{\Pi,A,S}(k) \leq \mathbf{Adv}^{\mathrm{css}}_{\Pi,A}(k) \ .$$

The running time of S is that of A plus the time to perform v encryptions. $\square$

Theorem 4.4 **[BB-SSS $\Rightarrow$ IND]** Let $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ be a PKE scheme. Let $I = (I_{\mathrm{c}}, I_{\mathrm{m}}, I_{\mathrm{g}}) \in \mathcal{I}^{\mu}_{\mathrm{ME}} \cap \mathcal{I}_{\lambda}$ be an IND-adversary. Let $\delta = 0$. Then there exists an SS-adversary $A = (A_{\mathrm{c}}, A_{\mathrm{m}}, A_{\mathrm{g}}) \in \mathcal{A}^{\mu}_{\mathrm{ME}} \cap \mathcal{A}_{\lambda} \cap \mathcal{A}^{\delta}_{\mathrm{BB}}$ such that for any simulator S and all $k \in \mathbb{N}$

$$\mathbf{Adv}^{\mathrm{ind}}_{\Pi,I}(k) \leq 2 \cdot \mathbf{Adv}^{\mathrm{sss}}_{\Pi,A,S}(k) \ .$$

The running time of A is that of I. $\square$

Theorem 4.5 **[IND $\Rightarrow$ BB-CSS]** Let $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ be a PKE scheme. Let $0 \leq \delta < 1/2$ and let $A = (A_{\mathrm{c}}, A_{\mathrm{m}}, A_{\mathrm{g}}) \in \mathcal{A}^{\mu}_{\mathrm{ME}} \cap \mathcal{A}_{\lambda} \cap \mathcal{A}^{\delta}_{\mathrm{BB}}$ be an SS-adversary. Then there exists an ind-adversary $I = (I_{\mathrm{c}}, I_{\mathrm{m}}, I_{\mathrm{g}}) \in \mathcal{I}^{\nu}_{\mathrm{ME}} \cap \mathcal{I}_{\lambda}$ such that for all $k \in \mathbb{N}$

$$\mathbf{Adv}^{\mathrm{css}}_{\Pi,A}(k) \leq 2 \cdot \mathbf{Adv}^{\mathrm{ind}}_{\Pi,I}(k) + 2^{-k} \ .$$

I has min-entropy $\nu(k) = \mu(k) - 1 + \log(1 - 2\delta)$ and its running time is that of A plus the time for $\lceil -(\log(2/(1+2\delta)))^{-1} \rceil (k+3) + 1$ executions of A_{m}. $\square$

5 Deterministic Encryption from Trapdoor Permutations

We construct a deterministic encryption scheme, without ROs, that meets our definitions in the case that the messages being encrypted are uniformly and independently distributed. It is based on the existence of

```
**algorithm** $\mathcal{K}(1^k)$:
$(\phi, \tau) \leftarrow_{\$} G(1^k)$
$s \leftarrow_{\$} \{0,1\}^k$
$(\overline{pk}, \overline{sk}) \leftarrow_{\$} \overline{\mathcal{K}}(1^k)$
$pk \leftarrow (\phi, \overline{pk}, s)$
$sk \leftarrow (\tau, \overline{sk})$
Ret $(pk, sk)$
```

```
**algorithm** $\mathcal{E}(1^k, pk, x)$:
$(\phi, \overline{pk}, s) \leftarrow pk$
$y \leftarrow F_{\phi}^{n(k)}(x)$
$\omega \leftarrow \mathcal{G}(1^k, 1^{n(k)}, \phi, x, s)$
$c \leftarrow \overline{\mathcal{E}}(1^k, \overline{pk}, y \,;\, \omega)$
Ret $c$
```

```
**algorithm** $\mathcal{D}(1^k, sk, c)$:
$(\tau, \overline{sk}) \leftarrow sk$
$y \leftarrow \overline{\mathcal{D}}(1^k, \overline{sk}, c)$
$x \leftarrow \overline{F}_{\tau}^{n(k)}(y)$
Ret $x$
```

Figure 3: Algorithms defining our deterministic encryption scheme $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$.

trapdoor permutations. In Appendix 6 we generalize the construction to independently distributed messages of high min-entropy μ, but under the (stronger and non-standard) assumption of the existence of trapdoor permutations that are one-way under all input distributions of min entropy μ.

PRIMITIVES. A family of trapdoor permutations $\mathcal{TP} = (G, F, \overline{F})$ is a triple of PT algorithms, with the last two being deterministic. On input 1^k, the key generation algorithm G returns a pair (ϕ, τ) of strings such that $F_{\phi}(\cdot) = F(\phi, \cdot)$ is a permutation on $\{0,1\}^k$ and $\overline{F}_{\tau}(\cdot) = \overline{F}(\tau, \cdot)$ is its inverse. If $f\colon \{0,1\}^k \rightarrow \{0,1\}^k$ then $f^i\colon \{0,1\}^k \rightarrow \{0,1\}^k$ is defined inductively by $f^0(x) = x$ and $f^{i+1}(x) = f(f^i(x))$ for $i \geq 0$ and $x \in \{0,1\}^k$. The Blum-Micali-Yao [10, 31], Goldreich-Levin [24] generator $\mathcal{G}_{\mathcal{TP}}$ takes input $1^k, 1^n, \phi$ and $x, s \in B_k$ and returns

$$\langle F_{\phi}^0(x), s\rangle \parallel \langle F_{\phi}^1(x), s\rangle \parallel \cdots \parallel \langle F_{\phi}^{n-1}(x), s\rangle \,.$$

To discuss the security of our scheme, we say that an SS-adversary is uniform if for every k and every st the components of $\mathbf{x}$ are uniformly and independently distributed over $\{0,1\}^k$ when $(\mathbf{x}, t) \leftarrow_{\$} A_{\mathrm{m}}(1^k, st)$. We let $\mathcal{A}_{\mathrm{UN}}$ be the class of all uniform SS-adversaries. If $f : B_k \rightarrow B_k$ then $f(\mathbf{x})$ denotes the vector whose i^{th} component is $f(\mathbf{x}[i])$. We let $\mathcal{G}_{\mathcal{TP}}(1^k, 1^n, \phi, \mathbf{x}, s)$ be the vector whose i^{th} component is $\mathcal{G}_{\mathcal{TP}}(1^k, 1^n, \phi, \mathbf{x}[i], s)$.

THE CONSTRUCTION. We fix a (randomized) encryption scheme $\overline{\Pi} = (\overline{\mathcal{K}}, \overline{\mathcal{E}}, \overline{\mathcal{D}})$. Assume that $\overline{\mathcal{E}}(1^k, \cdot, \cdot)$ draws its coins from $\{0,1\}^{n(k)}$, and write $\overline{\mathcal{E}}(1^k, pk, x \,;\, \omega)$ for the execution of $\overline{\mathcal{E}}$ on inputs $1^k, pk, x$ and coins ω. Let $\mathcal{TP} = (G, F, \overline{F})$ be a family of trapdoor permutations and $\mathcal{G}_{\mathcal{TP}}$ the associated generator. Our deterministic encryption scheme $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ is defined as shown in Figure 3. We refer to it as $\mathsf{DE1}$.

INTUITION. A weird aspect of our scheme is that one is encrypting, under the standard scheme $\overline{\mathcal{E}}$, a message y under coins ω that are related to y. The challenge is to show that this works assuming $\mathcal{TP}$ is one-way and $\overline{\Pi}$ is IND-CPA. So let $A = (A_{\mathrm{c}}, A_{\mathrm{m}}, A_{\mathrm{g}}) \in \mathcal{A}_{\mathrm{UN}} \cap \mathcal{A}_{\lambda}$ be an adversary with associated information length $\ell(\cdot)$ and number of messages $v(\cdot)$ that is successful in violating the A-CSS security of Π. It is not hard to see that the assumed security of $\overline{\Pi}$ allows us to reduce our task to showing that it is hard for a PT adversary D to have a non-negligible advantage in computing the challenge bit b in the following distinguishing game. The game generates $\phi, \tau, \overline{pk}, \overline{sk}, s$ as done by $\mathcal{K}(1^k)$ and lets $(\mathbf{x}, t) \leftarrow_{\$} A_{\mathrm{m}}(1^k, \lambda)$. It lets

$$\boldsymbol{\omega}_1 \leftarrow \mathcal{G}_{\mathcal{TP}}(1^k, 1^{n(k)}, \phi, \mathbf{x}, s) \quad \text{and} \quad \boldsymbol{\omega}_0 \leftarrow_{\$} B_{n(k)}^{v(k)} \,,$$

picks a random challenge bit b, and provides the adversary D with ϕ, s, $F_{\phi}^{n(k)}(\mathbf{x})$, $\boldsymbol{\omega}_b$, and t. Now, D's task would be merely the standard (and known to be hard) one of breaking the pseudorandomness of $\mathcal{G}_{\mathcal{TP}}$ (meaning, we would be done) but for one catch, namely that D has "help" information t about the seed(s) $\mathbf{x}$. If we could somehow remove it we would be done, but this seems hard to do directly. Instead, we first produce from D an adversary I' that solves (although still with help) a computational (rather than

```
Experiment $\mathbf{Exp}^{\mathrm{owf}}_{\mathcal{TP},J}(k)$
$(\phi,\tau) \leftarrow_{\$} G(1^k)$ ; $st \leftarrow_{\$} J_{\mathrm{c}}(1^k,\phi)$
$x \leftarrow_{\$} \{0,1\}^k$ ; $t \leftarrow_{\$} J_{\mathrm{p}}(1^k,x,\phi,st)$
$y \leftarrow F_{\phi}(x)$ ; $x' \leftarrow_{\$} J_{\mathrm{s}}(1^k,\phi,st,y,t)$
Ret $(x = x')$
```

```
Experiment $\mathbf{Exp}^{\mathrm{prg}\text{-}v}_{\mathcal{TP},D,n}(k)$
$(\phi,\tau) \leftarrow_{\$} G(1^k)$ ; $st \leftarrow_{\$} D_{\mathrm{c}}(1^k,\phi)$
$\mathbf{x} \leftarrow_{\$} B_k^{v(k)}$ ; $s \leftarrow_{\$} \{0,1\}^k$ ; $d \leftarrow_{\$} \{0,1\}$
$t \leftarrow_{\$} D_{\mathrm{p}}(1^k,\mathbf{x},\phi,st)$
$\boldsymbol{\omega}_1 \leftarrow \mathcal{G}_{\mathcal{TP}}(1^k,1^{n(k)},\phi,\mathbf{x},s)$
$\boldsymbol{\omega}_0 \leftarrow_{\$} B_{n(k)}^{v(k)}$
$d' \leftarrow_{\$} D_{\mathrm{g}}(1^k,\phi,st,F_{\phi}^{n(k)}(\mathbf{x}),\boldsymbol{\omega}_d,s,t)$
Ret $(d = d')$
```

Figure 4: **(Left)** Experiment defining one-wayness of $\mathcal{TP} = (G,F,\overline{F})$. **(Right)** Experiment defining pseudorandomness of $\mathcal{G}_{\mathcal{TP}}$.

decision) problem, namely that of inverting F_ϕ: given ϕ, $F_\phi(x)$, and $\ell(\cdot)$ bits of information about x, our adversary computes x. This is obtained by noting that the Goldreich-Levin [24] and Blum-Micali-Yao [10, 31] proof of pseudorandomness of $\mathcal{G}_{\mathcal{TP}}$ based on the one-wayness of $\mathcal{TP}$ generalizes to say that $\mathcal{G}_{\mathcal{TP}}$ remains pseudorandom in the presence of $\ell(\cdot)$ bits of help information about the seed assuming $\mathcal{TP}$ is one-way in the presence of $\ell(\cdot)$ bits of help information about the input. Now we need to turn I' into an adversary succeeding at the same task, but without help. We appeal to Theorem 4.1, which allows us to assume our starting adversary A was boolean, meaning $\ell(\cdot) = 1$. In this case it is easy to dispense with the help provided to I because we can try both values of it and lower our success probability by at most a factor of 2.

We remark that we have made crucial use of the fact that the adversary constructed by Theorem 4.1 has the same message space as the original one. This means that if the latter is in $\mathcal{A}_{\mathrm{UN}}$ then so is the former, so that B-CSS for uniform adversaries implies A-CSS for uniform adversaries. We now proceed to the full proof.

OWPS AND PRGS WITH HELP. For our proof, we will need to extend the usual frameworks of one-wayness and pseudorandomness to adversaries with "help." An inversion adversary $J = (J_{\mathrm{c}}, J_{\mathrm{p}}, J_{\mathrm{s}})$ is a triple of non-uniform algorithms. If $\mathcal{TP} = (G,F,\overline{F})$ is a family of trapdoor permutations we let

$$\mathbf{Adv}^{\mathrm{owf}}_{\mathcal{TP},J}(k) = \Pr\left[\,\mathbf{Exp}^{\mathrm{owf}}_{\mathcal{TP},J}(k) \Rightarrow \mathsf{true}\,\right]$$

where the experiment is shown in Figure 4. The running time of J is defined as the sum of the running times of J_{c} and J_{s}, so that J is PT if $J_{\mathrm{c}}, J_{\mathrm{s}}$ are PT. (J_{p} is not required to be PT.) We say that J has help-length $\ell(\cdot)$ if the output of $J_{\mathrm{p}}(1^k,\cdot,\cdot,\cdot)$ is always of length $\ell(k)$. We say that J is unaided if it has help length $\ell(\cdot) = 0$. We let $\mathcal{J}_\ell$ denote the class of all PT inversion adversaries with help length $\ell(\cdot)$. We say $\mathcal{TP}$ is one-way for help-length $\ell(\cdot)$ if $\mathbf{Adv}^{\mathrm{owf}}_{\mathcal{TP},J}(\cdot)$ is negligible for all $J \in \mathcal{J}_\ell$. We say that $\mathcal{TP}$ is one-way if it is one-way for help-length $\ell(\cdot) = 0$. The following, although trivial, will be very useful.

