MTAT.07.003 Cryptology II
Spring 2012 / Exercise session ?? / Example Solution

Exercise (Non-canonical knowledge-extractor for the QNR-ZKD protocol). *Let n be a composite number with a factorisation $n = pq$ known to the prover $\mathcal{P}$. Let $v \in \mathbb{Z}_n^*$ be a number for which the prover wants to prove that it is quadratic non-residue. Let $(\mathsf{Gen}, \mathsf{Com}, \mathsf{Open})$ be a perfectly binding and computationally hiding commitment. Them we can define the following zero-knowledge protocol*

$\mathcal{P}$		$\mathcal{V}$
$\mathsf{pk} \leftarrow \mathsf{Gen}$		$b \xleftarrow[u]{} \{0,1\}$,
	$\xleftarrow{\quad c = r^2 v^b \quad}$	$r \xleftarrow[u]{} \mathbb{Z}_n^*$
$\overline{b} \leftarrow \mathsf{IsQnr}_{p,q}(c)$		
$(\overline{c}, \overline{d}) \leftarrow \mathsf{Com}(\overline{b})$	$\xrightarrow{\quad \mathsf{pk}, \overline{c} \quad}$	
	$\xleftarrow{\quad r, b \quad}$	
$c \overset{?}{=} r^2 v^b$		
	$\xrightarrow{\quad \overline{d} \quad}$	$\mathsf{Open}_{\mathsf{pk}}(\overline{c}, \overline{d}) \overset{?}{=} b$

Show that the following knowledge extraction algorithm

```
$\mathcal{K}_\circ^{\mathcal{V}_*}(\phi, \ell; \omega)$
    $c \leftarrow \mathcal{V}_*(\phi; \omega)$
    For $i \in \{1, \ldots, l\}$ do
        $\mathsf{pk} \leftarrow \mathsf{Gen}$
        $\overline{b} \xleftarrow[u]{} \{0,1\}$
        $(\overline{c}, \overline{d}) \leftarrow \mathsf{Com}_{\mathsf{pk}}(\overline{b})$
        $(b, r) \leftarrow \mathcal{V}_*(\mathsf{pk}, \overline{c})$
        if $(c = r^2 v^b)$ then **return** $(r, b)$
    **return** $\perp$
```

is reasonably successful even if the commitment scheme is not hiding, e.g. the commitment digest reveals $\overline{b}$. Draw the corresponding time-success profile for $\mathcal{K}^{\mathcal{V}_*}$ and compare it with the standard knowledge extractor construction. Explain why the standard simulator construction described below

```
$\mathcal{V}_\circ(\phi)$
    $\omega \leftarrow \Omega$
    $c \leftarrow \mathcal{V}_*(\phi; \omega)$
    $(b_*, r_*) \leftarrow \mathcal{K}^{\mathcal{V}_*}(\phi, \ell; \omega)$
    if $(b_*, r_*) = \perp$ then $b_* \xleftarrow[u]{} \{0,1\}$
    $\mathsf{pk} \leftarrow \mathsf{Gen}$
    $(\hat{c}, \hat{d}) \leftarrow \mathsf{Com}_{\mathsf{pk}}(b_*)$
    $r, b \leftarrow \mathcal{V}_*(\mathsf{pk}, \hat{c})$
    if $r^2 b = c$ then **return** $\mathcal{V}_*(\hat{d})$
    else **return** $\mathcal{V}_*(\perp)$
```

will fail if the commitment is non-hiding. Finally show how a malicious verifier can gain some knowledge by interaction if the commitment is non-hiding.

Solution. To outline the the problem, we first do the standard analysis under the assumption that the commitment scheme is hiding and then provide the alternative analysis, which does not require hiding. Then we study why the simulation still fails although knowledge-extraction results are comparable in both cases.

Standard analysis. For the analysis of the knowledge extraction, it is instructive to contrast what happens in the single cycle with the protocol execution. For that we can construct the following games:

```
$\mathcal{G}_0^{\mathcal{V}_*}$
    $\mathsf{pk} \leftarrow \mathsf{Gen}$
    $\overline{b} \leftarrow \{0,1\}$
    $(\overline{c}, \overline{d}) \leftarrow \mathsf{Com}_{\mathsf{pk}}(\overline{b})$
    $(b, r) \leftarrow \mathcal{V}_*(\mathsf{pk}, \overline{c})$
    $\mathbf{return}\ [c \overset{?}{=} r^2 v^b]$
```

```
$\mathcal{G}_1^{\mathcal{V}_*}$
    $\mathsf{pk} \leftarrow \mathsf{Gen}$
    $\overline{b} \leftarrow \mathsf{IsQnr}_{p,q}(c)$
    $(\overline{c}, \overline{d}) \leftarrow \mathsf{Com}_{\mathsf{pk}}(\overline{b})$
    $(b, r) \leftarrow \mathcal{V}_*(\mathsf{pk}, \overline{c})$
    $\mathbf{return}\ [c \overset{?}{=} r^2 v^b]$
```

