A Computational Security Analysis of Signal's PQXDH Handshake

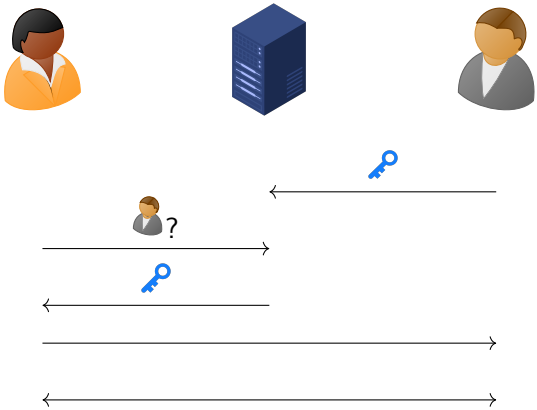
Rune Fiedler[1] Felix Günther[2]

[1]Technische Universität Darmstadt, Germany
rune.fiedler@cryptoplexity.de

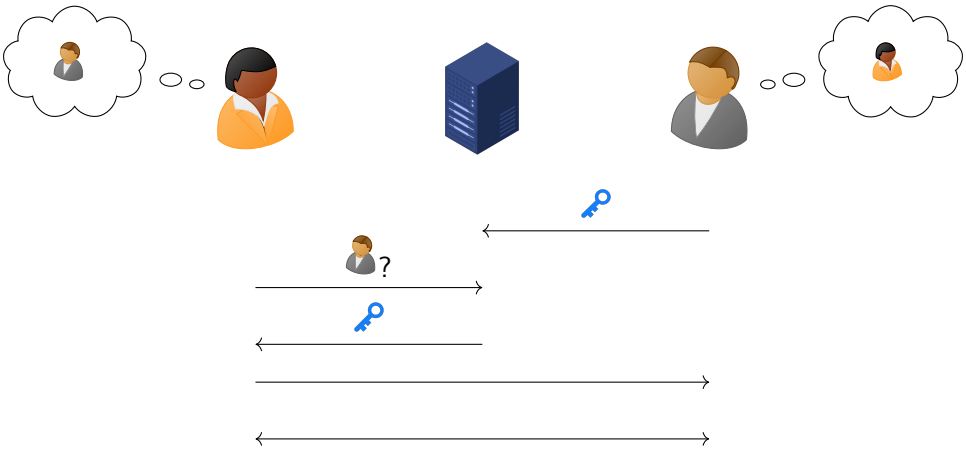
[2]IBM Research Europe – Zurich, Switzerland
mail@felixguenther.info