Proposition 5.1 Let $\mathcal{TP}$ be a family of trapdoor permutations and J an inversion adversary with help-length $\ell(\cdot)$. Then there is an inversion adversary J' with help-length 0 such that

$$\mathbf{Adv}^{\mathrm{owf}}_{\mathcal{TP},J}(k) \le 2^{\ell(k)} \cdot \mathbf{Adv}^{\mathrm{owf}}_{\mathcal{TP},J'}(k)$$

for all k, and the running time of J' is that of J plus $\mathcal{O}(\ell)$. □

Proof: Let $J = (J_{\mathrm{c}}, J_{\mathrm{p}}, J_{\mathrm{s}})$ and $J' = (J_{\mathrm{c}}, \Lambda, J'_{\mathrm{s}})$ where $J'_{\mathrm{s}}(1^k,\phi,st,y,\lambda)$ lets $t \leftarrow_{\$} \{0,1\}^{\ell(k)}$ and returns $J_{\mathrm{s}}(1^k,\phi,st,y,t)$. ▌

A PRG adversary $D = (D_{\mathrm{c}}, D_{\mathrm{p}}, D_{\mathrm{g}})$ is a triple of non-uniform algorithms. If $\mathcal{TP} = (G, F, \overline{F})$ is a family of trapdoor permutations and $\mathcal{G}_{\mathcal{TP}}$ is the corresponding generator we let

$$\mathbf{Adv}^{\text{prg-}v}_{\mathcal{TP},D,n}(k) = 2 \cdot \Pr\left[\,\mathbf{Exp}^{\text{prg-}v}_{\mathcal{TP},D,n}(k) \Rightarrow \mathsf{true}\,\right] - 1$$

where the experiment is shown in Figure 4 and $v(\cdot), n(\cdot) \colon\ \mathbb{N} \to \mathbb{N}$. The running time of D is defined as the sum of the running times of D_{c} and D_{g}, so that D is PT if $D_{\mathrm{c}}, D_{\mathrm{g}}$ are PT. (D_{p} is not required to be PT.) We say that D has help-length $\ell(\cdot)$ if the output of $D_{\mathrm{p}}(1^k, \cdot, \cdot, \cdot)$ is always of length $\ell(k)$. We let $\mathcal{D}_\ell$ denote the class of all PT PRG-adversaries with help length $\ell(\cdot)$. We say $\mathcal{G}_{\mathcal{TP}}$ is pseudorandom for help-length $\ell(\cdot)$ if $\mathbf{Adv}^{\text{prg-}v}_{\mathcal{TP},D,n}(\cdot)$ is negligible for all $D \in \mathcal{D}_\ell$ and all polynomials v, n. We say that $\mathcal{G}_{\mathcal{TP}}$ is pseudorandom if it is pseudorandom for help-length $\ell(\cdot) = 0$. We remark that it is important that D_{p} does not get s as input, meaning the help information is only about x. The following says that if $\mathcal{TP}$ is one-way for help-length $\ell(\cdot)$ then $\mathcal{G}_{\mathcal{TP}}$ is pseudorandom for help-length $\ell(\cdot)$. The case $\ell(\cdot) = 0$ is the standard result [10, 31, 24] saying that $\mathcal{G}_{\mathcal{TP}}$ is pseudorandom if $\mathcal{TP}$ is one-way. The proof of the following is in Appendix F.

Lemma 5.2 Let $\mathcal{TP} = (G, F, \overline{F})$ be a family of trapdoor permutations. Let $v(\cdot)$, $n(\cdot)$ be polynomials. Let D be a PRG-adversary with help-length $\ell(\cdot)$ and let $\epsilon(\cdot) = \mathbf{Adv}^{\text{prg-}v}_{\mathcal{TP},D,n}(\cdot) > 0$. Then there is an inversion adversary J with help-length $\ell(\cdot)$ such that

$$\epsilon(k) \leq 4n(k)v(k) \cdot \mathbf{Adv}^{\mathrm{owf}}_{\mathcal{TP},J}(k)$$

and the running time of J is

$$T_J = \mathcal{O}(k^3 n^4 v^4 \epsilon^{-4}) + \mathcal{O}(T_D + nvT_F)k^2 n^2 v^2 \epsilon^{-2}\ ,$$

where T_X is the running time of X. $\square$

IND-CPA. Associate to (randomized) encryption scheme $\overline{\Pi} = (\overline{\mathcal{K}}, \overline{\mathcal{E}}, \overline{\mathcal{D}})$ and adversary B the experiment $\mathbf{Exp}^{\text{ind-cpa}}_{\overline{\Pi},B}(k)$ defined by

$$b \leftarrow_{\$} \{0,1\}\ ;\ (\overline{pk}, \overline{sk}) \leftarrow_{\$} \overline{\mathcal{K}}(1^k)\ ;\ b' \leftarrow_{\$} B^{\overline{\mathcal{E}}_{\overline{pk}}(\mathrm{LR}(\cdot,\cdot,b))}(\overline{pk})\ ;\ \text{Ret } (b = b')$$

where $\mathrm{LR}(M_0, M_1, b) = M_b$. B is an IND-CPA adversary if all its oracle queries consist of equal length strings. Let

$$\mathbf{Adv}^{\text{ind-cpa}}_{\overline{\Pi},B}(k) = 2 \cdot \Pr\left[\,\mathbf{Exp}^{\text{ind-cpa}}_{\overline{\Pi},B}(k) \Rightarrow \mathsf{true}\,\right] - 1\ .$$

We say that $\overline{\Pi}$ is IND-CPA secure if $\mathbf{Adv}^{\text{ind-cpa}}_{\overline{\Pi},B}(\cdot)$ is negligible for all PT IND-CPA adversaries B.

SECURITY OF OUR SCHEME. The following says that our scheme is B-CSS secure for uniform adversaries assuming $\mathcal{TP}$ is one-way and $\overline{\Pi}$ is IND-CPA secure. By Theorem 4.1 it is A-CSS secure for uniform adversaries under the same assumptions and a constant factor loss in security. Since the existence of one-way trapdoor permutations implies the existence of IND-CPA secure encryption schemes we obtain the results under the sole assumption of the existence of one-way trapdoor permutations.

Theorem 5.3 Let $\mathcal{TP} = (G, F, \overline{F})$ be a family of trapdoor permutations and $\overline{\Pi} = (\overline{\mathcal{K}}, \overline{\mathcal{E}}, \overline{\mathcal{D}})$ an encryption scheme. Let $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ be the associated deterministic encryption scheme as per our construction above. Let $A = (A_{\mathrm{c}}, A_{\mathrm{m}}, A_{\mathrm{c}}) \in \mathcal{A}_{\mathrm{B}} \cap \mathcal{A}_{\lambda} \cap \mathcal{A}_{\mathrm{UN}}$ be an SS-adversary against Π with advantage $\epsilon(\cdot) = \mathbf{Adv}^{\mathrm{css}}_{\Pi,A}(\cdot) > 0$ and number of messages $v(\cdot)$. Then there is an unaided inversion adversary J and an IND-CPA adversary B such that for all $k \in \mathbb{N}$

$$\epsilon(k) \leq 2 \cdot \mathbf{Adv}^{\text{ind-cpa}}_{\overline{\Pi},B}(k) + 16n(k)v(k) \cdot \mathbf{Adv}^{\mathrm{owf}}_{\mathcal{TP},J}(k)\ . \tag{10}$$

The running time of B is that of A plus $\mathcal{O}(nT_F + T_{\mathcal{G}})$ and it makes $v(k)$ oracle queries. The running time of J is

$$\mathcal{O}(k^3 n^4 v^4 \epsilon^{-4}) + \mathcal{O}(T_A + T_{\overline{\mathcal{E}}} + T_{\overline{\mathcal{K}}} + nvT_F)\cdot k^2 n^2 v^2 \epsilon^{-2} \tag{11}$$

where T_X is the running time of X. $\Box$

Proof: Consider the experiments of Figure 5. There $\overline{\mathcal{E}}(1^k, \overline{pk}, \mathbf{y}\ ;\ \boldsymbol{\omega})$ is the vector whose i^{th} component is $\overline{\mathcal{E}}(1^k, \overline{pk}, \mathbf{y}[i]\ ;\ \boldsymbol{\omega}[i])$. Let

$$P_a = \Pr\left[\,\mathbf{Exp}^{d\text{-}a}_{\Pi,A}(k) \Rightarrow \textsf{true}\,\right]$$

for $a \in \{0,1\}$. Then

$$\mathbf{Adv}^{\mathrm{css}}_{\Pi,A}(k) = 2P_1 - 1 = 2(P_1 - P_0) + (2P_0 - 1)\ .$$

Adversary B is shown in Figure 5, and we omit the (easy) analysis establishing that

$$2P_0 - 1 \le \mathbf{Adv}^{\mathrm{ind\text{-}cpa}}_{\overline{\Pi},B}(k)\ .$$

Next we define PRG-adversary $D = (\Lambda, D_{\mathrm{p}}, D_{\mathrm{g}})$ with help length $\ell(\cdot)$ as shown in Figure 6 and claim that

$$P_1 - P_0 \le 2\cdot \mathbf{Adv}^{\mathrm{prg}\text{-}v}_{\mathcal{TP},D,n}(k)\ . \tag{12}$$

Before justifying this claim let us see how to conclude. Let J' be the inversion adversary obtained from D by Lemma 5.2. It also has help-length $\ell(\cdot)$. Now apply Proposition 5.1 to get inversion adversary J with help-length 0. Now, putting together the above would give us

$$\mathbf{Adv}^{\mathrm{css}}_{\Pi,A}(k) \le \mathbf{Adv}^{\mathrm{ind\text{-}cpa}}_{\overline{\Pi},B}(k) + 16n(k)v(k)\mathbf{Adv}^{\mathrm{owf}}_{\mathcal{TP},J}(k)\ . \tag{13}$$

However, (10) has an extra factor of 2 on the first right-hand-side term. This is to ensure that the running time of J is as claimed. To see this, consider two cases. The first is when $2(P_1 - P_0) \ge \epsilon(k)/2$. In this case, (12) implies that $\mathbf{Adv}^{\mathrm{prg}\text{-}v}_{\mathcal{TP},D,n}(k) \ge \epsilon(k)/8$, and hence the running time of J' (and hence J) is, up to a constant factor, as given by Lemma 5.2. However, in the second case, namely $2(P_1 - P_0) < \epsilon(k)/2$, the value of $\mathbf{Adv}^{\mathrm{prg}\text{-}v}_{\mathcal{TP},D,n}(k)$ could be very small and the running time of J' (and hence J) would not be as shown in (11). But also in this case we have $2P_0 - 1 \ge \epsilon(k)/2$ so

$$\mathbf{Adv}^{\mathrm{css}}_{\Pi,A}(k) \le 2(2P_0 - 1) \le 2\cdot \mathbf{Adv}^{\mathrm{ind\text{-}cpa}}_{\overline{\Pi},B}(k)$$

so (10) —but not (13)— is true regardless of the advantage of J in this case. Accordingly, we simply halt J' (and hence J) when its running time hits the bound (11).

It remains to justify (12). Let d be the challenge bit of $\mathbf{Exp}^{\mathrm{prg}\text{-}v}_{\mathcal{TP},D,n}(k)$ and d' the output of D_{g}. Then

$$\begin{aligned}
&\Pr\left[\, d' = 1 \ \mid\ d = 1 \,\right] \\
&\quad = \ \Pr\left[\, c = c' \ \mid\ d = 1 \,\right] \\
&\quad = \ \Pr\left[\, c' = 1 \ \mid\ c = 1 \wedge d = 1 \,\right]\frac{1}{2} + (1 - \Pr\left[\, c' = 1 \ \mid\ c = 0 \wedge d = 1 \,\right])\frac{1}{2} \\
&\quad = \ \frac{1}{2} + \frac{1}{2}\Pr\left[\, g = t_1 \ \mid\ c = 1 \wedge d = 1 \,\right] - \frac{1}{2}\Pr\left[\, g = t_1 \ \mid\ c = 0 \wedge d = 1 \,\right] \\
&\quad = \ \frac{1}{2} + \frac{1}{2}\Pr\left[\, \mathbf{Exp}^{d\text{-}1}_{\mathcal{TP},A}(k) \Rightarrow \textsf{true} \ \mid\ b = 1 \,\right] - \frac{1}{2}\Pr\left[\, \mathbf{Exp}^{d\text{-}1}_{\mathcal{TP},A}(k) \Rightarrow \textsf{false} \ \mid\ b = 0 \,\right] \\
&\quad = \ \frac{1}{2} + \frac{1}{2}P_1\ ,
\end{aligned}$$

```
**Experiment** $\mathbf{Exp}^{d\text{-}1}_{\Pi,A}(k)$ / $\boxed{\mathbf{Exp}^{d\text{-}0}_{\Pi,A}(k)}$
$b \leftarrow_{\$} \{0,1\}$
$(\mathbf{x}_0, t_0), (\mathbf{x}_1, t_1) \leftarrow_{\$} A_{\mathrm{m}}(1^k, \lambda)$
$(\phi, \tau) \leftarrow_{\$} G(1^k)$ ; $s \leftarrow_{\$} \{0,1\}^k$
$(\overline{pk}, \overline{sk}) \leftarrow_{\$} \overline{\mathcal{K}}(1^k)$ ; $pk \leftarrow (\phi, \overline{pk}, s)$
$\boldsymbol{\omega} \leftarrow \mathcal{G}_{\mathcal{TP}}(1^k, 1^{n(k)}, \phi, \mathbf{x}_b, s)$
$\boxed{\boldsymbol{\omega} \leftarrow_{\$} B^{v(k)}_{n(k)}}$
$\mathbf{y} \leftarrow F^{n(k)}_{\phi}(\mathbf{x}_b)$ ; $\mathbf{c} \leftarrow \overline{\mathcal{E}}(1^k, \overline{pk}, \mathbf{y} ; \boldsymbol{\omega})$
$g \leftarrow_{\$} A_{\mathrm{g}}(1^k, pk, \mathbf{c}, \lambda)$
If $g = t_1$ then $b' \leftarrow 1$ else $b' \leftarrow 0$
Ret $(b = b')$
```

```
**adversary** $B^{\overline{\mathcal{E}}_{\overline{pk}}(\mathrm{LR}(\cdot,\cdot,b))}(\overline{pk})$:
$(\mathbf{x}_0, t_0), (\mathbf{x}_1, t_1) \leftarrow_{\$} A_{\mathrm{m}}(1^k, \lambda)$
$(\phi, \tau) \leftarrow_{\$} G(1^k)$ ; $s \leftarrow_{\$} \{0,1\}^k$
$pk \leftarrow (\phi, \overline{pk}, s)$
$\mathbf{y}_0 \leftarrow F^{n(k)}_{\phi}(\mathbf{x}_0)$ ; $\mathbf{y}_1 \leftarrow F^{n(k)}_{\phi}(\mathbf{x}_1)$
For $i = 1, \ldots, v(k)$ do
    $\mathbf{c}[i] \leftarrow_{\$} \overline{\mathcal{E}}_{\overline{pk}}(\mathrm{LR}(\mathbf{y}_0[i], \mathbf{y}_1[i], b))$
$g \leftarrow_{\$} A_{\mathrm{g}}(1^k, pk, \mathbf{c}, \lambda)$
If $g = t_1$ then Ret 1 else Ret 0
```

