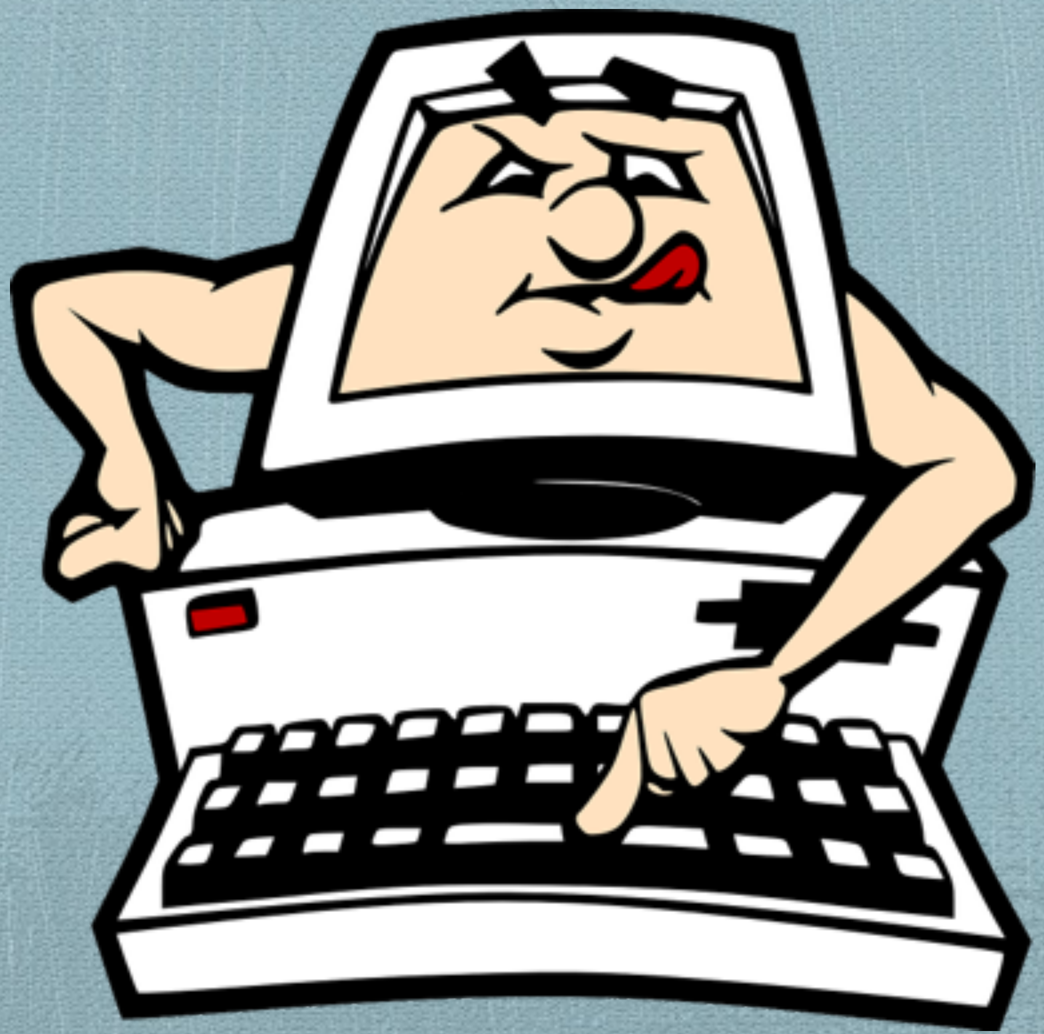
Black-box Separations for Differentially Private Protocols