CAW 2024

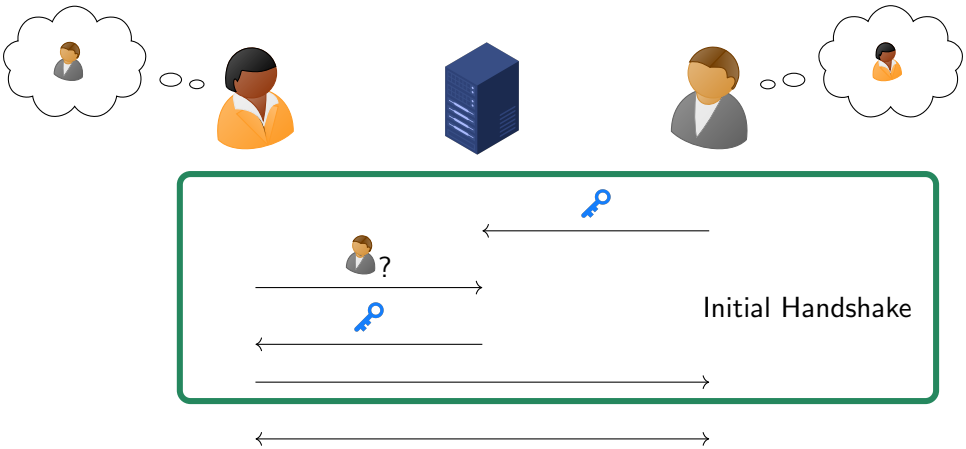
Asynchronous Authenticated Key Exchange



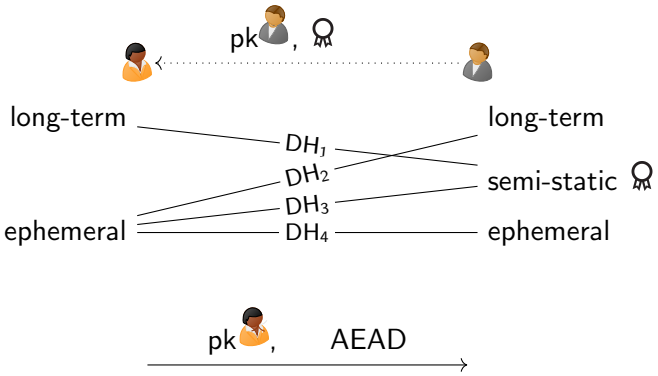
Asynchronous Authenticated Key Exchange



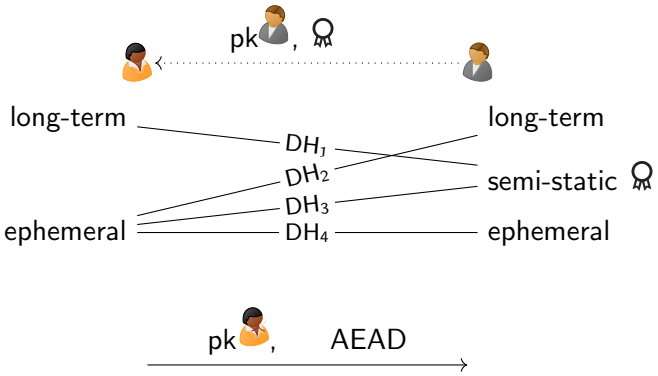
Asynchronous Authenticated Key Exchange



Signal's Initial Handshake(s): X3DH and PQXDH

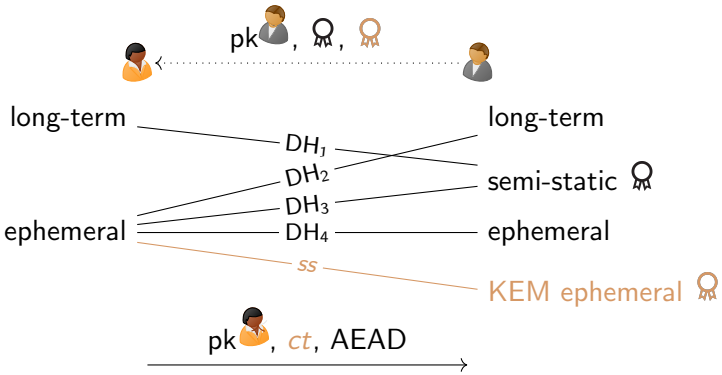


Signal's Initial Handshake(s): X3DH and PQXDH



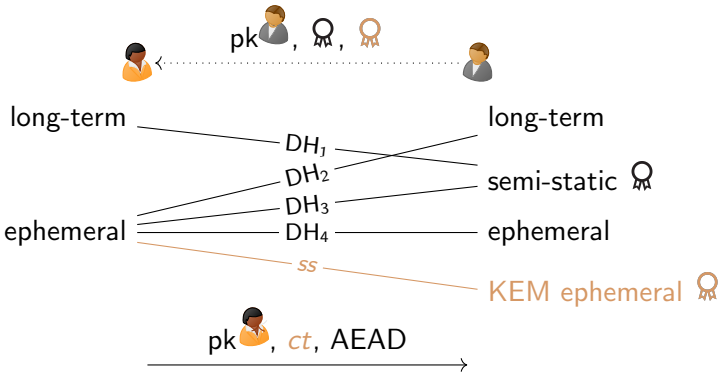
- session key: $\mathrm{KDF}(\mathrm{DH}_1 \| \ldots \| \mathrm{DH}_4 \quad)$

Signal's Initial Handshake(s): X3DH and PQXDH



- session key: $\mathsf{KDF}(\mathsf{DH}_1 \| \ldots \| \mathsf{DH}_4 \| ss)$

Signal's Initial Handshake(s): X3DH and PQXDH



- session key: $\mathrm{KDF}(\mathrm{DH}_1 \| \ldots \| \mathrm{DH}_4 \| ss)$
- reduced session: Bob without ephemeral keys, semi-static KEM

Analyzing Signal's Initial Handshake(s): X3DH and PQXDH

- reductionist analysis of X3DH [CCD+17] with a [BR94] style key exchange model
- tool-based analysis of PQXDH with ProVerif and CryptoVerif [BJK23, BJKS23]
  - (re-)discovered (potential) KEM re-encapsulation attack [CDM23]
  - corruption of long-term keys only

[BJK23, BJKS23] Barghavan, Jacomme, Kiefer, Schmidt, 2023
[BR94] Bellare, Rogaway, CRYPTO 1993
[BFG+22] Brendel, Fiedler, Günther, Janson, Stebila, PKC 2022
[CCD+17] Cohn-Gordon, Cremers, Dowling, Garratt, Stebila, EuroSP 2017
[CDM23] Cremers, Dax, Medinger, ePrint 2023
[FG24] Fiedler, Günther, ePrint 2024