Figure 5: **(Left)** Experiments used in the proof of Theorem 5.3. The experiment d-0 includes the boxed statement while d-1 does not. **(Right)** IND-CPA adversary for proof of Theorem 5.3.

where b is the challenge bit of the Figure 5 experiments. Similarly

$$\begin{aligned}
&\Pr\left[\, d' = 1 \mid d = 0 \,\right] \\
&\quad = \Pr\left[\, c = c' \mid d = 0 \,\right] \\
&\quad = \Pr\left[\, c' = 1 \mid c = 1 \wedge d = 0 \,\right]\frac{1}{2} + (1 - \Pr\left[\, c' = 1 \mid c = 0 \wedge d = 0 \,\right])\frac{1}{2} \\
&\quad = \frac{1}{2} + \frac{1}{2}\Pr\left[\, g = t_1 \mid c = 1 \wedge d = 0 \,\right] - \frac{1}{2}\Pr\left[\, g = t_1 \mid c = 0 \wedge d = 0 \,\right] \\
&\quad = \frac{1}{2} + \frac{1}{2}\Pr\left[\, \mathbf{Exp}^{d\text{-}0}_{\mathcal{TP},A}(k) \Rightarrow \mathsf{true} \mid b = 1 \,\right] - \frac{1}{2}\Pr\left[\, \mathbf{Exp}^{d\text{-}0}_{\mathcal{TP},A}(k) \Rightarrow \mathsf{false} \mid b = 0 \,\right] \\
&\quad = \frac{1}{2} + \frac{1}{2}P_0 \,.
\end{aligned}$$

So

$$\begin{aligned}
\mathbf{Adv}^{\text{prg-}v}_{\mathcal{TP},D,n}(k) &= \Pr\left[\, d' = 1 \mid d = 1 \,\right] - \Pr\left[\, d' = 1 \mid d = 0 \,\right] \\
&= \left(\frac{1}{2} + \frac{1}{2}P_1\right) - \left(\frac{1}{2} + \frac{1}{2}P_0\right) \;=\; \frac{1}{2}(P_1 - P_0)
\end{aligned}$$

establishing (12). ▌

INSTANTIATIONS. $\mathsf{DE1}$ admits quite efficient instantiations. Say we want to encrypt a 1024 bit (random) message. Let the trapdoor one-way permutation be squaring modulo a 1024-bit composite number N [8] that is part of the public key. Then the PRG requires n squarings, where n is the number of bits of randomness required by the (randomized) encryption scheme $\overline{\Pi}$. Let $\overline{\Pi}$ be the Blum-Goldwasser scheme [9], also using a 1024-bit modulus. (This modulus, also part of the public key, must be different from N.) Then encryption cost of $\mathsf{DE1}$ is that of Blum-Goldwasser (1024 squarings) plus $n = 1024$ squarings for the PRG to get coins for $\overline{\Pi}$. (We assume here, and below, an efficient mapping from bits to group elements, otherwise n increases by a small amount.) Decryption time also doubles, coming in at about 4 exponentiations modulo 512 bit numbers (less than one 1024 bit exponentiation!) using Chinese remainders. The ciphertext size is that of

```
algorithm $D_{\rm p}(1^k, \mathbf{x}, \phi, \lambda)$:
Repeat
    $(\mathbf{x}', t') \leftarrow_{\$} A_{\rm m}(1^k, \lambda)$
Until $(\mathbf{x}' = \mathbf{x})$
$t \leftarrow t'$
Ret $t$
```

```
algorithm $D_{\rm g}(1^k, \phi, \lambda, \mathbf{y}, \boldsymbol{\omega}, s, t)$:
$c \leftarrow_{\$} \{0,1\}$ ; $\mathbf{y}_1 \leftarrow \mathbf{y}$ ; $t_1 \leftarrow t$ ; $\boldsymbol{\omega}_1 \leftarrow \boldsymbol{\omega}$
$(\mathbf{x}_0, t_0) \leftarrow_{\$} A_{\rm m}(1^k, \lambda)$
$(\overline{pk}, \overline{sk}) \leftarrow_{\$} \overline{\mathcal{K}}(1^k)$ ; $pk \leftarrow (\phi, \overline{pk}, s)$
$\boldsymbol{\omega}_0 \leftarrow \mathcal{G}_{\mathcal{TP}}(1^k, 1^{n(k)}, \phi, \mathbf{x}_0, s)$ ; $\mathbf{y}_0 \leftarrow F_{\phi}^{n(k)}(\mathbf{x}_0)$
$\mathbf{c} \leftarrow \overline{\mathcal{E}}(1^k, \overline{pk}, \mathbf{y}_c ; \boldsymbol{\omega}_c)$
$g \leftarrow_{\$} A_{\rm g}(1^k, pk, \mathbf{c}, \lambda)$
If $(g = t_c)$ then $c' \leftarrow 1$ else $c' \leftarrow 0$
Ret $c \oplus c' \oplus 1$
```

Figure 6: PRG adversary for proof of Theorem 5.3.

Blum-Goldwasser, namely 2048 bits, and security rests solely on factoring. Alternatively, let $\overline{\Pi}$ be El Gamal hybrid encryption using a 160-bit group. (A universal hash of the DH key is used to one-time symmetrically encrypt the data.) Encryption time for $\mathsf{DE1}$ is that of hybrid El Gamal plus the time for $n = 320$ squarings modulo N, decryption time is 2 exponentiations modulo 512 bit numbers plus one 160-bit exponentiation. and the ciphertext size is only 1344 bits. The security assumption is now factoring + DDH.

DISCUSSION. One might ask why we did not work with IND rather than with CSS notions. The reason is that it is unclear how to meaningfully capture the case of uniformly and independently distributed messages with IND. We could certainly say that an IND-adversary $I = (I_{\rm c}, I_{\rm m}, I_{\rm g})$ is uniform if for every k and every st, b the components of $\mathbf{x}$ are uniformly distributed over $\{0,1\}^k$ when $\mathbf{x} \leftarrow_{\$} I_{\rm m}(1^k, b, st)$. But such an adversary would always have zero advantage.

6 Generalizing Our Construction to Non-Uniform Messages

Section 5 provides a deterministic encryption scheme for the A-CSS-secure encryption of independent, uniformly distributed messages assuming the existence of trapdoor one-way permutations. Here we explain how the same scheme provides A-CSS-secure encryption of independent messages that are not necessarily uniformly distributed but rather have high min-entropy μ, as long as the assumption is strengthened to the existence of trapdoor permutations one-way for distributions of min-entropy μ. We point out that a similar assumption was used by [19] in order to construct signature schemes getting only "imperfect" randomness. The main observation needed for the generalization is simply that min-entropy is preserved under permutation, meaning if a random variable X over B_k has min-entropy μ then so does $f(X)$ for any permutation f on B_k. In the following we make the result more precise and sketch how the previous proof approach generalizes.

EXTENDING THE FRAMEWORK. An inversion adversary $J = (J_{\rm m}, J_{\rm c}, J_{\rm p}, J_{\rm s})$ is now a 4-tuple where $J_{\rm m}$ is a non-uniform algorithm with $[J_{\rm m}(1^k)] \subseteq \{0,1\}^k$ and $J_{\rm c}, J_{\rm p}, J_{\rm s}$ are as before. We say that $J \in \mathcal{J}^{\mu}_{\rm ME}$ if the output of $J_{\rm m}$ has min-entropy μ. The running time of J is defined as the sum of the running times of $J_{\rm m}$, $J_{\rm c}$ and $J_{\rm s}$. A PRG-adversary $D = (D_{\rm m}, D_{\rm c}, D_{\rm p}, D_{\rm g})$ is similarly a 4-tuple where $D_{\rm m}$ is a non-uniform algorithm with $[D_{\rm m}(1^k, 1^v)] \subseteq B_k^v$ and $D_{\rm c}, D_{\rm p}, D_{\rm g}$ are as before. We say that $D \in \mathcal{D}^{\mu}_{\rm ME}$ if the components of the output of $D_{\rm m}$ are independently distributed, each with min-entropy μ. The running time of D is defined as the sum of the running times of $D_{\rm m}$, $D_{\rm c}$ and $D_{\rm g}$. The help length $\ell(\cdot)$ is defined as before and $\mathcal{J}_\ell, \mathcal{D}_\ell$ are the corresponding classes. J is unaided if it has help length 0. Experiment $\mathbf{Exp}^{\rm owf}_{\mathcal{TP},J}(k)$ of Figure 4 is modified by replacing $x \leftarrow_{\$} \{0,1\}^k$ by $x \leftarrow_{\$} J_{\rm m}(1^k)$. Experiment $\mathbf{Exp}^{\text{prg-}v}_{\mathcal{TP},D,n}(k)$ of Figure 4 is modified by replacing $\mathbf{x} \leftarrow\!\!\$\; B_k^{v(k)}$ by $\mathbf{x} \leftarrow\!\!\$\; D_{\mathrm{m}}(1^k, 1^{v(k)})$. The advantage functions are defined as before, and we say that $\mathcal{TP}$ is one-way for min-entropy μ if $\mathbf{Adv}^{\mathrm{owf}}_{\mathcal{TP},J}(\cdot)$ is negligible for all PT $J \in \mathcal{J}_0 \cap \mathcal{J}^{\mu}_{\mathrm{ME}}$. Proposition 5.1 generalizes so that if J is in $\mathcal{J}^{\mu}_{\mathrm{ME}}$ then so is J'. Lemma 5.2 generalizes so that if $D \in \mathcal{D}^{\mu}_{\mathrm{ME}}$ then $J \in \mathcal{J}^{\mu}_{\mathrm{ME}}$.

SECURITY OF OUR SCHEME. Theorem 5.3 generalizes as follows. In the preamble, instead of A being in $\mathcal{A}_{\mathrm{B}} \cap \mathcal{A}_{\lambda} \cap \mathcal{A}_{\mathrm{UN}}$, let it be in $\mathcal{A}_{\mathrm{B}} \cap \mathcal{A}_{\lambda} \cap \mathcal{A}^{\mu}_{\mathrm{ME}} \cap \mathcal{A}_{\times}$. Then, in the conclusion, the unaided inversion adversary J will be in $\mathcal{J}^{\mu}_{\mathrm{ME}}$. The theorem is saying that our scheme is B-CSS secure for independently distributed messages of min-entropy μ assuming $\mathcal{TP}$ is one-way for min-entropy μ and $\overline{\Pi}$ is IND-CPA. Since the transformation of Theorem 4.1 preserves the message distribution, the corollary is that our scheme is A-CSS secure under the same conditions. Since the existence of a family of trapdoor permutations one-way for min-entropy μ implies the existence of one-way trapdoor permutations, it also implies the existence of IND-CPA secure encryption schemes and so we obtain the results under the sole assumption of the existence of trapdoor permutations one-way for min-entropy μ.

7 From Deterministic to Randomized PKE

OVERVIEW. As observed in the introduction, any PRIV-secure deterministic scheme Π is trivially a one-way trapdoor injection, meaning an obvious method for building a secure randomized scheme $\overline{\Pi}$ is to use Π within a generic construction (i.e., [24, 21]) to derive an IND-CPA secure scheme. The equally obvious downside of such an approach is the lack of efficiency. For example, [24] requires large ciphertexts: $\mathcal{O}(k \cdot |M|)$ for security parameter k and message M. ([21] requires both large ciphertexts and large keys, though it meets CCA security.) One would expect to do better given a primitive that provides more than just one-wayness.

A tempting approach to achieve a more efficient construction is the following. Noting that Π meets a form of semantic-security whenever there is sufficient entropy in the message space, we could have $\overline{\Pi}$ encrypt by padding messages with an appropriate number of random bits, and then applying Π to the resulting padded string. This would ensure the scheme always enjoys PRIV security, even when messages have no entropy. But is $\overline{\Pi}$ also IND-CPA? In general the answer is no, due to the fact that Π only provides security when messages are chosen independently of the public key. On the other hand, the IND-CPA definition mandates security even against public-key dependent messages. One can easily build a scheme Π that is PRIV-secure but for which $\overline{\Pi}$ as described is *not* IND-CPA.

Fortunately we can circumvent the key-independency issue using a hybrid-encryption approach. Particularly, encryption first generates a fresh session key and a random pad. Then, it uses Π to encrypt the concatenation of the session key and pad followed by using a standard (one-time secure) encryption scheme to encrypt the actual message under the session key. This approach works even in the context of chosen-ciphertext attacks, see Appendix G.