where values c, p, q are hardwired. Note that hardwiring does not weaken the properties of the commitment as the commitment parameters pk are chosen independently from c, p, q. Formally, we can use the reduction construction

```
$\mathcal{B}^{\mathcal{V}_*}(\mathsf{pk})$
    $m_0 \leftarrow \{0,1\}$
    $m_1 \leftarrow \mathsf{IsQnr}_{p,q}(c)$
    $\mathbf{return}\ (m_0, m_1)$
```

```
$\mathcal{B}^{\mathcal{V}_*}(\overline{c})$
    $(b, r) \leftarrow \mathcal{V}_*(\mathsf{pk}, \overline{c})$
    if $c = r^2 v^b$ $\mathbf{return}$ 1
    else $\mathbf{return}$ 0
```

against hiding games. Since c, p, q are hardwired the overhead of $\mathcal{B}$ is only constant. Moreover, since c is fixed we can precompute the value $\mathsf{IsQnr}_{p,q}(c)$ ant hardwire only this value as m_1. Now it is straightforward to see that

$$\mathsf{Adv}^{\mathsf{hiding}}_{\mathfrak{C}}(\mathcal{B}) = |\Pr[\mathcal{G}_0^{\mathcal{V}_*} = 1] - \Pr[\mathcal{G}_0^{\mathcal{V}_*} = 1]|$$

and thus the success probability in the knowledge extraction cycle can drop only by $\varepsilon_\circ$ if the commitment is $(t, \varepsilon_\circ)$-hiding and the verifier $\mathcal{V}_*$ is t-time algorithm. Consequently, if $\varepsilon = \varepsilon(\phi, \omega)$ is the probability that $\mathcal{V}_*(\phi, \omega)$ passes the knowledge proof (verification $c = r^2 v^b$), then one iteration of the knowledge extractor will succeed with probability $\varepsilon - \varepsilon_\circ$. Hence, the expected running time of a knowledge extractor

$$\tau = \frac{1}{\varepsilon - \varepsilon_\circ}$$

and the probability that the knowledge extractor fails after ℓ tries

$$p_{\mathrm{fail}} \leq (1 + \varepsilon_\circ - \varepsilon)^\ell \enspace .$$

Notice that the larger the maximal hiding advantage $\varepsilon_\circ$ is the longer it takes to extract parameters on average and the higher is the failure probability. For the comparison, note that standard way to estimate failure probability through Markov inequality and expected running time gives the following bound

$$p_{\mathrm{fail}} \leq 2^{-\left\lfloor \frac{\ell(\varepsilon - \varepsilon_\circ)}{2} \right\rfloor} \enspace .$$

Figure 1 gives an illustrative comparison of bounds and shows the effect $\varepsilon_\circ$ on the failure probability. It easy to see that the value $\varepsilon_\circ$ alters the slope of the graph non-marginally only if

$$\alpha = \ln(1 + \varepsilon_\circ - \varepsilon) - \ln(1 - \varepsilon) \approx \frac{\varepsilon_\circ}{1 - \varepsilon}$$

is large. The latter can occur in two cases. The case $\varepsilon_\circ \gg 0$ implies that the commitment is non-hiding. The other case corresponds to the case where the verifier fails the proof of knowledge with probability that is comparable to $\varepsilon_\circ$. Then indeed the exact value of $\varepsilon_\circ$ has large impact on the success probability.

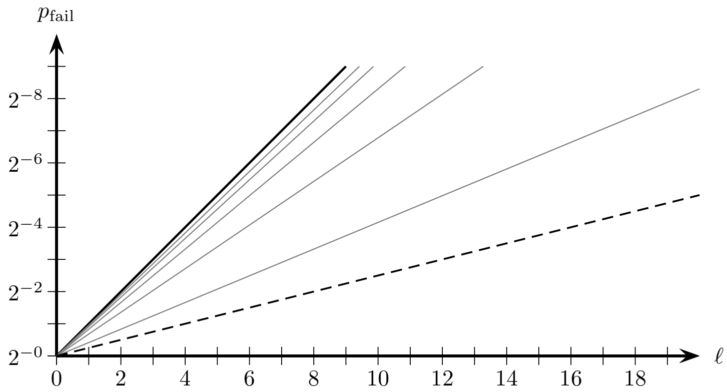


Figure 1: Knowledge-extraction failure as function of iterations ℓ. The upper black line corresponds to the verification success $\varepsilon = \frac{1}{2}$ and the lower dashed line corresponds to the lower bound obtained through the expected running-time. Grey lines between are the exact bounds for different values of $\varepsilon_\circ \in \{2^{-2}, 2^{-3}, 2^{-4}, 2^{-5}, 2^{-6}\}$. Clearly, the lower bound based on the expected running time is quite loose and the effect on the failure probability is marginal when $\varepsilon_\circ \ll \varepsilon$.