Analyzing Signal's Initial Handshake(s): X3DH and PQXDH

- reductionist analysis of X3DH [CCD+17] with a [BR94] style key exchange model
- tool-based analysis of PQXDH with ProVerif and CryptoVerif [BJK23, BJKS23]
  - (re-)discovered (potential) KEM re-encapsulation attack [CDM23]
  - corruption of long-term keys only
- our work: [FG24]
  - follows [CCD+17, BFG+22] but explicitly models signatures (albeit with distinct signing keys)
  - identifies precise requirements of the KEM
  - models maximum-exposure w. clean predicates: $\underbrace{\underbrace{\underbrace{\mathsf{clean}_{\mathsf{LTSS}}, \mathsf{clean}_{\mathsf{ELT}}, \mathsf{clean}_{\mathsf{ESS}}}_{\text{reduced}}, \mathsf{clean}_{\mathsf{EE}}}_{\text{X3DH full}}, \mathsf{clean}_{\mathsf{sigE}}}_{\text{PQXDH full}}$

[BJK23, BJKS23] Barghavan, Jacomme, Kiefer, Schmidt, 2023
[BR94] Bellare, Rogaway, CRYPTO 1993
[BFG+22] Brendel, Fiedler, Günther, Janson, Stebila, PKC 2022
[CCD+17] Cohn-Gordon, Cremers, Dowling, Garratt, Stebila, EuroSP 2017
[CDM23] Cremers, Dax, Medinger, ePrint 2023
[FG24] Fiedler, Günther, ePrint 2024

Concrete Bound for PQXDH

$$\mathsf{Adv}^{\mathsf{KI}}_{\mathsf{PQXDH}}(\mathcal{A}) \leq \frac{(n_p + n_p \cdot n_{ss} + n_s)^2}{q} + \gamma_{\mathsf{coll}}(n_p \cdot n_{ss} + n_s) + n_s \cdot \delta_{\mathsf{corr}} + \epsilon_{LEAK^+}$$

$$+ \left(\begin{array}{ll} \quad (n_p \cdot (\epsilon_{\mathsf{SIG}} + n_p \cdot n_{ss} \cdot \epsilon_{\mathsf{GDH}})) & // \, \mathsf{clean}_{\mathsf{LTSS}} \\ + (n_s \cdot n_p \cdot \epsilon_{\mathsf{GDH}}) & // \, \mathsf{clean}_{\mathsf{ELT}} \\ + (n_p \cdot (\epsilon_{\mathsf{SIG}} + n_{ss} \cdot n_s \cdot \epsilon_{\mathsf{GDH}})) & // \, \mathsf{clean}_{\mathsf{ESS}} \wedge \mathsf{type} = \mathsf{full} \\ + (n_p \cdot (\epsilon_{\mathsf{SIG}} + n_{ss} \cdot n_s \cdot \min(\epsilon_{\mathsf{GDH}}, \epsilon_{\mathsf{CCA}}))) & // \, \mathsf{clean}_{\mathsf{ESS}} \wedge \mathsf{type} = \mathsf{reduced} \\ + (n_s^2 \cdot \min(\epsilon_{\mathsf{GDH}}, \epsilon_{\mathsf{CCA}})) & // \, \mathsf{clean}_{\mathsf{EE}} \wedge \mathsf{clean}_{\mathsf{peerE}} \\ + (n_p \cdot (\epsilon_{\mathsf{SIG}} + n_s^2 \cdot q_{\mathrm{RO}} \cdot \epsilon_{\mathsf{CCA}})) & // \, \mathsf{clean}_{\mathsf{EE}} \wedge \mathsf{clean}_{\mathsf{sigE}} \end{array} \right)$$

Concrete Bound for PQXDH

$$\mathsf{Adv}^{\mathsf{KI}}_{\mathsf{PQXDH}}(\mathcal{A}) \leq \frac{(n_p + n_p \cdot n_{ss} + n_s)^2}{q} + \gamma_{\mathsf{coll}}(n_p \cdot n_{ss} + n_s) + n_s \cdot \delta_{\mathsf{corr}} + \epsilon_{LEAK^+}$$