KEY ENCAPSULATION. We will in fact show how to build a (randomized) key encapsulation mechanism (KEM) [15] from any PRIV-secure deterministic encryption scheme. Using the KEM formulation is simpler and sufficient: in conjunction with any (one-time secure) symmetric scheme, this provides an IND-CPA scheme [15]. Formally, a key-encapsulation mechanism $\Psi = (\mathcal{KK}, \mathcal{KE}, \mathcal{KD})$ is a triple of algorithms. The key generation algorithm $\mathcal{KK}$ takes input security parameter 1^k and outputs a public key, secret key pair. The key encapsulation algorithm $\mathcal{KE}$ takes input 1^k and public key pk and outputs a session key $K \in \{0,1\}^{s(k)}$ and a ciphertext. The function $s\colon \mathbb{N} \to \mathbb{N}$ specifies Ψ's *session-key length.* The key decapsulation algorithm $\mathcal{KD}$ takes as input 1^k, a secret key sk, and a ciphertext and outputs a session key. A KEM-adversary A is a non-uniform algorithm that takes inputs 1^k, a public key, a bit string, and a ciphertext and outputs a bit. We

```
Experiment $\mathbf{Exp}^{\mathrm{kem}}_{\Psi,A}(k)$
$b \leftarrow_{\$} \{0,1\}$ ; $(pk, sk) \leftarrow_{\$} \mathcal{KK}(1^k)$
$(K_1, C) \leftarrow_{\$} \mathcal{KE}(1^k, pk)$; $K_0 \leftarrow_{\$} \{0,1\}^{s(k)}$
$b' \leftarrow_{\$} A(1^k, pk, K_b, C)$
Ret $(b = b')$
```

Figure 7: Experiment defining advantage of a KEM adversary A.

define the KEM advantage of A against Ψ by

$$\mathbf{Adv}^{\mathrm{kem}}_{\Psi,A}(k) = 2 \cdot \Pr\left[\, \mathbf{Exp}^{\mathrm{kem}}_{\Psi,A}(k) \Rightarrow \mathsf{true} \,\right] - 1$$

where the kem experiment is defined in Figure 7.

The construction. Fix functions $\mu, s\colon \mathbb{N} \to \mathbb{N}$. Let $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ be a (deterministic) PKE scheme. We say Π is suitable if it encrypts messages of length $w(\cdot)$ such that $w(k) \geq \mu(k) + s(k)$ for all $k \in \mathbb{N}$. Let $\Psi = (\mathcal{KK}, \mathcal{KE}, \mathcal{KD})$ be the KEM with session key length $s(\cdot)$ defined by the subsequent three algorithms.

```
algorithm $\mathcal{KK}(1^k)$:
$(pk, sk) \leftarrow_{\$} \mathcal{K}(1^k)$
Ret $(pk, sk)$
```

```
algorithm $\mathcal{KE}(1^k, pk)$:
$R \leftarrow_{\$} \{0,1\}^{\mu(k)}$ ; $K \leftarrow_{\$} \{0,1\}^{s(k)}$
$c \leftarrow \mathcal{E}(1^k, pk, R \,\|\, K)$
Ret $(K, c)$
```

```
algorithm $\mathcal{KD}(1^k, sk, c)$:
$R \,\|\, K \leftarrow \mathcal{D}(sk, c)$
Ret $K$
```

The next theorem captures the security of Ψ.

Theorem 7.1 Let $\mu, s\colon \mathbb{N} \to \mathbb{N}$. Let $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ be a suitable PKE scheme. Let $\Psi = (\mathcal{KK}, \mathcal{KE}, \mathcal{KD})$ be the associated KEM scheme as per our construction. Let A be a KEM-adversary. Then there exists an IND-adversary $I = (I_{\mathrm{c}}, I_{\mathrm{m}}, I_{\mathrm{g}}) \in \mathcal{I}^{\mu}_{\mathrm{ME}} \cap \mathcal{I}_{\lambda}$, outputting a single message, such that for all $k \in \mathbb{N}$

$$\mathbf{Adv}^{\mathrm{kem}}_{\Psi,A}(k) \leq \mathbf{Adv}^{\mathrm{ind}}_{\Pi,I}(k) \ .$$

The running time of I is that of A. $\square$

Proof: Below we write μ for $\mu(k)$ and s for $s(k)$. We build $I^* = (I^*_{\mathrm{c}}, I^*_{\mathrm{m}}, I^*_{\mathrm{g}})$ using A, as shown below.

```
algorithm $I^*_{\mathrm{c}}(1^k)$:
$K \leftarrow_{\$} \{0,1\}^s$
Ret $K$
```

```
algorithm $I^*_{\mathrm{m}}(1^k, b, K)$:
$R \leftarrow_{\$} \{0,1\}^{\mu}$
If $b = 1$ then Ret $R \,\|\, K$
$K' \leftarrow_{\$} \{0,1\}^s$
Ret $R \,\|\, K'$
```

```
algorithm $I^*_{\mathrm{g}}(1^k, pk, c, K)$:
$b' \leftarrow_{\$} A(1^k, pk, K, c)$
Ret $b'$
```

I^* has min-entropy μ because of the selection of R. It is straightforward to verify that

$$\Pr\left[\, \mathbf{Exp}^{\mathrm{ind\text{-}cpa}}_{\Pi,I^*}(k) \Rightarrow \mathsf{true} \,\right] = \Pr\left[\, \mathbf{Exp}^{\mathrm{kem}}_{\Psi,A}(k) \Rightarrow \mathsf{true} \,\right] \ .$$

Finally, let I be the IND-adversary with trivial state function that works just like I^* except that K is replaced by a "best" value. ▌

Discussion. We make several observations about the construction. First, Ψ provides *witness-recovering* public-key encryption: all the randomness used to generate a ciphertext is recovered within $\mathcal{KD}$. Second, we only require Π to be secure against adversaries that output a single message. This is notable because, as discussed in Section 3, security against single-message adversaries is strictly weaker than multi-message adversaries. Finally, one might wonder if it is possible to dispense with the random padding R. In fact it is requisite to meet KEM security. Let Ψ' work just like our construction Ψ except that we omit R. But then there exists an easy KEM-adversary against Ψ': just compute $c' \leftarrow \mathcal{E}(1^k, pk, K)$ and output 1 iff $c' = c$. If $\mathcal{E}$ is deterministic the advantage of this adversary is $1 - 2^{-s(k)}$.

References

[1] M. Bellare. The Goldreich-Levin Theorem. Manuscript. `http://www-cse.ucsd.edu/users/mihir/papers/gl.pdf`

[2] M. Bellare, A. Boldyreva, and S. Micali. Public-key encryption in a multi-user setting: Security proofs and improvements. In *Advances in Cryptology – EUROCRYPT '00*, LNCS vol. 1807, pp. 259–274, 2000.

[3] M. Bellare, A. Boldyreva, and A. O'Neill. Deterministic and efficiently searchable encryption. In *Advances in Cryptology– CRYPTO '07*, LNCS vol. 4622, pp. 535–552, 2007.

[4] M. Bellare and P. Rogaway. Random oracles are practical: A paradigm for designing efficient protocols. In *Conference on Computer and Communications Security – CCS '93*, ACM, pp. 62–73, 1993.

[5] M. Bellare and P. Rogaway. Robust computational secret sharing and a unified account of classical secret-sharing goals. In *Conference on Computer and Communications Security – CCS '07*, ACM, pp. 172–184, 2007.

[6] M. Bellare and P. Rogaway. The security of triple encryption and a framework for code-based game-playing proofs. In *Advances in Cryptology – EUROCRYPT '06*, LNCS vol. 4004, pp. 409–426, 2006.

[7] M. Bellare and A. Sahai. Non-malleable encryption: Equivalence between two notions, and an indistinguishability-based characterization. In *Advances in Cryptology – CRYPTO '99*, LNCS vol. 1666, pp. 519–536, 1999.

[8] L. Blum, M. Blum, and M. Shub. A simple unpredictable pseudo-random number generator. *SIAM Journal on Computing*, vol. 15, pp. 364–383, 1986.

[9] M. Blum and S. Goldwasser. An efficient probabilistic public-key encryption scheme which hides all partial information. In *Advances in Cryptology – CRYPTO '84*, LNCS vol. 196, pp. 289–302, 1984.

[10] M. Blum and S. Micali. How to generate cryptographically strong sequences of pseudorandom bits. *SIAM Journal on Computing*, vol. 13, pp. 850–864, 1984.

[11] A. Boldyreva, S. Fehr, and A. O'Neill. On notions of security for deterministic encryption, and efficient constructions without random oracles. In *Advances in Cryptology – CRYPTO '08*, 2008, to appear.

[12] D. Boneh, G. Di Crescenzo, R. Ostrovsky, and G. Persiano. Public key encryption with keyword search. In *Advances in Cryptology – EUROCRYPT '04*, LNCS vol. 3027, pp. 506–522, 2004.

[13] R. Canetti. Towards realizing random oracles: Hash functions that hide all partial information. In *Advances in Cryptology – CRYPTO '97*, LNCS vol. 1294, pp. 455–469, 1997.

[14] R. Canetti, D. Micciancio, and O. Reingold. Perfectly one-way probabilistic hash functions (Preliminary version). In *Symposium on the Theory of Computation – STOC '98*, pp. 131–141, 1998.

[15] R. Cramer and V. Shoup. A practical public key cryptosystem provably secure against adaptive chosen ciphertext attack. In *Advances in Cryptology – CRYPTO '98*, LNCS vol. 1462, pp. 13–25, 1998.

[16] I. Damgaard, D. Hofheinz, E. Kiltz, and R. Thorbek. Public-key encryption with non-interactive opening. In *Topics in Cryptology – CT-RSA '08*, LNCS vol. 4964, pp. 239–255, 2008.

[17] S. Desrosiers. Entropic security in quantum cryptography. arXiv e-Print quant-ph/0703046, `http://arxiv.org/abs/quant-ph/0703046` , 2007.

[18] S. Desrosiers and F. Dupuis. Quantum entropic security and approximate quantum encryption. arXiv e-Print quant-ph/0707.0691, `http://arxiv.org/abs/0707.0691` , 2007.

[19] Y. Dodis, S. J. Ong, M. Prabhakaran, and A. Sahai. On the (im)possibility of cryptography with imperfect randomness. In *Symposium on the Foundations of Computer Science – FOCS '04*, IEEE, pp. 196–205, 2004.

[20] Y. Dodis and A. Smith. Entropic security and the encryption of high entropy messages. In *Theory of Cryptography Conference – TCC '05*, LNCS vol. 3378, pp. 556–577, 2005.

[21] D. Dolev, C. Dwork, and M. Naor. Nonmalleable cryptography. *SIAM Journal on Computing*, vol. 30, no. 2, pp. 391–437, 2000.

[22] T. El Gamal. A public-key cryptosystem and a signature scheme based on discrete logarithms. In *Advances in Cryptology – CRYPTO '84*, LNCS vol. 196, pp. 10–18, 1984.

[23] O. Goldreich. A uniform complexity treatment of encryption and zero-knowledge. *Journal of Cryptology*, vol. 6, pp. 21–53, 1993.

[24] O. Goldreich and L. Levin. A hard-core predicate for all one-way functions. In *Symposium on the Theory of Computation – STOC '89*, ACM, pp. 25–32, 1989.

[25] S. Goldwasser and S. Micali. Probabilistic encryption. *Journal of Computer and Systems Sciences*, vol. 28, no. 2, 1984, pp. 412–426.

[26] S. Micali, C. Rackoff, and R. Sloan. The notion of security for probabilistic cryptosystems. *SIAM Journal on Computing*, vol. 17, no. 2, April 1988, pp. 412–426.

[27] C. Peikert and B. Waters. Lossy trapdoor functions and their applications. In *Symposium on the Theory of Computing – STOC '08*, ACM, pp. 187–196, 2008.

[28] A. Russell and H. Wang. How to fool an unbounded adversary with a short key. In *Advances in Cryptology – EUROCRYPT'02*, LNCS vol. 2332, pp. 133–148, 2002.

[29] A. Smith. Personal correspondence. 2008.

[30] D. Song, D. Wagner, and A. Perrig. Practical techniques for searches on encrypted data. In *Symposium on Security and Privacy*, IEEE, pp. 44-55, 2000.

[31] A. Yao. Theory and applications of trapdoor functions. In *Symposium on Foundations of Computer Science – FOCS '82*, IEEE, pp. 80–91, 1982.

A Message-based Partial Information

In our css and sss experiments, the information t computed by A_{m} can depend on coins underlying the generation of $\mathbf{x}$ rather than merely on $\mathbf{x}$. Here we show that the two formulations are in fact equivalent and then explore the implications for single versus multi-message security that motivated this question.

EQUIVALENCE. An SS-adversary $A = (A_{\mathrm{c}}, A_{\mathrm{m}}, A_{\mathrm{g}})$ is said to be separable if there are non-uniform algorithms $A_{\mathrm{d}}, A_{\mathrm{p}}$ called the message space and partial information function, respectively, such that the outputs of the following are identically distributed for all $k \in \mathbb{N}$ and all st:

```
$(\mathbf{x}, t) \leftarrow\$ A_{\mathrm{m}}(1^k, st)$
Ret $(\mathbf{x}, t)$
```

```
$\mathbf{x} \leftarrow\$ A_{\mathrm{d}}(1^k, st)$
$t \leftarrow\$ A_{\mathrm{p}}(1^k, \mathbf{x}, st)$
Ret $(\mathbf{x}, t)$
```

Let $\mathcal{A}_{\mathrm{sep}}$ be the class of separable SS-adversaries. The following says that restricting attention to separable adversaries leaves the class of secure schemes unchanged.