Non-standard analysis. Let us again start by contrasting what happens in the single cycle with the protocol execution. For that we use the same games as in the standard analysis:

```
$\mathcal{G}_0^{\mathcal{V}_*}$
    $\mathsf{pk} \leftarrow \mathsf{Gen}$
    $\overline{b} \leftarrow \{0,1\}$
    $(\overline{c}, \overline{d}) \leftarrow \mathsf{Com}_{\mathsf{pk}}(\overline{b})$
    $(b, r) \leftarrow \mathcal{V}_*(\mathsf{pk}, \overline{c})$
    return $[c \overset{?}{=} r^2 v^b]$

$\mathcal{G}_1^{\mathcal{V}_*}$
    $\mathsf{pk} \leftarrow \mathsf{Gen}$
    $\overline{b} \leftarrow \mathsf{IsQnr}_{p,q}(c)$
    $(\overline{c}, \overline{d}) \leftarrow \mathsf{Com}_{\mathsf{pk}}(\overline{b})$
    $(b, r) \leftarrow \mathcal{V}_*(\mathsf{pk}, \overline{c})$
    return $[c \overset{?}{=} r^2 v^b]$ .
```

Since $\overline{b}$ is generated randomly on the left, its value will coincide with the correct value on the left with probability $\frac{1}{2}$. Consequently, if $\varepsilon = \varepsilon(\phi, \omega)$ is the probability that $\mathcal{V}_*(\phi, \omega)$ passes the knowledge proof (verification $c = r^2 v^b$), then one iteration of the knowledge extractor will succeed with probability

$$\Pr\left[\mathcal{G}_0^{\mathcal{V}_*} = 1\right] \geq \frac{1}{2} \cdot \Pr\left[\mathcal{G}_1^{\mathcal{V}_*} = 1\right] = \frac{\varepsilon}{2} \enspace .$$

Hence, the expected running time of a knowledge extractor is twice as slower than in standard case:

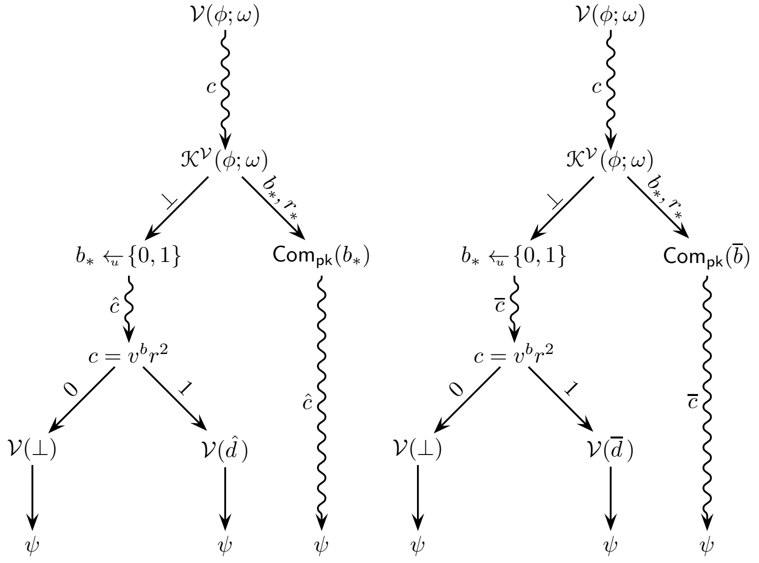
$$\tau = \frac{2}{\varepsilon}$$

and the probability that the knowledge extractor fails after ℓ tries

$$p_{\text{fail}} \leq \left(1 - \frac{\varepsilon}{2}\right)^{\ell}$$

has also comparable asymptotic behaviour in the process $\ell \to \infty$. In the range $\varepsilon_\circ \leq \varepsilon \leq 2\varepsilon_\circ$, the non-standard analysis provides actually better bounds on the failure probability.

Solution to the paradox. The analysis above established that knowledge extractor works reasonably well even if the commitment scheme is not hiding. However, it is clear that malicious verifier who creates a challenge with unknown quadratic residuosity gains new knowledge during the interaction with the honest prover – it learns the residuosity through leaking commitment. In order to clearly pinpoint where the problem occurs in the simulation, we align real protocol execution with the simulation. For that we consider a thought experiment where the knowledge extractor in the real world without using its output. As a result, we can decompose simulated and real execution into event trees with the same shape depicted in Figure 2.



$\mathcal{V}(\phi;\omega)$
c
$\mathcal{K}^{\mathcal{V}}(\phi;\omega)$
$\perp$
b_*, r_*
$b_* \xleftarrow{u} \{0,1\}$
$\mathsf{Com}_{\mathsf{pk}}(b_*)$
$\hat{c}$
$c = v^b r^2$
0
1
$\hat{c}$
$\mathcal{V}(\perp)$
$\mathcal{V}(\hat{d})$
ψ
ψ
ψ
$\mathcal{V}(\phi;\omega)$
c
$\mathcal{K}^{\mathcal{V}}(\phi;\omega)$
$\perp$
b_*, r_*
$b_* \xleftarrow{u} \{0,1\}$
$\mathsf{Com}_{\mathsf{pk}}(\overline{b})$
$\overline{c}$
$c = v^b r^2$
0
1
$\overline{c}$
$\mathcal{V}(\perp)$
$\mathcal{V}(\overline{d})$
ψ
ψ
ψ

Figure 2: Event trees for the protocol simulation and for real protocol execution.