$$+ \left(\begin{array}{ll} (n_p \cdot (\epsilon_{\mathsf{SIG}} + n_p \cdot n_{ss} \cdot \epsilon_{\mathsf{GDH}})) & // \ \mathsf{clean}_{\mathsf{LTSS}} \\ + (n_s \cdot n_p \cdot \epsilon_{\mathsf{GDH}}) & // \ \mathsf{clean}_{\mathsf{ELT}} \\ + (n_p \cdot (\epsilon_{\mathsf{SIG}} + n_{ss} \cdot n_s \cdot \epsilon_{\mathsf{GDH}})) & // \ \mathsf{clean}_{\mathsf{ESS}} \wedge \mathsf{type} = \mathsf{full} \\ + (n_p \cdot (\epsilon_{\mathsf{SIG}} + n_{ss} \cdot n_s \cdot \min(\epsilon_{\mathsf{GDH}}, \epsilon_{\mathsf{CCA}}))) & // \ \mathsf{clean}_{\mathsf{ESS}} \wedge \mathsf{type} = \mathsf{reduced} \\ + (n_s^2 \cdot \min(\epsilon_{\mathsf{GDH}}, \epsilon_{\mathsf{CCA}})) & // \ \mathsf{clean}_{\mathsf{EE}} \wedge \mathsf{clean}_{\mathsf{peerE}} \\ + (n_p \cdot (\epsilon_{\mathsf{SIG}} + n_s^2 \cdot q_{\mathrm{RO}} \cdot \epsilon_{\mathsf{CCA}})) & // \ \mathsf{clean}_{\mathsf{EE}} \wedge \mathsf{clean}_{\mathsf{sigE}} \end{array} \right)$$

Concrete Bound for PQXDH

$$\mathsf{Adv}^{\mathsf{KI}}_{\mathsf{PQXDH}}(\mathcal{A}) \leq \frac{(n_p + n_p \cdot n_{ss} + n_s)^2}{q} + \gamma_{\mathsf{coll}}(n_p \cdot n_{ss} + n_s) + n_s \cdot \delta_{\mathsf{corr}} + \epsilon_{LEAK^+}$$

$$+ \left(\begin{array}{ll} (n_p \cdot (\epsilon_{\mathsf{SIG}} + n_p \cdot n_{ss} \cdot \epsilon_{\mathsf{GDH}})) & // \, \mathsf{clean}_{\mathsf{LTSS}} \\ + (n_s \cdot n_p \cdot \epsilon_{\mathsf{GDH}}) & // \, \mathsf{clean}_{\mathsf{ELT}} \\ + (n_p \cdot (\epsilon_{\mathsf{SIG}} + n_{ss} \cdot n_s \cdot \epsilon_{\mathsf{GDH}})) & // \, \mathsf{clean}_{\mathsf{ESS}} \wedge \mathsf{type} = \mathsf{full} \\ + (n_p \cdot (\epsilon_{\mathsf{SIG}} + n_{ss} \cdot n_s \cdot \min(\epsilon_{\mathsf{GDH}}, \epsilon_{\mathsf{CCA}}))) & // \, \mathsf{clean}_{\mathsf{ESS}} \wedge \mathsf{type} = \mathsf{reduced} \\ + (n_s^2 \cdot \min(\epsilon_{\mathsf{GDH}}, \epsilon_{\mathsf{CCA}})) & // \, \mathsf{clean}_{\mathsf{EE}} \wedge \mathsf{clean}_{\mathsf{peerE}} \\ + (n_p \cdot (\epsilon_{\mathsf{SIG}} + n_s^2 \cdot q_{\mathrm{RO}} \cdot \epsilon_{\mathsf{CCA}})) & // \, \mathsf{clean}_{\mathsf{EE}} \wedge \mathsf{clean}_{\mathsf{sigE}} \end{array} \right)$$

Concrete Bound for PQXDH

$$\mathsf{Adv}^{\mathsf{KI}}_{\mathsf{PQXDH}}(\mathcal{A}) \leq \frac{(n_p + n_p \cdot n_{ss} + n_s)^2}{q} + \gamma_{\mathsf{coll}}(n_p \cdot n_{ss} + n_s) + n_s \cdot \delta_{\mathsf{corr}} + \epsilon_{LEAK^+}$$