Theorem A.1 Let $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ be a PKE scheme. Let $A = (A_{\mathrm{c}}, A_{\mathrm{m}}, A_{\mathrm{g}}) \in \mathcal{A}_{\mathrm{ME}}^{\mu}$ be an SS-adversary with information length $\ell(\cdot)$. Then there is a separable SS-adversary $A' = (A_{\mathrm{c}}, A'_{\mathrm{m}}, A'_{\mathrm{g}}) \in \mathcal{A}_{\mathrm{ME}}^{\mu} \cap \mathcal{A}_{\mathrm{sep}}$ with information length $\ell(\cdot)$ such that for all $k \in \mathbb{N}$

$$\mathbf{Adv}_{\Pi,A}^{\mathrm{css}}(k) \leq \mathbf{Adv}_{\Pi,A'}^{\mathrm{css}}(k) \ .$$

The running time of A' is that of A plus $\mathcal{O}(\ell + \mu)$. If A is in $\mathcal{A}_{\mathrm{BB}}^{\delta}$ then so is A'. □

Proof: Let $m(k) = \lceil \mu(k) \rceil$ and let $v(\cdot)$ be the number of messages output by A. We obtain $A' = (A_{\mathrm{c}}, A'_{\mathrm{m}}, A'_{\mathrm{g}})$, which will output $v(\cdot) + 1$ messages, by defining

```
algorithm $A'_{\mathrm{m}}(1^k, st)$:
$(\mathbf{x}, t) \leftarrow\$ A_{\mathrm{m}}(1^k, st)$
$r \leftarrow\$ \{0,1\}^{m(k)}$
$\mathbf{x}[v(k)+1] \leftarrow t \,\|\, r$
Ret $(\mathbf{x}, t)$
```

```
algorithm $A'_{\mathrm{g}}(1^k, pk, \mathbf{c}', st)$:
$\mathbf{c} \leftarrow (\mathbf{c}'[1], \ldots, \mathbf{c}'[v(k)])$
$g \leftarrow\$ A_{\mathrm{g}}(1^k, pk, \mathbf{c}, st)$
Ret $g$
```

That is, A'_{m} simply puts t into the message vector, randomizing it to ensure the min-entropy of the adversary is not reduced. it is easy to see that $A' = (A_{\mathrm{c}}, A'_{\mathrm{m}}, A'_{\mathrm{g}})$ is separable and has the same advantage as A. ▮

WHY SEPARABILITY? The following says that in the context of separable adversaries producing independently distributed messages, security of single and multi message encryption are equivalent. The proof is a simple hybrid argument.

Proposition A.2 Let $\Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D})$ be a PKE scheme. Let $A = (A_{\mathrm{c}}, A_{\mathrm{m}}, A_{\mathrm{g}}) \in \mathcal{A}_{\mathrm{ME}}^{\mu} \cap \mathcal{A}_{\mathrm{sep}} \cap \mathcal{A}_{\times}$ be an SS-adversary with information length $\ell(\cdot)$ outputting $v(\cdot)$ messages. Then there is an SS-adversary $A' = (A'_{\mathrm{c}}, A'_{\mathrm{m}}, A'_{\mathrm{g}}) \in \mathcal{A}_{\mathrm{ME}}^{\mu} \cap \mathcal{A}_{\mathrm{sep}}$ with information length $\ell(\cdot)$ outputting $v'(\cdot) = 1$ message such that for all $k \in \mathbb{N}$

$$\mathbf{Adv}_{\Pi,A}^{\mathrm{css}}(k) \leq v(k) \cdot \mathbf{Adv}_{\Pi,A'}^{\mathrm{css}}(k) \ .$$

The running time of A' is that of A plus $\mathcal{O}(v)$. If A is in $\mathcal{A}_{\mathrm{BB}}^{\delta}$ then so is A'. □

This leads to the following possible way to simplify the proof of Theorem 5.3. First, by Theorem A.1, restrict attention to separable adversaries. Second, by Proposition A.2, assume $v(\cdot) = 1$. The catch is that

Theorem A.1 does not preserve message independence, meaning even if $A \in \mathcal{A}_\times$, adversary A' need not be in $\mathcal{A}_\times$. This is why Theorem 5.3 explicitly considers arbitrary $v(\cdot)$.

OPEN QUESTIONS. The above leads to several interesting open questions. The first is whether there is a reduction to separated adversaries that preserves independence, meaning an analog of Theorem A.1 in which $A \in \mathcal{A}_\times$ implies $A' \in \mathcal{A}_\times$. Barring this another open question is whether Proposition A.2 extends to non-separable adversaries. In case that the answer to either question is "no" it would also be interesting to see counter-examples.

B BB-CSS $\Rightarrow$ B-CSS: Proof of Theorem 4.2

Let $n : \mathbb{N} \to \mathbb{N}$ be a function to be specified later. Below we write n for $n(k)$. Let $H : \{0,1\}^s \times \{0,1\} \to \{0,1\}$ be a family of pairwise independent hash functions where each key $K \in \{0,1\}^s$ specifies a particular function $H_K : \{0,1\} \to \{0,1\}$. (Specifically let $s = 2$ so that a key $K = a \parallel b$ is a pair of bits and let $H_K(x) = ax \oplus b$.) Let $S^n = \{0,1\}^s \times \cdots \times \{0,1\}^s$ where $\{0,1\}^s$ is repeated n times. Since $A \in \mathcal{A}_\lambda$ its state function $A_{\rm c}$ always outputs λ, which is the last input to both $A_{\rm m}$ and $A_{\rm g}$. Let $A^* = (A^*_{\rm c}, A^*_{\rm m}, A^*_{\rm g})$ where

```
algorithm $A^*_{\rm c}(1^k)$:
$\mathbf{K} \leftarrow\$ S^n$
Ret $\mathbf{K}$
```

```
algorithm $A^*_{\rm m}(1^k, \mathbf{K})$:
$(\mathbf{x}, t) \leftarrow\$ A_{\rm m}(1^k, \lambda)$
$i \leftarrow\$ [1\,..\,n]$
Ret $(\mathbf{x}, H(\mathbf{K}[i], t))$
```

```
algorithm $A^*_{\rm g}(1^k, pk, \mathbf{c}, \mathbf{K})$:
$g \leftarrow\$ A_{\rm g}(1^k, pk, \mathbf{c}, \lambda)$
$j \leftarrow\$ [1\,..\,n]$
Ret $H(\mathbf{K}[j], g)$
```

For $t \in \{0,1\}$ let $Z_t(\mathbf{K}) = \Pr\left[H(\mathbf{K}[i], t) = 0 \;:\; i \leftarrow\$ [1\,..\,n]\right]$ and let

$$G_1 = \left\{ \mathbf{K} \in S^n \;:\; \left| Z_t(\mathbf{K}) - \frac{1}{2} \right| \geq \frac{1}{4} \;\text{ for some } t \in \{0,1\} \right\} .$$

Claim B.1 $\Pr\left[\, \mathbf{K} \in G_1 \;:\; \mathbf{K} \leftarrow\$ S^n \,\right] \leq 4e^{-n/32}$ □

The proof of the above will use the following standard Chernoff bound.

Lemma B.2 Let $X_1, \ldots, X_n$ be independent random variables taking values in $[0,1]$ and let $X = X_1 + \cdots + X_n$. Then for any $a \geq 0$

$$\Pr\left[\, |X - \mathbf{E}[X]| \geq a \,\right] \leq 2e^{-a^2/2n} \quad \square$$

Proof of Claim B.1: Let $X_{t,i}(\mathbf{K}) = 1 - H(\mathbf{K}[i], t)$. Let $X_t = \sum_{i=1}^n X_{t,i}$. Then $\mathbf{E}[X_{t,i}] = 1/2$ and $\mathbf{E}[X_t] = n/2$. Observe that $Z_t(\mathbf{K}) = X_t(\mathbf{K})/n$. So, with probabilities taken over $\mathbf{K} \leftarrow\$ S^n$,

$$\Pr\left[\left| Z_t - \frac{1}{2} \right| \geq \frac{1}{4} \right] = \Pr\left[\left| X_t - \frac{n}{2} \right| \geq \frac{n}{4} \right] = \Pr\left[\left| X_t - \mathbf{E}[X_t] \right| \geq \frac{n}{4} \right] .$$

But $\{X_{t,i}\}_{i=1}^n$ are independent so we can use Lemma B.2

$$\Pr\left[\left| X_t - \mathbf{E}[X_t] \right| \geq \frac{n}{4} \right] \;\leq\; 2e^{-\left(\frac{n}{4}\right)^2/2n} \;=\; 2e^{-n/32} .$$

Finally, we can apply a union bound to get

$$\Pr\left[\, \mathbf{K} \in G_1 \;:\; \mathbf{K} \leftarrow\$ S^n \,\right] \leq \sum_{t \in \{0,1\}} \Pr\left[\left| X_t - \frac{n}{2} \right| \geq \frac{n}{4} \;:\; \mathbf{K} \leftarrow\$ S^n \right] \leq 4e^{-n/32} . \blacksquare$$

Claim B.3 $\mathbf{Adv}^{\rm css}_{\Pi,A^*}(k) = \dfrac{1}{2n} \mathbf{Adv}^{\rm css}_{\Pi,A}(k)$ □

Proof: Let $P_1 = \Pr\left[\,\mathbf{Exp}^{\text{css-1}}_{\Pi,A}(k) \Rightarrow \mathsf{true}\,\right]$ and $P_0 = \Pr\left[\,\mathbf{Exp}^{\text{css-0}}_{\Pi,A}(k) \Rightarrow \mathsf{false}\,\right]$. Then

$$\begin{aligned}\Pr\left[\,\mathbf{Exp}^{\text{css-1}}_{\Pi,A^*}(k) \Rightarrow \mathsf{true}\,\right] &= \frac{P_1}{n}\cdot 1 + \left(1-\frac{P_1}{n}\right)\frac{1}{2} \quad \text{and}\\ \Pr\left[\,\mathbf{Exp}^{\text{css-0}}_{\Pi,A^*}(k) \Rightarrow \mathsf{false}\,\right] &= \frac{P_0}{n}\cdot 1 + \left(1-\frac{P_0}{n}\right)\frac{1}{2}\end{aligned}$$

where we have used that H is pairwise independent and so

$$\begin{aligned}\mathbf{Adv}^{\text{css}}_{\Pi,A^*}(k) &= \frac{P_1}{n} - \frac{P_0}{n} + \frac{1}{2}\left[\left(1-\frac{P_1}{n}\right) - \left(1-\frac{P_0}{n}\right)\right]\\ &= \frac{1}{2}\frac{P_1}{n} - \frac{1}{2}\frac{P_0}{n} = \frac{1}{2n}\cdot\mathbf{Adv}^{\text{css}}_{\Pi,A}(k)\ . \ \blacksquare\end{aligned}$$

Let $Y(\mathbf{K})$ be the css advantage of A^* when we do not choose st at random in the game but instead use $\mathbf{K}$. Then $\mathbf{Adv}^{\text{css}}_{\Pi,A^*}(k) = \mathbf{E}[Y]$, where the expectation is over $\mathbf{K} \leftarrow\$\, S^n$. Let $P = 2^{-sn}$ be the probability of picking a particular $\mathbf{K}$. Then we use the definition of expectation and Claim B.1 to get that

$$\begin{aligned}\mathbf{E}[Y] &= \sum_{\mathbf{K}\notin G_1} Y(\mathbf{K})\cdot P + \sum_{\mathbf{K}\in G_1} Y(\mathbf{K})\cdot P\\ &\le \sum_{\mathbf{K}\notin G_1} Y(\mathbf{K})\cdot P + \Pr\left[\,\mathbf{K}\in G_1 \ :\ \mathbf{K}\leftarrow\$\, S^n\,\right] \le \sum_{\mathbf{K}\notin G_1} Y(\mathbf{K})\cdot P + 4e^{-n/32}\ .\end{aligned}$$

Rearranging, applying Claim B.3, and recalling that $\epsilon = \mathbf{Adv}^{\text{css}}_{\Pi,A}(k)$ gives

$$\sum_{\mathbf{K}\notin G_1} Y(\mathbf{K})\cdot P \ge \mathbf{E}[Y] - 4e^{-n/32} = \mathbf{Adv}^{\text{css}}_{\Pi,A^*}(k) - \frac{4}{e^{n/32}} = \frac{\epsilon}{2n} - \frac{4}{e^{n/32}}\ . \tag{14}$$

Then choosing n so that $4e^{-n/32} \le \epsilon/4n$ ensures that the difference in (14) is greater than or equal to $\epsilon/4n$. This ensures that there exists a $\mathbf{K}$ such that $Y(\mathbf{K}) \ge \epsilon/4n$ and also $\mathbf{K} \notin G_1$. Let A' be the adversary that runs like A^* except that it always uses such a $\mathbf{K}$. Then A' has trivial state function, is 1/4-balanced, and has advantage at least $\epsilon/4n$.

Now we determine a suitable value for n. We need that $4ne^{-n/32} \le \epsilon/4$. We can first find an N so that $4n \le e^{n/64}$ for all $n \ge N$. This holds for $N = 485$. We can then find an $n \ge 485$ such that $e^{-n/64} \le \epsilon/4$. This concludes the proof.

C A-CSS ⇒ A-SSS: Proof of Theorem 4.3

We define the simulator S below.

```
Algorithm $S(1^k, pk, \lambda)$:
$(\mathbf{x}_0, t_0) \leftarrow\$\, A_{\mathrm{m}}(1^k, st)$
$\mathbf{c} \leftarrow\$\, \mathcal{E}(1^k, pk, \mathbf{x}_0)$
$g \leftarrow\$\, A_{\mathrm{g}}(1^k, pk, \mathbf{c}, st)$
Ret $g$
```

Then we have that

$$\Pr\left[\,\mathbf{Exp}^{\text{sss-1}}_{\Pi,A,S}(k) \Rightarrow \mathsf{true}\,\right] = \Pr\left[\,\mathbf{Exp}^{\text{css-1}}_{\Pi,A}(k) \Rightarrow \mathsf{true}\,\right]$$

because the experiments are exactly the same in the case that $b = 1$. By the construction of S we also have that

$$\Pr\left[\,\mathbf{Exp}^{\text{sss-0}}_{\Pi,A,S}(k) \Rightarrow \mathsf{false}\,\right] = \Pr\left[\,\mathbf{Exp}^{\text{css-0}}_{\Pi,A}(k) \Rightarrow \mathsf{false}\,\right]$$

by the same reasoning. The theorem statement follows.