If the knowledge extraction succeeds the simulation is perfect, since b_* must coincide with $\overline{b}$. Problems occur if the simulation fails. In this case, b_* is randomly chosen and corresponding commitment value is fed to the verifier $\mathcal{V}_*$, while $\mathcal{V}_*$ obtains a commitment of $\overline{b}$ in a real protocol run. If the commitment is non-hiding the difference in the commitment strings is enough to disturb the output of $\mathcal{V}_*$. In the extreme case where the commitment is perfectly hiding, $\hat{c}$ and $\overline{c}$ have the same distribution. Consequently, the simulation is perfect until the simulator is forced to open $\hat{c}$. The latter does not occur, if the verifier fails to release consistent values of r, b, This corresponds to the left-most path on the event tree.

Hence, we have established that the simulation might fail only for the middle path if the commitment is perfectly hiding. For brevity, let $\mathsf{SimFail}$ denote the event that the knowledge error fails while verification succeeds. Then it is easy to prove that the statistical distance between the output distributions of real world execution and simulation is

$$\Pr\left[\mathsf{SimFail}\right] = \varepsilon \cdot p_{\text{fail}} = \varepsilon(1-\varepsilon)^{\ell} \quad \leq \frac{1}{\ell} \cdot \left(1 - \frac{1}{\ell+1}\right)^{\ell+1} \approx \frac{1}{\ell e},$$

which guaranteed to be negligible for reasonably chosen ℓ.

Detailed analysis. The analysis above highlighted why the commitment must be hiding but did not quantify how much non-perfect hiding changes the bounds. Before, we go into the analysis recall a useful bound on the distinguishing advantage. If the distribution is generated by a two-stage sampling procedure:

$$\mathsf{Adv}^{\mathsf{ind}}_{f,g}(\mathcal{A}) = \left|\Pr\left[\omega \leftarrow \Omega, \psi \leftarrow f(\omega) : \mathcal{A}(\psi) = 1\right] - \Pr\left[\omega \leftarrow \Omega, \psi \leftarrow g(\omega) : \mathcal{A}(\psi) = 1\right]\right|$$

then the advantage can be bounded

$$\mathsf{Adv}^{\mathsf{ind}}_{f,g}(\mathcal{A}) \leq \sum_{\omega \in \Omega} \Pr\left[\omega\right] \mathsf{Adv}^{\mathsf{ind}}_{f(\omega),g(\omega)}(\mathcal{A}) \leq \max_{\omega \in \Omega} \mathsf{Adv}^{\mathsf{ind}}_{f(\omega),g(\omega)}(\mathcal{A})$$

where a sub-advantages are defined naturally:

$$\mathsf{Adv}^{\mathsf{ind}}_{f(\omega),g(\omega)}(\mathcal{A}) = \left|\Pr\left[\psi \leftarrow f(\omega) : \mathcal{A}(\psi) = 1\right] - \Pr\left[\psi \leftarrow g(\omega) : \mathcal{A}(\psi) = 1\right]\right|$$

and thus if all sub-distributions are (τ, ε_*)-indistinguishable then so is the main distribution.

Hence, it is sufficient to consider the sub-distributions where ω is fixed. Let $\mathcal{A}$ be a τ-time distinguisher for the verifiers output. Since the sampling procedure is still a two-stage procedure, where in the stage the knowledge-extractors output is generated, we can consider sub-distributions generated for different knowledge extractor outputs. If the output is different from $\perp$ then $b_* = \overline{c}$ and the distributions are identical. If the knowledge extractor fails, we are left with procedures

```
$D_1$
    $c \leftarrow \mathcal{V}_*(\phi; \omega)$
    $b_* \underset{u}{\leftarrow} \{0, 1\}$
    $\mathsf{pk} \leftarrow \mathsf{Gen}$
    $(\hat{c}, \hat{d}) \leftarrow \mathsf{Com}_{\mathsf{pk}}(b_*)$
    $r, b \leftarrow \mathcal{V}_*(\mathsf{pk}, \hat{c})$
    if $r^2 b = c$ then **return** $\mathcal{V}_*(\hat{d})$
    else **return** $\mathcal{V}_*(\perp)$
```

```
$D_2$
    $c \leftarrow \mathcal{V}_*(\phi; \omega)$
    $\overline{b} \leftarrow \mathsf{IsQnr}_{p,q}(c)$
    $\mathsf{pk} \leftarrow \mathsf{Gen}$
    $(\overline{c}, \overline{d}) \leftarrow \mathsf{Com}_{\mathsf{pk}}(\overline{b})$
    $r, b \leftarrow \mathcal{V}_*(\mathsf{pk}, \overline{c})$
    if $r^2 b = c$ then **return** $\mathcal{V}_*(\overline{d})$
    else **return** $\mathcal{V}_*(\perp)$
```