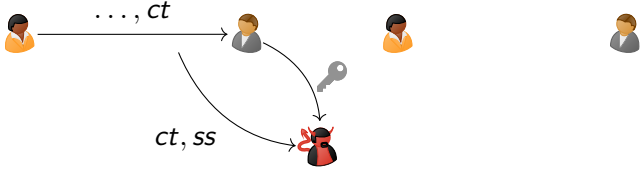
$$+ \left(\begin{array}{ll} (n_p \cdot (\epsilon_{\mathsf{SIG}} + n_p \cdot n_{ss} \cdot \epsilon_{\mathsf{GDH}})) & // \mathsf{clean}_{\mathsf{LTSS}} \\ + (n_s \cdot n_p \cdot \epsilon_{\mathsf{GDH}}) & // \mathsf{clean}_{\mathsf{ELT}} \\ + (n_p \cdot (\epsilon_{\mathsf{SIG}} + n_{ss} \cdot n_s \cdot \epsilon_{\mathsf{GDH}})) & // \mathsf{clean}_{\mathsf{ESS}} \wedge \mathsf{type} = \mathsf{full} \\ + (n_p \cdot (\epsilon_{\mathsf{SIG}} + n_{ss} \cdot n_s \cdot \min(\epsilon_{\mathsf{GDH}}, \epsilon_{\mathsf{CCA}}))) & // \mathsf{clean}_{\mathsf{ESS}} \wedge \mathsf{type} = \mathsf{reduced} \\ + (n_s^2 \cdot \min(\epsilon_{\mathsf{GDH}}, \epsilon_{\mathsf{CCA}})) & // \mathsf{clean}_{\mathsf{EE}} \wedge \mathsf{clean}_{\mathsf{peerE}} \\ + (n_p \cdot (\epsilon_{\mathsf{SIG}} + n_s^2 \cdot q_{\mathrm{RO}} \cdot \epsilon_{\mathsf{CCA}})) & // \mathsf{clean}_{\mathsf{EE}} \wedge \mathsf{clean}_{\mathsf{sigE}} \end{array} \right)$$

KEM Re-Encapsulation Attack [CDM23, BJK23, BJKS23]

- two sessions with same DH public keys, distinct KEM keys, both reduced

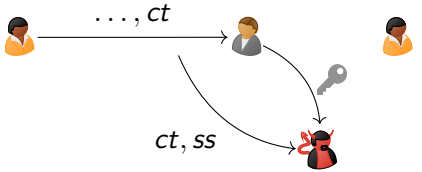
KEM Re-Encapsulation Attack [CDM23, BJK23, BJKS23]

- two sessions with same DH public keys, distinct KEM keys, both reduced



KEM Re-Encapsulation Attack [CDM23, BJK23, BJKS23]

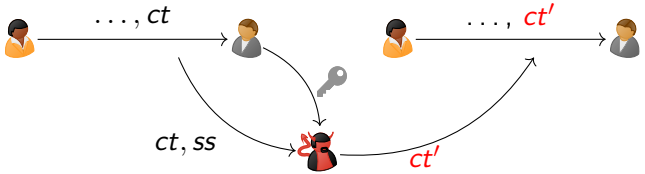
- two sessions with same DH public keys, distinct KEM keys, both reduced



$\underline{KEM.encaps(\mathrm{pk}) :}$
$ss \leftarrow \{0,1\}^{256}$
$ct \leftarrow PKE.encrypt(\mathrm{pk}, ss)$
return (ct, ss)

KEM Re-Encapsulation Attack [CDM23, BJK23, BJKS23]

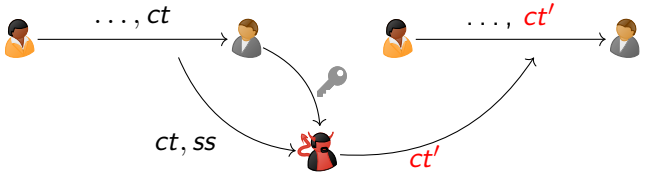
- two sessions with same DH public keys, distinct KEM keys, both reduced



$\underline{KEM.encaps(\mathrm{pk}) :}$
$ss \leftarrow \{0,1\}^{256}$
$ct \leftarrow PKE.encrypt(\mathrm{pk}, ss)$
return (ct, ss)

KEM Re-Encapsulation Attack [CDM23, BJK23, BJKS23]

- two sessions with same DH public keys, distinct KEM keys, both reduced

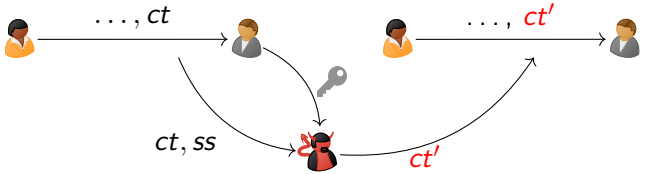


$\underline{KEM.encaps(\mathrm{pk}) :}$
$ss \leftarrow \{0,1\}^{256}$
$ct \leftarrow PKE.encrypt(\mathrm{pk}, ss)$
return (ct, ss)

session key: $\mathrm{KDF}(\mathrm{DH}_1 \| \ldots \| \mathrm{DH}_4 \| ss)$

KEM Re-Encapsulation Attack [CDM23, BJK23, BJKS23]