D BB-SSS ⇒ IND: Proof of Theorem 4.4

We define $A = (\Lambda, A_{\mathrm{m}}, A_{\mathrm{g}})$ via

```
algorithm $A_{\mathrm{m}}(1^k, \lambda)$:
    $t \leftarrow\$ \{0,1\}$
    $\mathbf{x} \leftarrow\$ I_{\mathrm{m}}(1^k, t, \lambda)$
    Ret $(\mathbf{x}, t)$

algorithm $A_{\mathrm{g}}(1^k, pk, \mathbf{c}, \lambda)$:
    $t' \leftarrow\$ I_{\mathrm{g}}(1^k, pk, \mathbf{c}, \lambda)$
    Ret $t'$
```

Note that A is perfectly balanced since it chooses d uniformly. Let S be an arbitrary simulator. Let $A_{\mathrm{g}} \Rightarrow t$ be the event that A_{g} outputs t. in $\mathbf{Exp}^{\mathrm{sss}}_{\Pi,A,S}(k)$. Let $S \not\Rightarrow t$ be the event that S outputs $1-t$ in $\mathbf{Exp}^{\mathrm{sss}}_{\Pi,A,S}(k)$. Then,

$$\begin{aligned}
\mathbf{Adv}^{\mathrm{sss}}_{\Pi,A,S}(k) &= 2 \cdot \Pr\left[\,\mathbf{Exp}^{\mathrm{sss}}_{\Pi,A,S}(k) \Rightarrow \mathsf{true}\,\right] - 1 \\
&= 2 \cdot (\Pr[\,A_{\mathrm{g}} \Rightarrow t \mid b = 1\,] \cdot \Pr[\,b=1\,] + \Pr[\,S \not\Rightarrow t \mid b = 0\,] \cdot \Pr[\,b = 0\,]) - 1 \\
&= 2 \cdot \left(\frac{1}{2}\Pr\left[\,\mathbf{Exp}^{\mathrm{ind}}_{\Pi,I}(k) \Rightarrow \mathsf{true}\,\right] + \frac{1}{2}\cdot\frac{1}{2}\right) - 1 \qquad (15) \\
&= \Pr\left[\,\mathbf{Exp}^{\mathrm{ind}}_{\Pi,I}(k) \Rightarrow \mathsf{true}\,\right] - \frac{1}{2} \\
&= \frac{1}{2} + \frac{1}{2}\cdot \mathbf{Adv}^{\mathrm{ind}}_{\Pi,I}(k) - \frac{1}{2} \qquad (16) \\
&= \frac{1}{2}\cdot \mathbf{Adv}^{\mathrm{ind}}_{\Pi,I}(k)\ .
\end{aligned}$$

In the case that $b = 1$, the experiment $\mathbf{Exp}^{\mathrm{sss}}_{\Pi,A,S}(k)$ simulates for I exactly the experiment $\mathbf{Exp}^{\mathrm{ind}}_{\Pi,I}(k)$. In the case that $b = 0$, the simulator S receives no information about the bit t. Thus, the probability that it outputs a bit not equal to t is $1/2$. Together these facts justify (15). Equation (16) is derived by applying (7).

E IND ⇒ BB-CSS: Proof of Theorem 4.5

First, to give an idea of the efficiency of the reduction relative to δ, note that for $\delta = 1/4$, the running time of I is increased over that of A by the time to perform $4k + 13$ executions of A_{m}.

Let $n(\cdot)\colon \mathbb{N} \to \mathbb{N}$ to be defined later; below we write n for $n(\cdot)$. We define two IND-adversaries $I = (\Lambda, I_{\mathrm{m}}, I_{\mathrm{g}})$ and $I' = (\Lambda, I'_{\mathrm{m}}, I_{\mathrm{g}})$, both with trivial state functions and with the other algorithms defined below.

```
Algorithm $I_{\mathrm{m}}(1^k, b, \lambda)$:
For $i = 1, \ldots, n$ do
    $(\mathbf{x}, t) \leftarrow\$ A_{\mathrm{m}}(1^k, \lambda)$
    If $t = b$ then Ret $\mathbf{x}$
Ret $\mathbf{x}$

Algorithm $I'_{\mathrm{m}}(1^k, b, \lambda)$:
Do $(\mathbf{x}, t) \leftarrow\$ A_{\mathrm{m}}(1^k, \lambda)$
Until $(t = b)$
Ret $\mathbf{x}$

Algorithm $I_{\mathrm{g}}(1^k, pk, \mathbf{c}, \lambda)$:
$g \leftarrow\$ A_{\mathrm{g}}(1^k, pk, \mathbf{c}, \lambda)$
Ret $g$
```

Note that I'_{m} may not be PT. I_{m} is an approximation to it that is PT. We first state three claims and use them to conclude, then proceed to prove the claims.

Claim E.1 $\mathbf{Adv}^{\mathrm{ind}}_{\Pi,I'}(k) \;\le\; \mathbf{Adv}^{\mathrm{ind}}_{\Pi,I}(k) + 4\cdot\left(\frac{1}{2}+\delta\right)^{n-1}$

Claim E.2 $\mathbf{Adv}^{\mathrm{css}}_{\Pi,A}(k) \;\le\; 2\cdot\mathbf{Adv}^{\mathrm{ind}}_{\Pi,I'}(k)$

Combining Claim E.1 and Claim E.2 gives that

$$\mathbf{Adv}^{\mathrm{css}}_{\Pi,A}(k) \;=\; 2\cdot\mathbf{Adv}^{\mathrm{ind}}_{\Pi,I}(k) + 8\left(\frac{1}{2}+\delta\right)^{n-1} \;\le\; 2\cdot\mathbf{Adv}^{\mathrm{ind}}_{\Pi,I}(k) + 2^{-k}\ ,$$

the last part by setting

$$n(k) = \left\lceil -\frac{1}{\log\left(\frac{1}{2}+\delta\right)}\right\rceil\cdot(k+3)+1\ .$$

Our final claim is that the min-entropy of I is close to that of A.

Claim E.3 I has min-entropy $\mu'(k)=\mu(k)-2+\log(1-2\delta)$.

Before justifying the claims, we fix some notation. Let $\mathbf{x}\in[A_{\mathrm{m}}(1^k,\lambda)]$ and let $b\in\{0,1\}$. Let

$$\begin{aligned}
p_1 &= \Pr\left[\,t=1\ :\ (\mathbf{y},t)\leftarrow_{\$} A_{\mathrm{m}}(1^k,\lambda)\,\right]\\
p_0 &= \Pr\left[\,t=0\ :\ (\mathbf{y},t)\leftarrow_{\$} A_{\mathrm{m}}(1^k,\lambda)\,\right]=1-p_1\\
\alpha_b(\mathbf{x}) &= \Pr\left[\,(\mathbf{y},t)=(\mathbf{x},b)\ :\ (\mathbf{y},t)\leftarrow_{\$} A_{\mathrm{m}}(1^k,\lambda)\,\right]\\
\gamma_b(\mathbf{x}) &= \Pr\left[\,\mathbf{y}=\mathbf{x}\ :\ \mathbf{y}\leftarrow_{\$} I_{\mathrm{m}}(1^k,b,\lambda)\,\right]\\
\gamma'_b(\mathbf{x}) &= \Pr\left[\,\mathbf{y}=\mathbf{x}\ :\ \mathbf{y}\leftarrow_{\$} I'_{\mathrm{m}}(1^k,b,\lambda)\,\right]\ .
\end{aligned}$$

Then, we have

$$\begin{aligned}
\gamma_b(\mathbf{x}) &= p_{1-b}^{n-1}\cdot\alpha_{1-b}(\mathbf{x})+\sum_{i=0}^{n-1}p_{1-b}^{i}\cdot\alpha_b(\mathbf{x}) \;=\; p_{1-b}^{n-1}\cdot\alpha_{1-b}(\mathbf{x})+\alpha_b(\mathbf{x})\cdot\sum_{i=0}^{n-1}p_{1-b}^{i}\\
&= p_{1-b}^{n-1}\cdot\alpha_{1-b}(\mathbf{x})+\alpha_b(\mathbf{x})\cdot\frac{1-p_{1-b}^{n}}{p_b}\ .
\end{aligned}$$

and

$$\gamma'_b(\mathbf{x}) \;=\; \sum_{i=0}^{\infty}p_{1-b}^{i}\cdot\alpha_b(\mathbf{x}) \;=\; \frac{\alpha_b(\mathbf{x})}{p_b}\ .$$

So,

$$\begin{aligned}
\left|\gamma'_b(\mathbf{x})-\gamma_b(\mathbf{x})\right| &= \left|\frac{\alpha_b(\mathbf{x})}{p_b}-p_{1-b}^{n-1}\cdot\alpha_{1-b}(\mathbf{x})-\alpha_b(\mathbf{x})\cdot\frac{1-p_{1-b}^{n}}{p_b}\right|\\
&= \left|\frac{\alpha_b(\mathbf{x})p_{1-b}^{n}}{p_b}-p_{1-b}^{n-1}\cdot\alpha_{1-b}(\mathbf{x})\right|\ .
\end{aligned}$$

We now turn to proving the claims.

Proof of Claim E.1: For any $b \in \{0,1\}$,

$$\left|\Pr\left[\mathbf{Exp}^{\text{ind-b}}_{\Pi,I'}(k) \Rightarrow \mathsf{true}\right] - \Pr\left[\mathbf{Exp}^{\text{ind-b}}_{\Pi,I}(k) \Rightarrow \mathsf{true}\right]\right| \leq \text{SD}\left(I'_{\text{m}}(1^k,b), I_{\text{m}}(1^k,b)\right)$$

and

$$\begin{aligned}
\text{SD}\left(I'_{\text{m}}(1^k,b), I_{\text{m}}(1^k,b)\right) &= \sum_{\mathbf{x}} \left|\gamma'_b(\mathbf{x}) - \gamma_b(\mathbf{x})\right| \\
&= \sum_{\mathbf{x}} \left|\frac{\alpha_b(\mathbf{x})p^n_{1-b}}{p_b} - p^{n-1}_{1-b} \cdot \alpha_{1-b}(\mathbf{x})\right| \\
&\leq \sum_{\mathbf{x}} \frac{\alpha_b(\mathbf{x})p^n_{1-b}}{p_b} + \sum_{\mathbf{x}} p^{n-1}_{1-b} \cdot \alpha_{1-b}(\mathbf{x}) \\
&= p^n_{1-b}\underbrace{\Big(\sum_{\mathbf{x}} \frac{\alpha_b(\mathbf{x})}{p_b}\Big)}_{=1} + p^{n-1}_{1-b}\underbrace{\Big(\sum_{\mathbf{x}} \alpha_{1-b}(\mathbf{x})\Big)}_{\leq 1} \\
&\leq p^{n-1}_{1-b}(p_{1-b}+1) \\
&\leq 2p^{n-1}_{1-b} \\
&\leq 2 \cdot \left(\frac{1}{2}+\delta\right)^{n-1}
\end{aligned}$$

where we have used the fact that $p_b \leq 1/2 + \delta$, which follows from the fact that A is δ-balanced. The above implies that

$$\mathbf{Adv}^{\text{ind}}_{\Pi,I'}(k) - \mathbf{Adv}^{\text{ind}}_{\Pi,I}(k) \leq 4 \cdot \left(\frac{1}{2}+\delta\right)^{n-1}$$

which proves the claim. ▌

To prove Claim E.2 we will utilize the following lemma:

Lemma E.4 Let $\Pi = (\mathcal{K},\mathcal{E},\mathcal{D})$ be an encryption scheme and $A = (A_{\text{c}}, A_{\text{m}}, A_{\text{g}}) \in \mathcal{A}_{\text{SS}} \cap \mathcal{A}_{\lambda} \cap \mathcal{A}_{\text{B}}$ be a boolean ss-adversary. Consider the following experiment, where $k \in \mathbb{N}$:

$$(\mathbf{x},t) \leftarrow\$ \, A_{\text{m}}(1^k,\lambda) \; ; \; (pk,sk) \leftarrow\$ \, \mathcal{K}(1^k) \; ; \; \mathbf{c} \leftarrow\$ \, \mathcal{E}(pk,\mathbf{x}) \; ; \; g \leftarrow\$ \, A_{\text{g}}(1^k,pk,\mathbf{c},\lambda) \; .$$

Let $a_1 = \Pr\left[\, g=1 \mid t=1 \,\right]$ and $b_1 = \Pr\left[\, g=1 \mid t=0 \,\right]$ and $c_1 = \Pr\left[\, t=1 \,\right]$. Then

$$\mathbf{Adv}^{\text{css}}_{\Pi,A}(k) = (a_1 - b_1)(2c_1 - 2c_1^2) \quad \square$$

Proof of Lemma E.4: We extend the experiment of the lemma statement with the additional step

$$(\mathbf{x}',t') \leftarrow\$ \, A_{\text{m}}(1^k,\lambda) \; .$$

Then let

$$\begin{aligned}
a_0 &= \Pr\left[\, g=0 \mid t=1 \,\right] = 1 - a_1 \\
b_0 &= \Pr\left[\, g=0 \mid t=0 \,\right] = 1 - b_1 \\
c_0 &= \Pr\left[\, t=0 \,\right] = 1 - c_1
\end{aligned}$$

Then

$$
\begin{aligned}
\Pr\left[\,\mathbf{Exp}^{\text{css-1}}_{\Pi,A}(k) \Rightarrow \mathsf{true}\,\right] &= \Pr\left[\,g = t\,\right] \\
&= \Pr\left[\,g = 1 \mid t = 1\,\right]\Pr\left[\,t = 1\,\right] + \Pr\left[\,g = 0 \mid t = 0\,\right]\Pr\left[\,t = 0\,\right] \\
&= a_1 c_1 + b_0 c_0 \; .
\end{aligned}
$$