which output we want to distinguish. This could still be viewed as a two-stage sampling procedure, where in the first stage $\beta_* \underset{u}{\leftarrow} \{0, 1\}$ is drawn and on the second stage it is used to further (in the right game the value of b_* is ignored). There are two sub-game pairs in one $b_* = \overline{b}$ and in the other they are different. Since the game pair is indistinguishable, we have to look only the second game pair. As the order of the games does not change the distinguishability, it is sufficient to estimate the distance of the following games:

```
$D_3$
    $c \leftarrow \mathcal{V}_*(\phi; \omega)$
    $\mathsf{pk} \leftarrow \mathsf{Gen}$
    $(\hat{c}, \hat{d}) \leftarrow \mathsf{Com}_{\mathsf{pk}}(0)$
    $r, b \leftarrow \mathcal{V}_*(\mathsf{pk}, \hat{c})$
    if $r^2 b = c$ then **return** $\mathcal{V}_*(\hat{d})$
    else **return** $\mathcal{V}_*(\perp)$
```

```
$D_4$
    $c \leftarrow \mathcal{V}_*(\phi; \omega)$
    $\mathsf{pk} \leftarrow \mathsf{Gen}$
    $(\overline{c}, \overline{d}) \leftarrow \mathsf{Com}_{\mathsf{pk}}(1)$
    $r, b \leftarrow \mathcal{V}_*(\mathsf{pk}, \overline{c})$
    if $r^2 b = c$ then **return** $\mathcal{V}_*(\overline{d})$
    else **return** $\mathcal{V}_*(\perp)$ .
```

Let $\mathcal{A}$ be a τ-time distinguisher. Then we can decompose its advantage

$$\begin{aligned}\mathsf{Adv}^{\mathsf{ind}}_{D_3,D_4}(\mathcal{A}) \le{}& \left|\Pr\left[\psi \leftarrow D_3 : \mathcal{A}(\psi) = 1 \wedge r^2 b \neq c\right] - \Pr\left[\psi \leftarrow D_4 : \mathcal{A}(\psi) = 1 \wedge r^2 b \neq c\right]\right| \\ &+ \left|\Pr\left[\psi \leftarrow D_3 : \mathcal{A}(\psi) = 1 \wedge r^2 b = c\right] - \Pr\left[\psi \leftarrow D_4 : \mathcal{A}(\psi) = 1 \wedge r^2 b = c\right]\right| \enspace .\end{aligned}$$

If the running time t_v of $\mathcal{V}_*$ is small enough so that the total running time of D_3 and D_4 without the first if branch is below t then

$$\left|\Pr\left[\psi \leftarrow D_3 : \mathcal{A}(\psi) = 1 \wedge r^2 b \neq c\right] - \Pr\left[\psi \leftarrow D_3 : \mathcal{A}(\psi) = 1 \wedge r^2 b \neq c\right]\right| \le \varepsilon_\circ \enspace ,$$

or otherwise we can use the adversary

```
$\mathcal{A}(\mathsf{pk})$
    $c \leftarrow \mathcal{V}_*(\phi; \omega)$
    **return** $(0, 1)$
```

```
$\mathcal{A}(c)$
    $r, b \leftarrow \mathcal{V}_*(\mathsf{pk}, \overline{c})$
    if $r^2 b = c$ then **return** $\perp$
    else **return** $\mathcal{B}(\mathcal{V}_*(\perp))$
```

as the t-time distinguisher against hiding games. For the second term we use just a trivial upper bound. As a result, we get

$$\mathsf{Adv}^{\mathsf{ind}}_{D_3,D_4}(\mathcal{A}) \leq \max\left\{\Pr\left[\psi \leftarrow D_3 : r^2 b = c\right], \Pr\left[\psi \leftarrow D_4 : r^2 b = c\right]\right\} + \varepsilon_\circ \enspace .$$

However, the cannot bound the probabilities under the maximum operator, the malicious verifier can have arbitrary success in releasing r, b. Hence, we cannot eliminate knowledge extraction from the distribution generation in order to get reasonable bound on the distinguishing advantage.

As a result we have to analyse the computational distance of following distributions

D_1

```
$c \leftarrow \mathcal{V}_*(\phi;\omega)$
$\overline{b} \leftarrow \mathsf{IsQnr}_{p,q}(c)$
$(b_*, r_*) \leftarrow \mathcal{K}^{\mathcal{V}_*}(\phi, \ell; \omega)$
if $(b_*, r_*) = \perp$ then $b_* \xleftarrow[u]{} \{0,1\}$
$\mathsf{pk} \leftarrow \mathsf{Gen}$
$(\hat{c}, \hat{d}) \leftarrow \mathsf{Com}_{\mathsf{pk}}(b_*)$
$r, b \leftarrow \mathcal{V}_*(\mathsf{pk}, \hat{c})$
if $r^2 b = c$ then return $\mathcal{V}_*(\hat{d})$
else return $\mathcal{V}_*(\perp)$
```

D_2

```
$c \leftarrow \mathcal{V}_*(\phi;\omega)$
$\overline{b} \leftarrow \mathsf{IsQnr}_{p,q}(c)$
$(b_*, r_*) \leftarrow \mathcal{K}^{\mathcal{V}_*}(\phi, \ell; \omega)$
if $(b_*, r_*) = \perp$ then $b_* \xleftarrow[u]{} \{0,1\}$
$\mathsf{pk} \leftarrow \mathsf{Gen}$
$(\overline{c}, \overline{d}) \leftarrow \mathsf{Com}_{\mathsf{pk}}(\overline{b})$
$r, b \leftarrow \mathcal{V}_*(\mathsf{pk}, \overline{c})$
if $r^2 b = c$ then return $\mathcal{V}_*(\overline{d})$
else return $\mathcal{V}_*(\perp)$
```

where the distribution corresponding to the ideal execution ignores the value $\overline{b}$ and the distribution corresponding to the real execution ignores the value b_*. As before view can view the sampling procedure as a two-stage procedure where in the first phase b_* is sampled. Hence, we can consider two sub-distributions. One with $b_* = \overline{b}$ and the other with $b_* \neq \overline{b}$. As distributions are identical in the first case, we must analyse only the distance between