- two sessions with same DH public keys, distinct KEM keys, both reduced



$\underline{KEM.encaps(\mathrm{pk}) :}$
$ss \leftarrow \{0,1\}^{256}$
$ct \leftarrow PKE.encrypt(\mathrm{pk}, ss)$
return (ct, ss)

session key: $\mathrm{KDF}(\mathrm{DH}_1 \| \ldots \| \mathrm{DH}_4 \| ss)$

- $\Rightarrow$ two sessions with same session key: adversary can reveal one and test the other
- [BJK23, BJKS23] models the KEM public key into the Associated Data of the AEAD
- which KEM property needed?
- proposed protocol fix: session context (KEM public key, ciphertext) in key derivation

KEM Binding Notion $LEAK^{+}$-BIND-SS-$\{CT, PK\}$ (extending [CDM23])

$$((\mathsf{pk}, \mathsf{sk}, r)_1, \ldots, (\mathsf{pk}, \mathsf{sk}, r)_n)$$

$$(\mathsf{pk}_i, ct_i) \neq (\mathsf{pk}_j, ct_j)$$

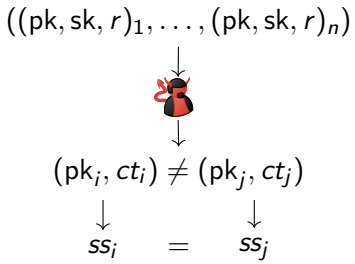
KEM Binding Notion $LEAK^{+}$-BIND-SS-$\{CT, PK\}$ (extending [CDM23])

$$((\mathsf{pk}, \mathsf{sk}, r)_1, \ldots, (\mathsf{pk}, \mathsf{sk}, r)_n)$$

$$(\mathsf{pk}_i, ct_i) \neq (\mathsf{pk}_j, ct_j)$$

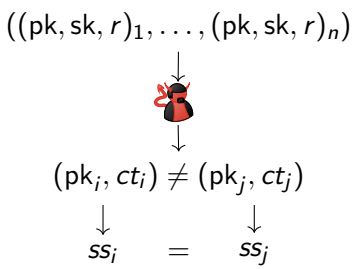
$$ss_i \quad = \quad ss_j$$

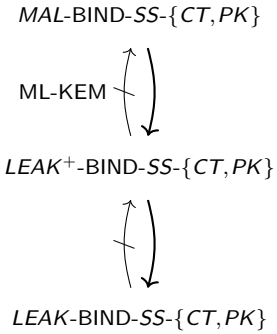
KEM Binding Notion $LEAK^{+}$-BIND-SS-$\{CT, PK\}$ (extending [CDM23])



$LEAK^{+}$-BIND-SS-$\{CT, PK\}$

KEM Binding Notion $LEAK^{+}$-BIND-SS-$\{CT, PK\}$ (extending [CDM23])

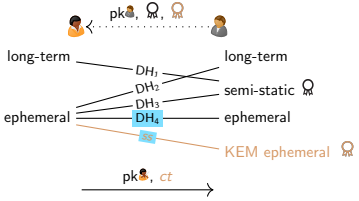




Concrete Hybrid Bound for PQXDH

$$\mathsf{Adv}^{\mathsf{KI}}_{\mathsf{PQXDH}}(\mathcal{A}) \leq \frac{(n_p + n_p \cdot n_{ss} + n_s)^2}{q} + \gamma_{\mathsf{coll}}(n_p \cdot n_{ss} + n_s) + n_s \cdot \delta_{\mathsf{corr}} + \epsilon_{LEAK^+}$$