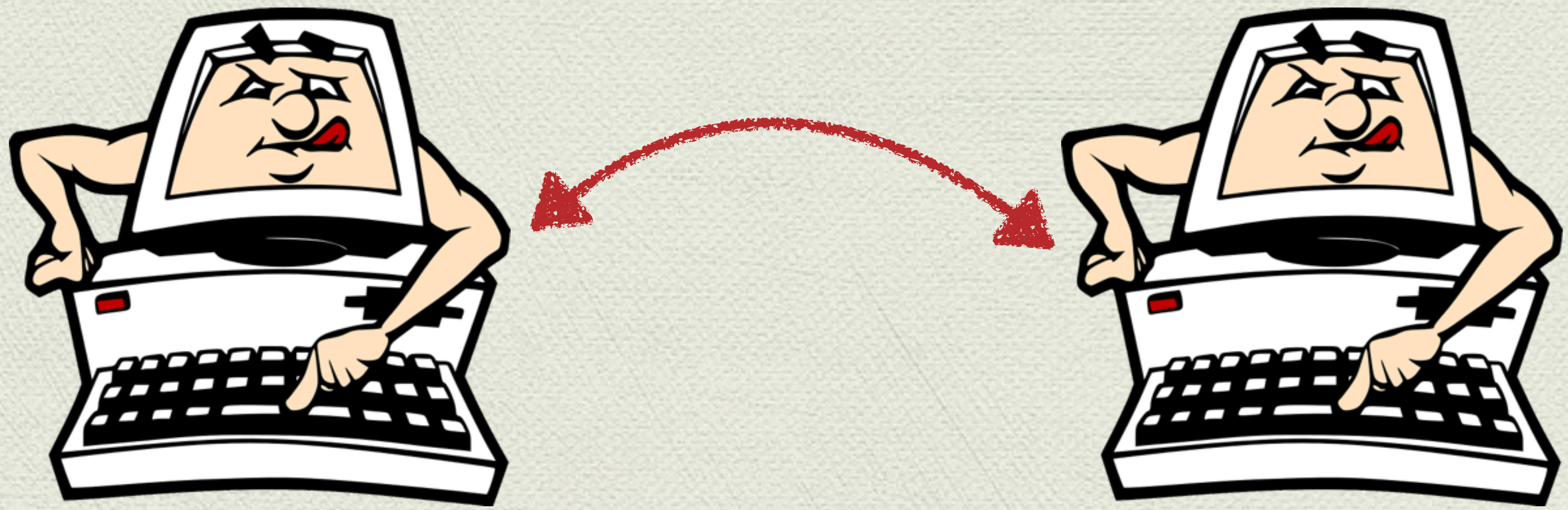
Dakshita Khurana, Hemanta K. Maji, Amit Sahai