Also, using the fact that $g, t, t' \in \{0, 1\}$,

$$
\begin{aligned}
\Pr\left[\,\mathbf{Exp}^{\text{css-0}}_{\Pi,A}(k) \Rightarrow \mathsf{false}\,\right] &= \Pr\left[\,g = t'\,\right] \\
&= \Pr\left[\,g = t \wedge t = t'\,\right] + \Pr\left[\,g \neq t \wedge t \neq t'\,\right] \\
&= \Pr\left[\,g = 1 \wedge t = 1 \wedge t' = 1\,\right] + \Pr\left[\,g = 0 \wedge t = 0 \wedge t' = 0\,\right] \\
&\quad + \Pr\left[\,g = 1 \wedge t = 0 \wedge t' = 1\,\right] + \Pr\left[\,g = 0 \wedge t = 1 \wedge t' = 0\,\right] \; .
\end{aligned}
$$

But the event "$t' = 1$" is independent of its conjuncts, and similarly for "$t' = 0$" so

$$
\begin{aligned}
\Pr\left[\,\mathbf{Exp}^{\text{css-0}}_{\Pi,A}(k) \Rightarrow \mathsf{false}\,\right] &= \Pr\left[\,g = 1 \wedge t = 1\,\right]\Pr\left[\,t' = 1\,\right] + \Pr\left[\,g = 0 \wedge t = 0\,\right]\Pr\left[\,t' = 0\,\right] + \\
&\quad \Pr\left[\,g = 1 \wedge t = 0\,\right]\Pr\left[\,t' = 1\,\right] + \Pr\left[\,g = 0 \wedge t = 1\,\right]\Pr\left[\,t' = 0\,\right] \\
&= \Pr\left[\,g = 1 \mid t = 1\,\right]\Pr\left[\,t = 1\,\right]\Pr\left[\,t' = 1\,\right] + \\
&\quad \Pr\left[\,g = 0 \mid t = 0\,\right]\Pr\left[\,t = 0\,\right]\Pr\left[\,t' = 0\,\right] + \\
&\quad \Pr\left[\,g = 1 \mid t = 0\,\right]\Pr\left[\,t = 0\,\right]\Pr\left[\,t' = 1\,\right] + \\
&\quad \Pr\left[\,g = 0 \mid t = 1\,\right]\Pr\left[\,t = 1\,\right]\Pr\left[\,t' = 0\,\right] \\
&= a_1 c_1^2 + b_0 c_0^2 + b_1 c_0 c_1 + a_0 c_1 c_0 \; .
\end{aligned}
$$

The last equality uses the facts that $\Pr[t' = 1] = \Pr[t = 1] = c_1$ and $\Pr[t' = 0] = \Pr[t = 0] = c_0$. Now

$$
\begin{aligned}
\mathbf{Adv}^{\text{css}}_{\Pi,A}(k) &= \Pr\left[\,\mathbf{Exp}^{\text{css-1}}_{\Pi,A}(k) \Rightarrow \mathsf{true}\,\right] - \Pr\left[\,\mathbf{Exp}^{\text{css-0}}_{\Pi,A}(k) \Rightarrow \mathsf{false}\,\right] \\
&= a_1 c_1 + b_0 c_0 - a_1 c_1^2 - b_0 c_0^2 - b_1 c_0 c_1 - a_0 c_1 c_0 \\
&= a_1(c_1 - c_1^2) + b_0(c_0 - c_0^2) - (a_0 + b_1) c_0 c_1
\end{aligned}
$$

but $c_0 c_1 = c_1 - c_1^2$ and also $c_0 c_1 = c_0 - c_0^2$ so

$$
\begin{aligned}
\mathbf{Adv}^{\text{css}}_{\Pi,A}(k) &= (a_1 + b_0 - b_1 - a_0)(c_1 - c_1^2) \\
&= (2a_1 - 2b_1)(c_1 - c_1^2) \\
&= (a_1 - b_1)(2c_1 - 2c_1^2) \; . \blacksquare
\end{aligned}
$$

Proof of Claim E.2: Consider the experiment of Lemma E.4. Then $\Pr[\mathbf{Exp}^{\text{ind-1}}_{\Pi,I'}(k) \Rightarrow \mathsf{true}] = a_1$ and $\Pr[\mathbf{Exp}^{\text{ind-0}}_{\Pi,I'}(k) \Rightarrow \mathsf{false}] = b_1$ so $\mathbf{Adv}^{\text{ind}}_{\Pi,I'}(k) = a_1 - b_1$. By Lemma E.4

$$
\mathbf{Adv}^{\text{css}}_{\Pi,A}(k) = \mathbf{Adv}^{\text{ind}}_{\Pi,I'}(k) \cdot (2c_1 - 2c_1^2) \leq 2 \cdot \mathbf{Adv}^{\text{ind}}_{\Pi,I'}(k)
$$

where we have used the fact that $2c_1 - 2c_1^2$ is maximized when $c_1 = 1/2$. $\blacksquare$

Proof of Claim E.3: Fix $b \in \{0,1\}$, $x \in \{0,1\}^*$, and $i \in [1..v(k)]$. Let $S(x,i) = \{\, \mathbf{x} \,:\, \mathbf{x}[i] = x \,\}$. Then

$$\begin{aligned}
\Pr\left[\,\mathbf{x}[i] = x :\ \mathbf{x} \leftarrow\!\!\$\ I_{\mathrm{m}}(1^k, b)\,\right] &= \sum_{\mathbf{x} \in S(x,i)} \gamma_b(\mathbf{x}) \\
&\leq \sum_{\mathbf{x} \in S(x,i)} \alpha_b(\mathbf{x}) \frac{1 - p_{1-b}^n}{p_b} + p_{1-b}^{n-1} \alpha_{1-b}(\mathbf{x}) \\
&= \frac{1 - p_{1-b}^n}{p_b} \cdot \sum_{\mathbf{x} \in S(x,i)} \alpha_b(\mathbf{x}) + p_{1-b}^{n-1} \cdot \sum_{\mathbf{x} \in S(x,i)} \alpha_{1-b}(\mathbf{x}) \\
&\leq \frac{1}{p_b} \sum_{\mathbf{x} \in S(x,i)} (\alpha_b(\mathbf{x}) + \alpha_{1-b}(\mathbf{x})) \\
&= \frac{1}{p_b} \Pr\left[\,\mathbf{x}[i] = x :\ (\mathbf{x}, t) \leftarrow\!\!\$\ A_{\mathrm{m}}(1^k)\,\right] \\
&\leq \frac{1}{p_b} \cdot 2^{-\mu(k)} \\
&\leq \frac{1}{1/2 - \delta} \cdot 2^{-\mu(k)} \\
&= 2^{-\mu(k)+1-\log(1-2\delta)}
\end{aligned}$$

and since $\delta < 1/2$ this is well-defined. ▌

F Proof of Lemma 5.2

Let $D = (D_{\mathrm{c}}, D_{\mathrm{p}}, D_{\mathrm{g}})$. The following allows us to reduce to the case $v(\cdot) = 1$.

Claim F.1 There is a PRG-adversary $D' = (D'_{\mathrm{c}}, D'_{\mathrm{p}}, D'_{\mathrm{g}})$ with help length $\ell(\cdot)$ such that for all $k \in \mathbb{N}$

$$\epsilon(k) \leq v(k) \cdot \mathbf{Adv}^{\text{prg-1}}_{\mathcal{TP},D',n}(k)\ .$$

The running time of D' is $T_D + \mathcal{O}(nv) \cdot T_F$. □

Proof: The proof is a simple hybrid argument. Adversary D' is shown in Figure 8. We highlight the fact that $\mathbf{x}, i$ need to be chosen by D'_{c} and put into st'. This is important to ensure that the help-length of D' stays equal to that of D. We omit the analysis. ▌

A prediction adversary $P = (P_{\mathrm{c}}, P_{\mathrm{p}}, P_{\mathrm{g}})$ is a triple of algorithms. We let

$$\mathbf{Adv}^{\mathrm{pre}}_{\mathcal{TP},P}(k) = 2 \cdot \Pr\left[\,\mathbf{Exp}^{\mathrm{pre}}_{\mathcal{TP},P}(k) \Rightarrow \mathsf{true}\,\right] - 1$$

where the experiment is shown in Figure 9. The running time of P is defined as the sum of the running times of P_{c} and P_{g}, so that P is PT if $P_{\mathrm{c}}, P_{\mathrm{g}}$ are PT. (P_{p} is not required to be PT.) We say that P has help-length $\ell(\cdot)$ if the output of $P_{\mathrm{p}}(1^k, \cdot, \cdot, \cdot)$ is always of length $\ell(k)$.

Claim F.2 Let $D' = (D_{\mathrm{c}}, D_{\mathrm{p}}, D_{\mathrm{g}})$ be a PRG-adversary. Then there is a prediction adversary $P = (P_{\mathrm{c}}, P_{\mathrm{p}}, P_{\mathrm{g}})$ such that for all $k \in \mathbb{N}$

$$\mathbf{Adv}^{\text{prg-1}}_{\mathcal{TP},D',n}(k) \leq n(k) \cdot \mathbf{Adv}^{\mathrm{pre}}_{\mathcal{TP},P}(k)\ .$$

The running time of P is $T_{D'} + \mathcal{O}(n) \cdot T_F$. □

```
algorithm $D'_{\mathrm{c}}(1^k, \phi)$:
$i \leftarrow\$ \{1, \ldots, v(k)\}$
$\mathbf{x} \leftarrow\$ B_k^{v(k)}$
$st \leftarrow\$ D_{\mathrm{c}}(1^k, \phi)$
$st' \leftarrow (st, \mathbf{x}, i)$
Ret $st'$
```

```
algorithm $D'_{\mathrm{p}}(1^k, x, \phi, st')$:
$(st, \mathbf{x}, i) \leftarrow st'$
$\mathbf{x}[i] \leftarrow x$
$t \leftarrow\$ D_{\mathrm{p}}(1^k, \mathbf{x}, \phi, st)$
Ret $t$
```

```
algorithm $D'_{\mathrm{g}}(1^k, \phi, y, \omega, s, t)$:
$(st, \mathbf{x}, i) \leftarrow st'$
For $j = 1, \ldots, v(k)$ do
    $\mathbf{y}[j] \leftarrow F_\phi^{n(k)}(\mathbf{x}[j])$
    If $j \le i - 1$ then
        $\boldsymbol{\omega}[j] \leftarrow \mathcal{G}_{\mathcal{TP}}(1^k, 1^{n(k)}, \phi, \mathbf{x}[j], s)$
    Else $\boldsymbol{\omega}[j] \leftarrow\$ B_{n(k)}$
$\mathbf{y}[i] \leftarrow y$ ; $\boldsymbol{\omega}[i] \leftarrow \omega$
$b' \leftarrow\$ D_{\mathrm{g}}(1^k, \phi, st, \mathbf{y}, \boldsymbol{\omega}, s, t)$
Ret $b'$
```

Figure 8: Adversary D' for proof of Claim F.1.

```
Experiment $\mathbf{Exp}^{\mathrm{pre}}_{\mathcal{TP},P}(k)$
$(\phi, \tau) \leftarrow\$ G(1^k)$ ; $st \leftarrow\$ P_{\mathrm{c}}(1^k, \phi)$
$x \leftarrow\$ \{0,1\}^k$ ; $t \leftarrow\$ P_{\mathrm{p}}(1^k, x, \phi, st)$
$y \leftarrow F_\phi(x)$ ; $s \leftarrow\$ \{0,1\}^k$
$c \leftarrow\$ P_{\mathrm{g}}(1^k, \phi, st, y, s, t)$
Ret $(c = \langle x, s \rangle)$
```

Figure 9: Experiment defining advantage of prediction adversary $P = (P_{\mathrm{c}}, P_{\mathrm{p}}, P_{\mathrm{g}})$.

Proof: Adversary P is shown in Figure 10. We focus on the aspects related to help information, meaning what is different from the standard argument. In this regard we note that P_{p} runs D_{p} but on a value x' obtained by iterating F_ϕ backwards on x some number of times. This means P_{p} needs to invert F_ϕ and is not PT, but we allowed that. (Its running time is not counted in that of P.) Also the guess index i is chosen by P_{c} and put in st so that P can keep its help length equal to that of D'. We omit the hybrid argument used in the analysis. ▌

The final step uses the Goldreich-Levin theorem [24] whose core is captured by the following.