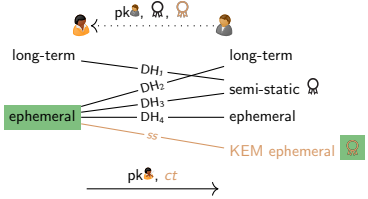
$$+ \left(\begin{array}{ll} (n_p \cdot (\epsilon_{\mathsf{SIG}} + n_p \cdot n_{ss} \cdot \epsilon_{\mathsf{GDH}})) & // \, \mathsf{clean}_{\mathsf{LTSS}} \\ + (n_s \cdot n_p \cdot \epsilon_{\mathsf{GDH}}) & // \, \mathsf{clean}_{\mathsf{ELT}} \\ + (n_p \cdot (\epsilon_{\mathsf{SIG}} + n_{ss} \cdot n_s \cdot \epsilon_{\mathsf{GDH}})) & // \, \mathsf{clean}_{\mathsf{ESS}} \wedge \mathsf{type} = \mathsf{full} \\ + (n_p \cdot (\epsilon_{\mathsf{SIG}} + n_{ss} \cdot n_s \cdot \min(\epsilon_{\mathsf{GDH}}, \epsilon_{\mathsf{CCA}}))) & // \, \mathsf{clean}_{\mathsf{ESS}} \wedge \mathsf{type} = \mathsf{reduced} \\ + (n_s^2 \cdot \min(\epsilon_{\mathsf{GDH}}, \epsilon_{\mathsf{CCA}})) & // \, \mathsf{clean}_{\mathsf{EE}} \wedge \mathsf{clean}_{\mathsf{peerE}} \\ + (n_p \cdot (\epsilon_{\mathsf{SIG}} + n_s^2 \cdot q_{\mathrm{RO}} \cdot \epsilon_{\mathsf{CCA}})) & // \, \mathsf{clean}_{\mathsf{EE}} \wedge \mathsf{clean}_{\mathsf{sigE}} \end{array} \right)$$



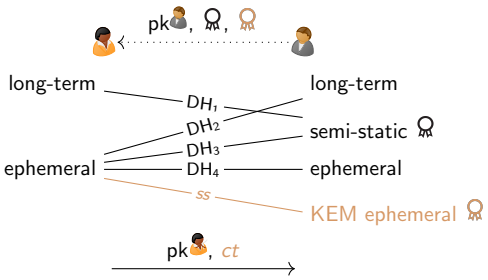
Concrete Bound for PQXDH Against Active-Later-Quantum Adversaries

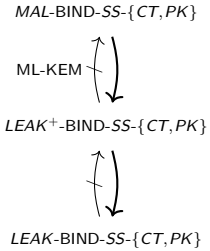
$$\mathsf{Adv}^{\mathsf{KI}}_{\mathsf{PQXDH}}(\mathcal{A}) \leq \frac{(n_p + n_p \cdot n_{ss} + n_s)^2}{q} + \gamma_{\mathsf{coll}}(n_p \cdot n_{ss} + n_s) + n_s \cdot \delta_{\mathsf{corr}} + \epsilon_{LEAK^+}$$

$$+ \left(\begin{array}{ll} (n_p \cdot (\epsilon_{\mathsf{SIG}} + n_p \cdot n_{ss} \cdot \epsilon_{\mathsf{GDH}})) & // \ \mathsf{clean}_{\mathsf{LTSS}} \\ + (n_s \cdot n_p \cdot \epsilon_{\mathsf{GDH}}) & // \ \mathsf{clean}_{\mathsf{ELT}} \\ + (n_p \cdot (\epsilon_{\mathsf{SIG}} + n_{ss} \cdot n_s \cdot \epsilon_{\mathsf{GDH}})) & // \ \mathsf{clean}_{\mathsf{ESS}} \wedge \mathsf{type} = \mathsf{full} \\ + (n_p \cdot (\epsilon_{\mathsf{SIG}} + n_{ss} \cdot n_s \cdot \min(\epsilon_{\mathsf{GDH}}, \epsilon_{\mathsf{CCA}}))) & // \ \mathsf{clean}_{\mathsf{ESS}} \wedge \mathsf{type} = \mathsf{reduced} \\ + (n_s^2 \cdot \min(\epsilon_{\mathsf{GDH}}, \epsilon_{\mathsf{CCA}})) & // \ \mathsf{clean}_{\mathsf{EE}} \wedge \mathsf{clean}_{\mathsf{peerE}} \\ + (n_p \cdot (\epsilon_{\mathsf{SIG}} + n_s^2 \cdot q_{\mathsf{RO}} \cdot \epsilon_{\mathsf{CCA}})) & // \ \mathsf{clean}_{\mathsf{EE}} \wedge \mathsf{clean}_{\mathsf{sigE}} \end{array} \right)$$