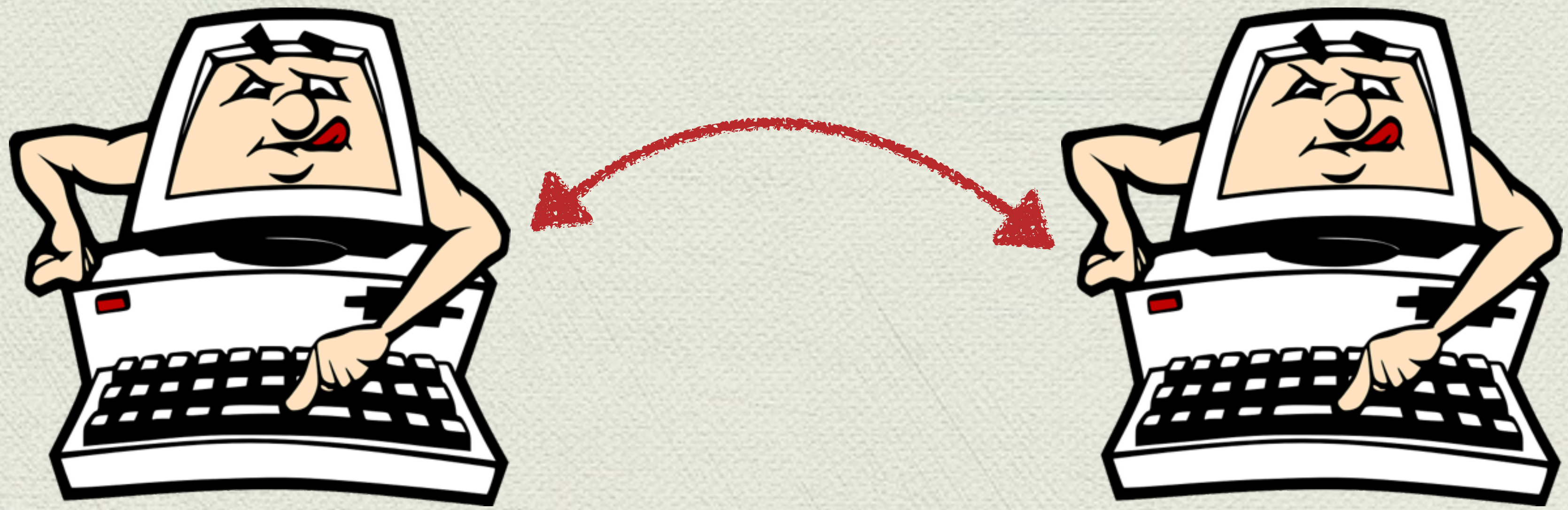
The Setting

The Setting

The Setting

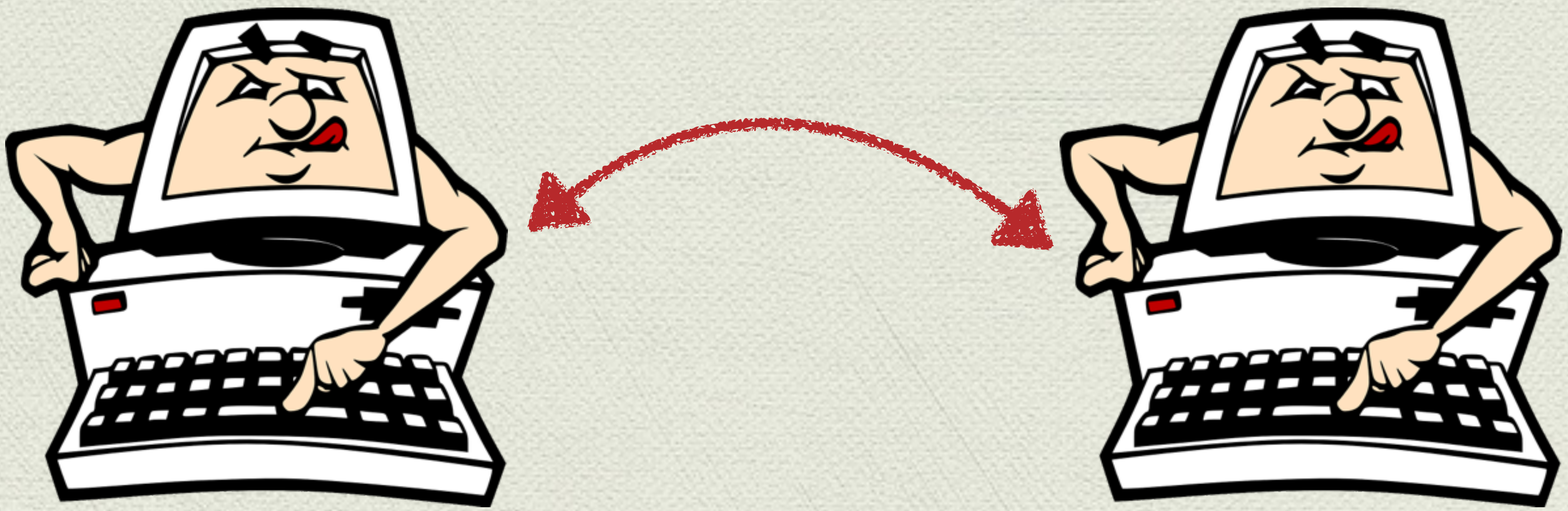


The Setting



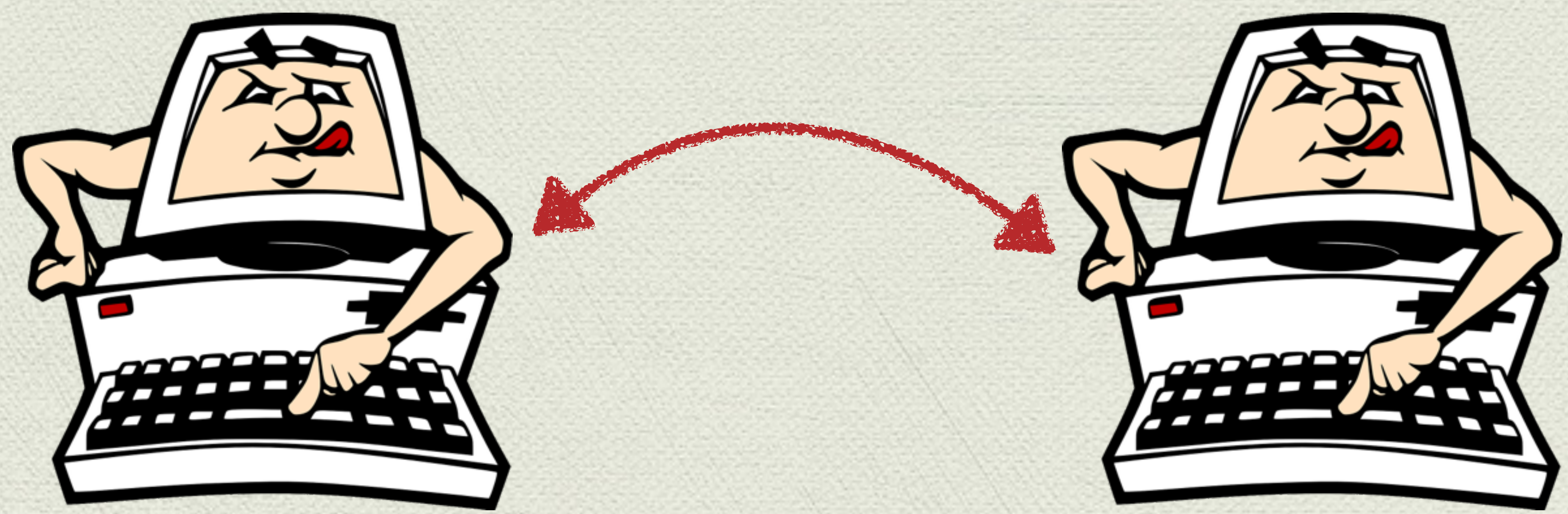
- Mutually distrusting medical servers with databases (d_1, d_2)