D_3

```
$c \leftarrow \mathcal{V}_*(\phi;\omega)$
$\overline{b} \leftarrow \mathsf{IsQnr}_{p,q}(c)$
$(b_*, r_*) \leftarrow \mathcal{K}^{\mathcal{V}_*}(\phi, \ell; \omega)$
if $(b_*, r_*) = \perp$ then $b_* \xleftarrow[u]{} \neg\overline{b}$
$\mathsf{pk} \leftarrow \mathsf{Gen}$
$(\hat{c}, \hat{d}) \leftarrow \mathsf{Com}_{\mathsf{pk}}(b_*)$
$r, b \leftarrow \mathcal{V}_*(\mathsf{pk}, \hat{c})$
if $r^2 b = c$ then return $\mathcal{V}_*(\hat{d})$
else return $\mathcal{V}_*(\perp)$
```

D_4

```
$c \leftarrow \mathcal{V}_*(\phi;\omega)$
$\overline{b} \leftarrow \mathsf{IsQnr}_{p,q}(c)$
$(b_*, r_*) \leftarrow \mathcal{K}^{\mathcal{V}_*}(\phi, \ell; \omega)$
if $(b_*, r_*) = \perp$ then $b_* \xleftarrow[u]{} \neg\overline{b}$
$\mathsf{pk} \leftarrow \mathsf{Gen}$
$(\overline{c}, \overline{d}) \leftarrow \mathsf{Com}_{\mathsf{pk}}(\overline{b})$
$r, b \leftarrow \mathcal{V}_*(\mathsf{pk}, \overline{c})$
if $r^2 b = c$ then return $\mathcal{V}_*(\overline{d})$
else return $\mathcal{V}_*(\perp)$
```

Note that for fixed ϕ, ω pair, the knowledge extractor fails with the same but probability $p_{\mathrm{fail}} = p_{\mathrm{fail}}(\mathcal{V}_*)$.

Hence, we can simplify the distributions:

```
$D_5$
$c \leftarrow \mathcal{V}_*(\phi;\omega)$
$\overline{b} \leftarrow \mathsf{IsQnr}_{p,q}(c)$
$s \xleftarrow[u]{} [0,1]$
if $s \leq p_{\text{fail}}(\mathcal{V}_*)$ then $b_* \xleftarrow[u]{} \neg\overline{b}$
else $b_* \leftarrow \overline{b}$
$\mathsf{pk} \leftarrow \mathsf{Gen}$
$(\hat{c}, \hat{d}) \leftarrow \mathsf{Com}_{\mathsf{pk}}(b_*)$
$r, b \leftarrow \mathcal{V}_*(\mathsf{pk}, \hat{c})$
if $r^2 b = c$ then **return** $\mathcal{V}_*(\hat{d})$
else **return** $\mathcal{V}_*(\perp)$
```

```
$D_6$
$c \leftarrow \mathcal{V}_*(\phi;\omega)$
$\overline{b} \leftarrow \mathsf{IsQnr}_{p,q}(c)$
$s \xleftarrow[u]{} [0,1]$
if $s \leq p_{\text{fail}}(\mathcal{V}_*)$ then $b_* \xleftarrow[u]{} \neg\overline{b}$
else $b_* \leftarrow \overline{b}$
$\mathsf{pk} \leftarrow \mathsf{Gen}$
$(\overline{c}, \overline{d}) \leftarrow \mathsf{Com}_{\mathsf{pk}}(\overline{b})$
$r, b \leftarrow \mathcal{V}_*(\mathsf{pk}, \overline{c})$
if $r^2 b = c$ then **return** $\mathcal{V}_*(\overline{d})$
else **return** $\mathcal{V}_*(\perp)$ .
```

Let $\mathcal{A}$ be a τ-time distinguisher. Then we can decompose its advantage

$$\begin{aligned}\mathsf{Adv}^{\mathsf{ind}}_{D_3,D_4}(\mathcal{A}) \leq \big|&\Pr\left[\psi \leftarrow D_3 : \mathcal{A}(\psi) = 1 \wedge r^2 b \neq c\right] - \Pr\left[\psi \leftarrow D_4 : \mathcal{A}(\psi) = 1 \wedge r^2 b \neq c\right]\big| \\ &+ \big|\Pr\left[\psi \leftarrow D_3 : \mathcal{A}(\psi) = 1 \wedge r^2 b = c\right] - \Pr\left[\psi \leftarrow D_4 : \mathcal{A}(\psi) = 1 \wedge r^2 b = c\right]\big| \enspace .\end{aligned}$$