PQXDH Provides Hybrid Security





or include the session context in key derivation

$$\mathsf{Adv}^{\mathsf{KI}}_{\mathsf{PQXDH}}(\mathcal{A}) \leq \frac{(n_p + n_p \cdot n_{ss} + n_s)^2}{q} + \gamma_{\mathsf{coll}}(n_p \cdot n_{ss} + n_s) + n_s \cdot \delta_{\mathsf{corr}} + \epsilon_{\mathit{LEAK}^+}$$

$$+ \left(\begin{array}{ll} (n_p \cdot (\epsilon_{\mathsf{SIG}} + n_p \cdot n_{ss} \cdot \epsilon_{\mathsf{GDH}})) & // \ \mathsf{clean}_{\mathsf{LTSS}} \\ + (n_s \cdot n_p \cdot \epsilon_{\mathsf{GDH}}) & // \ \mathsf{clean}_{\mathsf{ELT}} \\ + (n_p \cdot (\epsilon_{\mathsf{SIG}} + n_{ss} \cdot n_s \cdot \epsilon_{\mathsf{GDH}})) & // \ \mathsf{clean}_{\mathsf{ESS}} \wedge \mathsf{type} = \mathsf{full} \\ + (n_p \cdot (\epsilon_{\mathsf{SIG}} + n_{ss} \cdot n_s \cdot \min(\epsilon_{\mathsf{GDH}}, \epsilon_{\mathsf{CCA}}))) & // \ \mathsf{clean}_{\mathsf{ESS}} \wedge \mathsf{type} = \mathsf{reduced} \\ + (n_s^2 \cdot \min(\epsilon_{\mathsf{GDH}}, \epsilon_{\mathsf{CCA}})) & // \ \mathsf{clean}_{\mathsf{EE}} \wedge \mathsf{clean}_{\mathsf{peerE}} \\ + (n_p \cdot (\epsilon_{\mathsf{SIG}} + n_s^2 \cdot q_{\mathsf{RO}} \cdot \epsilon_{\mathsf{CCA}})) & // \ \mathsf{clean}_{\mathsf{EE}} \wedge \mathsf{clean}_{\mathsf{sigE}} \end{array} \right)$$

combine tool-based and reductionist analyses
⇒ detect attacks and identify requirements

`https://eprint.iacr.org/2024/702`

rune.fiedler@cryptoplexity.de

References I

[BFG+22] Jacqueline Brendel, Rune Fiedler, Felix Günther, Christian Janson, and Douglas Stebila.
Post-quantum asynchronous deniable key exchange and the Signal handshake.
In Goichiro Hanaoka, Junji Shikata, and Yohei Watanabe, editors, *PKC 2022: 25th International Conference on Theory and Practice of Public Key Cryptography, Part II*, volume 13178 of *Lecture Notes in Computer Science*, pages 3–34, Virtual Event, March 8–11, 2022. Springer, Heidelberg, Germany.

[BJK23] Karthikean Barghavan, Charlie Jacomme, and Franziskus Kiefer.
Formal analysis of the PQXDH protocol, 2023.
https://github.com/Inria-Prosecco/pqxdh-analysis.

[BJKS23] Karthikean Barghavan, Charlie Jacomme, Franziskus Kiefer, and Rolfe Schmidt.
An analysis of Signal's PQXDH, October 2023.
https://cryspen.com/post/pqxdh/.

[BR94] Mihir Bellare and Phillip Rogaway.
Entity authentication and key distribution.
In Douglas R. Stinson, editor, *Advances in Cryptology – CRYPTO'93*, volume 773 of *Lecture Notes in Computer Science*, pages 232–249, Santa Barbara, CA, USA, August 22–26, 1994. Springer, Heidelberg, Germany.

References II

[CCD+17] Katriel Cohn-Gordon, Cas J. F. Cremers, Benjamin Dowling, Luke Garratt, and Douglas Stebila.
A formal security analysis of the Signal messaging protocol.
In *IEEE European Symposium on Security and Privacy, EuroS&P 2017*, pages 451–466, 2017.

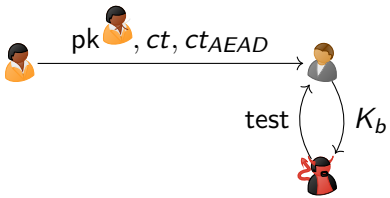
[CDM23] Cas Cremers, Alexander Dax, and Niklas Medinger.
Keeping up with the KEMs: Stronger security notions for KEMs.
Cryptology ePrint Archive, Paper 2023/1933, 2023.
Version 1.0.6 (April 3, 2024), https://eprint.iacr.org/2023/1933.

[FG24] Rune Fiedler and Felix Günther.
Security analysis of Signal's PQXDH handshake.
Cryptology ePrint Archive, Paper 2024/702, 2024.
https://eprint.iacr.org/2024/702.

Picture References

- server icon by Alexiuz AS
- public key icon by Yannick Lung
- secret key icon by Yannick Lung
- signature icon by PINPOINT.WORLD

What if we model the AEAD ciphertext?



if $AEAD.Dec(K_b, ct_{AEAD}) \neq \perp$: **return** 0
else return 1

- $\Rightarrow$ adversary trivially wins (except with negligible probability)