The Setting



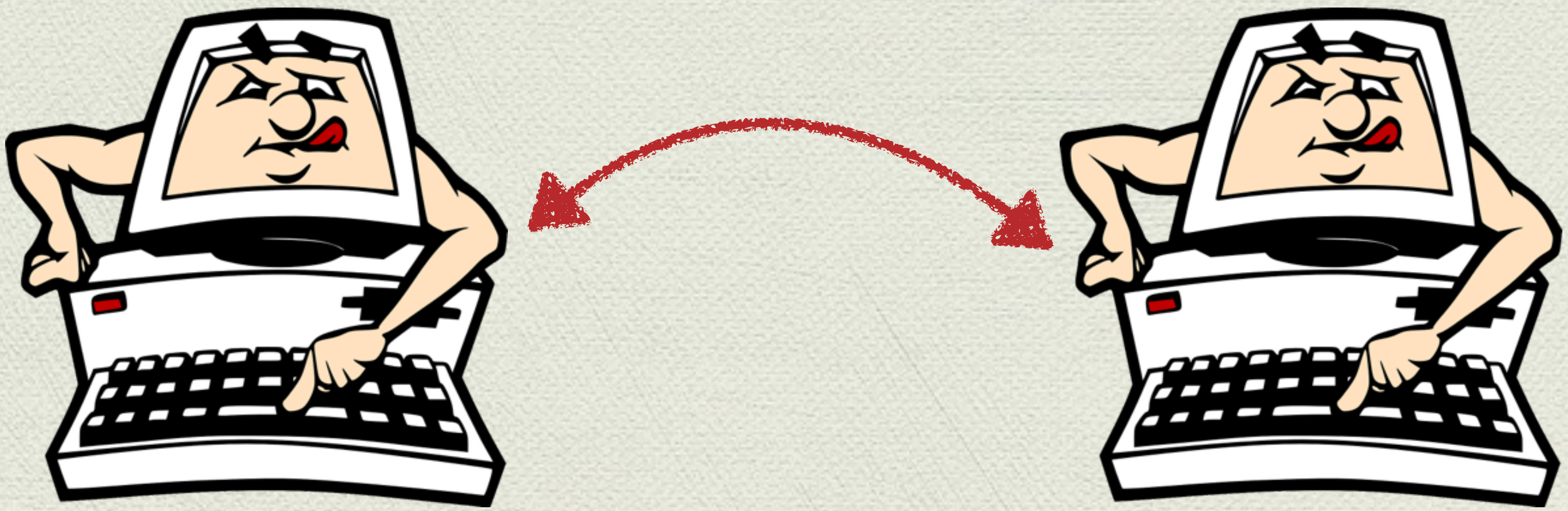
- Mutually distrusting medical servers with databases (d_1, d_2)
- Together compute joint function of input databases d_1, d_2