Let us first analyse the second term Δ_2. Clearly, if $s > p_{\text{fail}}(\mathcal{V}_*)$ then the distributions provide identical outputs and thus the advantage of $\mathcal{A}$ can be expressed by the following games:

```
$\mathcal{G}_0^{\mathcal{A}}$
$c \leftarrow \mathcal{V}_*(\phi;\omega)$
$\overline{b} \leftarrow \mathsf{IsQnr}_{p,q}(c)$
$s \xleftarrow[u]{} [0,1]$
if $s > p_{\text{fail}}(\mathcal{V}_*)$ then **return** $0$
$b_* \xleftarrow[u]{} \neg\overline{b}$
$\mathsf{pk} \leftarrow \mathsf{Gen}$
$(\hat{c}, \hat{d}) \leftarrow \mathsf{Com}_{\mathsf{pk}}(b_*)$
$r, b \leftarrow \mathcal{V}_*(\mathsf{pk}, \hat{c})$
if $r^2 b = c$ then **return** $\mathcal{A}(\mathcal{V}_*(\hat{d}))$
else **return** $0$
```

```
$\mathcal{G}_1^{\mathcal{A}}$
$c \leftarrow \mathcal{V}_*(\phi;\omega)$
$\overline{b} \leftarrow \mathsf{IsQnr}_{p,q}(c)$
$s \xleftarrow[u]{} [0,1]$
if $s > p_{\text{fail}}(\mathcal{V}_*)$ then **return** $0$
$b_* \xleftarrow[u]{} \neg\overline{b}$
$\mathsf{pk} \leftarrow \mathsf{Gen}$
$(\overline{c}, \overline{d}) \leftarrow \mathsf{Com}_{\mathsf{pk}}(\overline{b})$
$r, b \leftarrow \mathcal{V}_*(\mathsf{pk}, \overline{c})$
if $r^2 b = c$ then **return** $\mathcal{A}(\mathcal{V}_*(\overline{d}))$
else **return** $0$ .
```

Now, it is evident that we have to consider only the probability differences in the branches $s > p_{\text{fail}}$ and thus

$$\Delta_2 = p_{\text{fail}} \cdot \left|\Pr\left[\mathcal{G}_2^{\mathcal{A}} = 1\right] - \Pr\left[\mathcal{G}_3^{\mathcal{A}} = 1\right]\right|$$

where

```
$\mathcal{G}_2^{\mathcal{A}}$
$c \leftarrow \mathcal{V}_*(\phi;\omega)$
$\mathsf{pk} \leftarrow \mathsf{Gen}$
$(\hat{c}, \hat{d}) \leftarrow \mathsf{Com}_{\mathsf{pk}}(0)$
$r, b \leftarrow \mathcal{V}_*(\mathsf{pk}, \hat{c})$
if $r^2 b = c$ then **return** $\mathcal{A}(\mathcal{V}_*(\hat{d}))$
else **return** $0$
```

```
$\mathcal{G}_3^{\mathcal{A}}$
$c \leftarrow \mathcal{V}_*(\phi;\omega)$
$\mathsf{pk} \leftarrow \mathsf{Gen}$
$(\hat{c}, \hat{d}) \leftarrow \mathsf{Com}_{\mathsf{pk}}(1)$
$r, b \leftarrow \mathcal{V}_*(\mathsf{pk}, \overline{c})$
if $r^2 b = c$ then **return** $\mathcal{A}(\mathcal{V}_*(\hat{d}))$
else **return** $0$ .
```

From these games it is evident that

$$\left|\Pr\left[\mathcal{G}_2^{\mathcal{A}}=1\right]-\Pr\left[\mathcal{G}_3^{\mathcal{A}}=1\right]\right|\leq\max\left\{\Pr\left[r,b\leftarrow D_7: r^2v^b=c\right],\Pr\left[r,b\leftarrow D_8: r^2v^b=c\right]\right\}$$

where procedure D_7 and D_8 correspond to the first fragments of the game:

```
$D_7$
  $c \leftarrow \mathcal{V}_*(\phi;\omega)$
  $\mathsf{pk} \leftarrow \mathsf{Gen}$
  $(\hat{c},\hat{d}) \leftarrow \mathsf{Com}_{\mathsf{pk}}(0)$
  $r,b \leftarrow \mathcal{V}_*(\mathsf{pk},\hat{c})$
  return $(r,b)$
```

```
$D_8$
  $c \leftarrow \mathcal{V}_*(\phi;\omega)$
  $\mathsf{pk} \leftarrow \mathsf{Gen}$
  $(\hat{c},\hat{d}) \leftarrow \mathsf{Com}_{\mathsf{pk}}(1)$
  $r,b \leftarrow \mathcal{V}_*(\mathsf{pk},\hat{c})$
  return $(r,b)$ .
```

If the running time t_v of $\mathcal{V}_*$ together with the running time τ of $\mathcal{A}$ is below t, then the advantages corresponding to these games can differ only by $\varepsilon_\circ$. As one of these runs corresponds to the same fragment is the real world execution and has thus probability ε, we have established that

$$\left|\Pr\left[\mathcal{G}_2^{\mathcal{A}}=1\right]-\Pr\left[\mathcal{G}_3^{\mathcal{A}}=1\right]\right|\leq\varepsilon+\varepsilon_\circ\enspace.$$

This implies

$$\mathsf{Adv}^{\mathsf{ind}}_{D_3,D_4}(\mathcal{A})\leq\left|\Pr\left[\psi\leftarrow D_3:\mathcal{A}(\psi)=1\wedge r^2b\neq c\right]-\Pr\left[\psi\leftarrow D_4:\mathcal{A}(\psi)=1\wedge r^2b\neq c\right]\right|+p_{\mathrm{fail}}(\mathcal{V}_*)\cdot(\varepsilon+\varepsilon_\circ)\enspace.$$