Lemma F.3 There is an algorithm REC such that for all $k \in \mathbb{N}$ and all $x \in \{0,1\}^k$ the following is true. Let $\mathsf{B}\colon \{0,1\}^k \to \{0,1\}$ be an oracle such that

$$2 \cdot \Pr\left[\, \mathsf{B}(s) = \langle x, s \rangle \;:\; s \leftarrow\$ \{0,1\}^k \,\right] - 1 \ge \delta > 0 \,. \tag{17}$$

Let Eq be an oracle that on any input w returns true if $x = w$ and false otherwise. Then

$$\Pr\left[\, \mathsf{REC}^{\mathsf{B},\mathsf{Eq}}(1^k) = x \,\right] \ge \frac{1}{2} \,. \tag{18}$$

The running time of REC is $\mathcal{O}(k^3 \cdot \delta^{-4})$. It makes $\mathcal{O}(k^2 \cdot \delta^{-2})$ calls to oracle B and $\mathcal{O}(k\delta^{-2})$ calls to oracle Eq. □

The running time above is in the model where an oracle call has unit cost. The probability in (18) is over the coins of REC only, and that in (17) is over the choice of s only. (B is deterministic.) A proof of the above, following Rackoff's simplification to [24], can be found in [1]. Let $P = (P_{\mathrm{c}}, P_{\mathrm{p}}, P_{\mathrm{g}})$ be the prediction adversary given by Claim F.2 applied to the PRG-adversary D' of Claim F.1. Let $\gamma(\cdot) = \mathbf{Adv}^{\mathrm{pre}}_{\mathcal{TP},P}(\cdot) > 0$. Let $I = (P_{\mathrm{c}}, P_{\mathrm{p}}, I_{\mathrm{s}})$ where algorithm I_{s} is shown in Figure 11. We claim that this inversion adversary

```
**adversary** $P_{\mathrm{c}}(1^k, \phi)$:
$i \leftarrow\$ \{1, \ldots, v(k)\}$
$st' \leftarrow\$ D'_{\mathrm{c}}(1^k, \phi)$
$st \leftarrow (i, st')$
Ret $st$
```

```
**adversary** $P_{\mathrm{p}}(1^k, x, \phi, st)$:
$(i, st') \leftarrow st$
$x' \leftarrow F_{\phi}^{-(i-1)}(x)$
$t \leftarrow\$ D'_{\mathrm{p}}(1^k, x', \phi, st')$
Ret $t$
```

```
**adversary** $P_{\mathrm{g}}(1^k, \phi, st, y, s, t)$:
$(i, st') \leftarrow st$
For $j = 1, \ldots, n(k)$ do
    If $j \leq i$ then $\omega_j \leftarrow\$ \{0, 1\}$
    Else $\omega_j \leftarrow \langle F_{\phi}^{j-i-1}(y), s \rangle$
$\omega \leftarrow \omega_1 \cdots \omega_{n(k)}$
$b' \leftarrow\$ D'_{\mathrm{g}}(1^k, \phi, st', F_{\phi}^{n(k)-i}(y), \omega, s, t)$
Ret $b' \oplus \omega_i \oplus 1$
```

Figure 10: Prediction adversary $P = (P_{\mathrm{c}}, P_{\mathrm{p}}, P_{\mathrm{g}})$ for proof of Claim F.2.

```
**algorithm** $I_{\mathrm{s}}(1^k, \phi, st, y, t)$:
$R_g \leftarrow\$ \mathsf{Coins}^{P_{\mathrm{g}}}(1^k)$

  **oracle** $\mathsf{B}(s)$:
  $c \leftarrow P_{\mathrm{g}}(1^k, \phi, st, y, s, t\,;\ R_g)$
  Ret $c$

  **oracle** $\mathsf{Eq}(w)$:
  Ret $(F_{\phi}(w) = y)$

$x' \leftarrow\$ \mathsf{REC}^{\mathsf{B},\mathsf{Eq}}(1^k)$
Ret $x'$
```

Figure 11: Algorithm I_{s} for proof of Lemma 5.2, where $\mathsf{Coins}^{P_{\mathrm{g}}}(1^k)$ is the space of coins for $P_{\mathrm{g}}(1^k, \cdot, \cdot, \cdot, \cdot, \cdot)$.

satisfies the conditions of Lemma 5.2. For the analysis let us enumerate the coins underlying $\mathbf{Exp}^{\mathrm{pre}}_{\mathcal{TP},P}(k)$ as R_G, R_c, x, R_p, s, R_g where R_G, R_c, R_p, R_g are the coins of G, P_{c}, P_{p} and P_{g} respectively. Let

$$\Gamma(R_G, R_c, x, R_p, R_g) = 2 \cdot \Pr\left[\, c = \langle x, s \rangle \,\right] - 1$$

where the probability is over the choice of s alone and the other coins in the experiment are fixed to the given values. Then $\gamma(k) = \mathbf{E}\left[\Gamma\right]$. So a standard averaging argument says there is a set Ω of choices of (R_G, R_c, x, R_p, R_g) that has probability at least $\gamma(k)/2$ and

$$\Gamma(R_G, R_c, x, R_p, R_g) \geq \frac{\gamma(k)}{2}$$

for all $(R_G, R_c, x, R_p, R_g) \in \Omega$. Now Lemma F.3 with $\delta = \gamma(k)/2$ implies

$$\mathbf{Adv}^{\mathrm{owf}}_{\mathcal{TP},I}(k) \geq \frac{\gamma(k)}{2} \cdot \frac{1}{2} = \frac{\gamma(k)}{4} \,.$$

Putting everything together we have

$$\begin{aligned} \epsilon(k) &\leq v(k) \cdot \mathbf{Adv}^{\text{prg-1}}_{\mathcal{TP},D',n}(k) \\ &\leq v(k) \cdot n(k) \cdot \mathbf{Adv}^{\mathrm{pre}}_{\mathcal{TP},P}(k) \\ &= v(k) \cdot n(k) \cdot \gamma(k) \\ &\leq 4v(k) \cdot n(k) \cdot \mathbf{Adv}^{\mathrm{owf}}_{\mathcal{TP},I}(k) \,. \end{aligned}$$

With regard to running time we have

$$\begin{aligned}
T_I &= \mathcal{O}(k^3\gamma^{-4}) + \mathcal{O}(k^2\gamma^{-2})\cdot T_P + \mathcal{O}(k\gamma^{-2})\cdot T_F \\
&= \mathcal{O}(k^3\gamma^{-4}) + \mathcal{O}(k^2\gamma^{-2})\cdot [T_D + \mathcal{O}(nv)\cdot T_F] + \mathcal{O}(k\gamma^{-2})\cdot T_F \\
&= \mathcal{O}(k^2\gamma^{-2})\cdot T_D + \mathcal{O}(k^2\gamma^{-2}nv)\cdot T_F + \mathcal{O}(k^3\gamma^{-4}) \\
&= \mathcal{O}(k^2(nv/\epsilon)^2)\cdot T_D + \mathcal{O}(k^2(nv/\epsilon)^2 nv)\cdot T_F + \mathcal{O}(k^3(nv/\epsilon)^4) \\
&= \mathcal{O}(k^3n^4v^4\epsilon^{-4}) + \mathcal{O}(T_D + nvT_F)k^2n^2v^2\epsilon^{-2} \ .
\end{aligned}$$

G Deterministic Encryption under Chosen Ciphertext Attacks

The definitions of Section 3 lift in a natural way to model chosen-ciphertext attacks. We denote the CCA versions of the experiments by a -CCA suffix. Let $A = (A_{\mathrm{c}}, A_{\mathrm{m}}, A_{\mathrm{g}})$ be an SS-adversary and let S be a simulator. Then the css experiment is modified to give A_{g} access to a decryption oracle. The sss experiment is modified to give both A_{g} and S access to a decryption oracle. We now say that adversary A is legitimate if —in addition to the existing constraints detailed in Section 3— $A_{\mathrm{g}}(1^k, pk, \mathbf{c}, st)$ does not query the decryption oracle on any $c \in \mathbf{c}$. Let $I = (I_{\mathrm{c}}, I_{\mathrm{m}}, I_{\mathrm{g}})$ be an IND-adversary. The ind experiment is modified to allow I_{g} access to a decryption oracle. The definition of legitimacy is similarly adapted. It is straightforward to modify the theorem statements and proofs of our implications in Section 4 and Appendices B, C, D, and E to the CCA setting.

The construction of randomized PKE from deterministic PKE in Appendix 7 also holds in the CCA setting. Specifically, if Π is secure against IND-CCA adversaries, then Ψ is secure as a KEM against CCA attacks. Modifying the proof just requires adding a decryption oracle in the appropriate places. As discussed in the introduction, this is of particular interest because our construction implies that building a CCA-secure deterministic encryption scheme is at least as hard as building a witness-recovering CCA encryption scheme.

H Related Work

To discuss prior work we let us say a that an adversary is efficient if it is polynomial time and a message space is efficient if it is poly-time sampleable. Then a reduction is efficient if, whenever the starting adversary and message space are efficient, so are the resulting adversary and message space. The reason efficient message spaces are important is because they are necessary whenever one uses the definitions in computationally-bounded settings. For example, the reductions in [3] require efficiently-sampleable message spaces.

ENTROPIC SECURITY. The core concern of Dodis and Smith's work on entropic security [20] is the same as ours, namely the encryption of high min-entropy plaintexts. But there are important differences between the settings, namely that of entropic security is information theoretic and symmetric while ours is computational and public-key. The first difference means that adversaries and message spaces in the entropic security setting need not be efficient while in our setting, in contrast, they must be efficient. They introduce notions that are analogous to our notions IND, B-SSS and BB-SSS, to complement the A-SSS-like definition of Russell and Wang [28], and they show equivalences in their setting. Note that in their definitions adversaries are restricted to seeing the encryption of a single message, which is not in general equivalent to our multi-message definitions. It becomes equivalent when one restricts our definitions to the case of independent and separable adversaries, as discussed in Appendix A. Below we implicitly mean our definitions so restricted.

Dodis and Smith [20] provide implications showing that IND, A-SSS, and BB-SSS are equivalent, but their non-trivial reductions from BB-SSS to B-SSS and B-SSS to A-SSS are inefficient, so their results do not imply equivalences in our setting. They do provide an efficient reduction from B-SSS to A-SSS

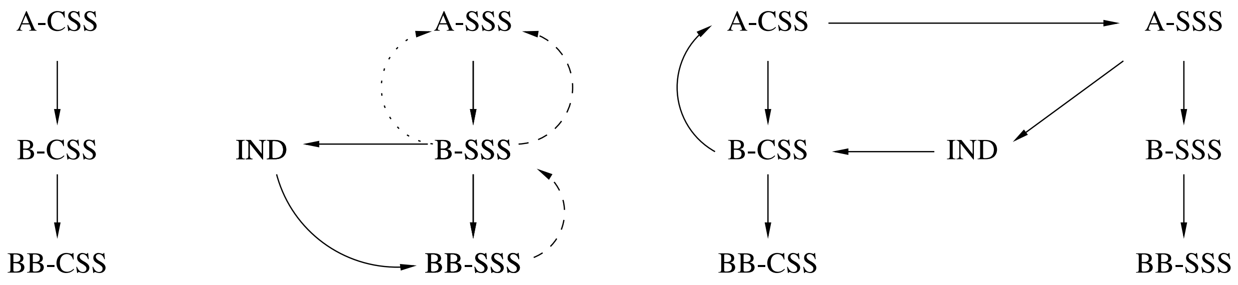

Figure 12: **(Left)** Diagram illustrating implications shown in [20]. Dashed lines are implications relying on an inefficient reduction. The dotted line represents an implication relying on an unnatural adversarial constraint. **(Right)** Diagram illustrating implications shown in [17, 18].

when the probability of the most likely output of the partial information function has a certain relation to the adversary advantage, but this does not show that the one notion implies the other in general. The left diagram in Figure 12 diagrams these results. Note that we show A-CSS, B-CSS, and BB-CSS and their associated trivial implications in the diagram for completeness, but these were not considered in [20].

QUANTUM ENTROPIC SECURITY. Desrosiers [17] and Desrosiers and Dupuis [18] adapted entropic security to the quantum setting. Moreover, they define notions analogous to A-CSS and B-CSS. As above, these are single-message definitions. They provide an efficient reduction showing that A-CSS implies A-SSS. (For completeness we also give a theorem and its proof for this straightforward implication; see Appendix C.) They use a Goldreich-Levin predicate as the main tool of an efficient reduction showing that B-CSS implies A-CSS. Of greater technical interest is their proof that IND implies B-CSS. Here they aim to build an IND-adversary $I = (I_{\rm c}, I_{\rm m}, I_{\rm g})$ from a boolean SS-adversary $A = (A_{\rm c}, A_{\rm m}, A_{\rm g})$. Associate to $A_{\rm m}$ the (implicitly defined) message distribution $\mathcal{M}$ and the (implicitly defined) boolean function f. Let $\mathcal{M}_b$ be $\mathcal{M}$ but conditioned on a message being in the preimage of f for b. Ideally, $I_{\rm m}$, when run with input bit b, could sample a message from $\mathcal{M}_b$ and return it. However, since A, and therefore f, are not balanced, $\mathcal{M}_b$ might be low entropy. Instead, they have $I_{\rm m}$ sample from a convex combination of $\mathcal{M}_b$ and $\mathcal{M}$, which ensures high min-entropy, but nevertheless allows $I_{\rm g}$ to utilize $A_{\rm g}$ to infer the bit b with probability close to A's advantage. Note that this does not change the balance of f, but cleverly modifies the way messages are chosen by $A_{\rm m}$ to compensate for f's lack of balance. The reduction is efficient as long as the combination of $\mathcal{M}_b$ and $\mathcal{M}$ is efficiently sampleable, which appears to be the case [29].

The right diagram of Figure 12 shows the relationships established in [17, 18]. Note that we show BB-CSS, B-SSS, and BB-SSS and their associated trivial implications in the diagram for completeness, but these were not considered in [17, 18]. Indeed, their implication that IND implies B-CSS (as sketched above) side-steps the issue of balanced predicates entirely.

In light of the implications provided by [17, 18], one might ask why bother with the BB notions at all, since IND can apparently be shown equivalent to the others without them. There are several reasons we nevertheless consider them. The work of [20], and also our implications, highlight balance as a useful tool for understanding the relationships between security notions for deterministic encryption. Most importantly, the BB-CSS and BB-SSS notions are conceptually close to IND, and the balance feature allows an obviously efficient reduction from IND to BB-CSS. (Whereas adapting the [17, 18] reduction from IND to B-CSS to the computationally-efficient setting requires a non-obvious sampling algorithm.) Moreover, intuition might predict that the BB notions are not as strong as their boolean or arbitrary counterparts. Our results show otherwise. Finally, we expect that these notions might be useful in future applications of deterministic encryption.

PERFECTLY ONE-WAY HASH FUNCTIONS. Canetti introduced perfectly one-way hash functions (POWHF) [13], which were further studied by Canetti, Micciancio, and Reingold [14]. These are randomized hash functions that produce publically-verifiable outputs (i.e., given a message and a hash value, any party can check if the hash corresponds to the message). The security required is that no adversary, given only the output of the hash applied to some unpredictable message, should be able to compute any partial information about the message. In [13] a notion analogous to B-SSS is introduced. Several other definitions are offered in [13, 14], along with some equivalences, but these definitions, and the implications, are only meaningful for randomized primitives. Our definitions and equivalences can be adapted to work in this setting.