To estimate the remaining term, we can consider the following games

```
$\mathcal{G}_5^{\mathcal{A}}$
  $c \leftarrow \mathcal{V}_*(\phi;\omega)$
  $\overline{b} \leftarrow \mathsf{IsQnr}_{p,q}(c)$
  $s \xleftarrow[u]{} [0,1]$
  if $s > p_{\mathrm{fail}}(\mathcal{V}_*)$ then return $0$
  $b_* \xleftarrow[u]{} \neg\overline{b}$
  $\mathsf{pk} \leftarrow \mathsf{Gen}$
  $(\hat{c},\hat{d}) \leftarrow \mathsf{Com}_{\mathsf{pk}}(b_*)$
  $r,b \leftarrow \mathcal{V}_*(\mathsf{pk},\hat{c})$
  if $r^2b = c$ then return $0$
  else return $\mathcal{A}(\mathcal{V}_*(\perp))$
```

```
$\mathcal{G}_6^{\mathcal{A}}$
  $c \leftarrow \mathcal{V}_*(\phi;\omega)$
  $\overline{b} \leftarrow \mathsf{IsQnr}_{p,q}(c)$
  $s \xleftarrow[u]{} [0,1]$
  if $s > p_{\mathrm{fail}}(\mathcal{V}_*)$ then return $0$
  $b_* \xleftarrow[u]{} \neg\overline{b}$
  $\mathsf{pk} \leftarrow \mathsf{Gen}$
  $(\overline{c},\overline{d}) \leftarrow \mathsf{Com}_{\mathsf{pk}}(\overline{b})$
  $r,b \leftarrow \mathcal{V}_*(\mathsf{pk},\overline{c})$
  if $r^2b = c$ then return $0$
  else return $\mathcal{A}(\mathcal{V}_*(\perp))$ .
```

Clearly the difference can only grow if we omit the first check and consider games

```
$\mathcal{G}_7^{\mathcal{A}}$
  $c \leftarrow \mathcal{V}_*(\phi;\omega)$
  $\mathsf{pk} \leftarrow \mathsf{Gen}$
  $(\hat{c},\hat{d}) \leftarrow \mathsf{Com}_{\mathsf{pk}}(0)$
  $r,b \leftarrow \mathcal{V}_*(\mathsf{pk},\hat{c})$
  if $r^2b = c$ then return $0$
  else return $\mathcal{A}(\mathcal{V}_*(\perp))$
```

```
$\mathcal{G}_8^{\mathcal{A}}$
  $c \leftarrow \mathcal{V}_*(\phi;\omega)$
  $\mathsf{pk} \leftarrow \mathsf{Gen}$
  $(\hat{c},\hat{d}) \leftarrow \mathsf{Com}_{\mathsf{pk}}(0)$
  $r,b \leftarrow \mathcal{V}_*(\mathsf{pk},\hat{c})$
  if $r^2b = c$ then return $0$
  else return $\mathcal{A}(\mathcal{V}_*(\perp))$ .
```

As the decommitment value is never released, we can reduce these games to hiding game. Thus if the sum of running times $t_v+\tau\leq t$, then

$$\left|\Pr\left[\mathcal{G}_7^{\mathcal{A}}=1\right]-\Pr\left[\mathcal{G}_8^{\mathcal{A}}=1\right]\right|\leq\varepsilon_\circ\enspace,$$

which leads to the following final estimate:

$$\mathsf{Adv}^{\mathsf{ind}}_{D_3,D_4}(\mathcal{A}) \leq \varepsilon_\circ + p_{\mathrm{fail}}(\mathcal{V}_*) \cdot (\varepsilon + \varepsilon_\circ) \ \ .$$

Final remarks. Let us now compare the results for perfectly hiding and computationally hiding commitments. Let ψ denote the outdot distribution for the real and $\psi_\circ$ the output distribution for the simulation. Then we can express the bounds on the distinguishing advantage as follows:

$$\begin{aligned}\mathsf{sd}(\psi,\psi_\circ) &\leq \max_{\varepsilon\in[0,1]} \varepsilon(1-\varepsilon)^\ell \\ \mathsf{cd}^\tau(\psi,\psi_\circ) &\leq \varepsilon_\circ + \max_{\varepsilon\in[0,1]} (\varepsilon+\varepsilon_\circ)(1+\varepsilon_\circ-\varepsilon)^\ell \ \ .\end{aligned}$$

The form of the inequalities indicates that the value $\varepsilon_\circ$ influences the bound in two ways. First it offsets the final bound by the $\varepsilon_\circ$. Second, it offsets the simulation error term. This effect is bounded. One can clearly see that it can offset the second term only by $\varepsilon_\circ$ at most.

Another thing that is worth mentioning is that our analysis excluded the complexity of the knowledge extractor. Hence, by increasing ℓ we can suppress the second term regardless of the bound t. Of course, this increases the overhead of simulator and thus is unwanted but at least we do not have to choose the commitment scheme parameters on planned overhead.