The Setting



- Mutually distrusting medical servers with databases (d_1, d_2)
- Together compute joint function of input databases d_1, d_2
- Maintain privacy against each other

The Setting



- Mutually distrusting medical servers with databases (d_1, d_2)
- Together compute joint function of input databases d_1, d_2
- Maintain privacy against each other

Differential Privacy

Differential Privacy [Dwork '11]

[DN'04, BDMN '05, DMNS '06]

Differential Privacy

[Dwork '11]

Differential Privacy

[DN'04, BDMN '05, DMNS '06]

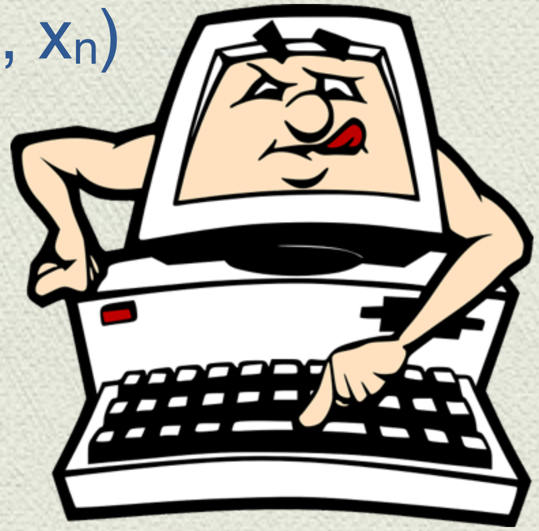
[Dwork '11]

Differential Privacy

[DN'04, BDMN '05, DMNS '06]

[Dwork '11]

$d = (x_1, x_2, \ldots, x_n)$

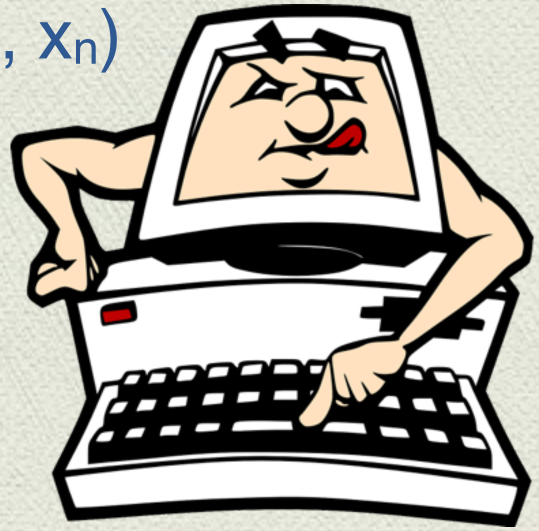


[DN’04, BDMN ’05, DMNS ’06]

Differential Privacy

[Dwork ’11]

$d = (x_1, x_2, \ldots, x_n)$

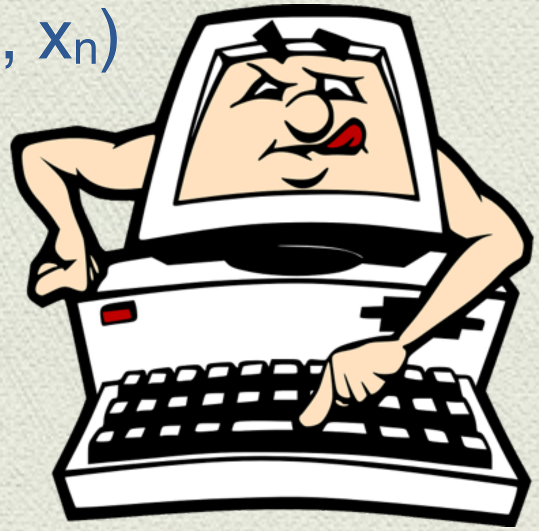


Differential Privacy

[DN'04, BDMN '05, DMNS '06]

[Dwork '11]

$d = (x_1, x_2, \ldots, x_n)$



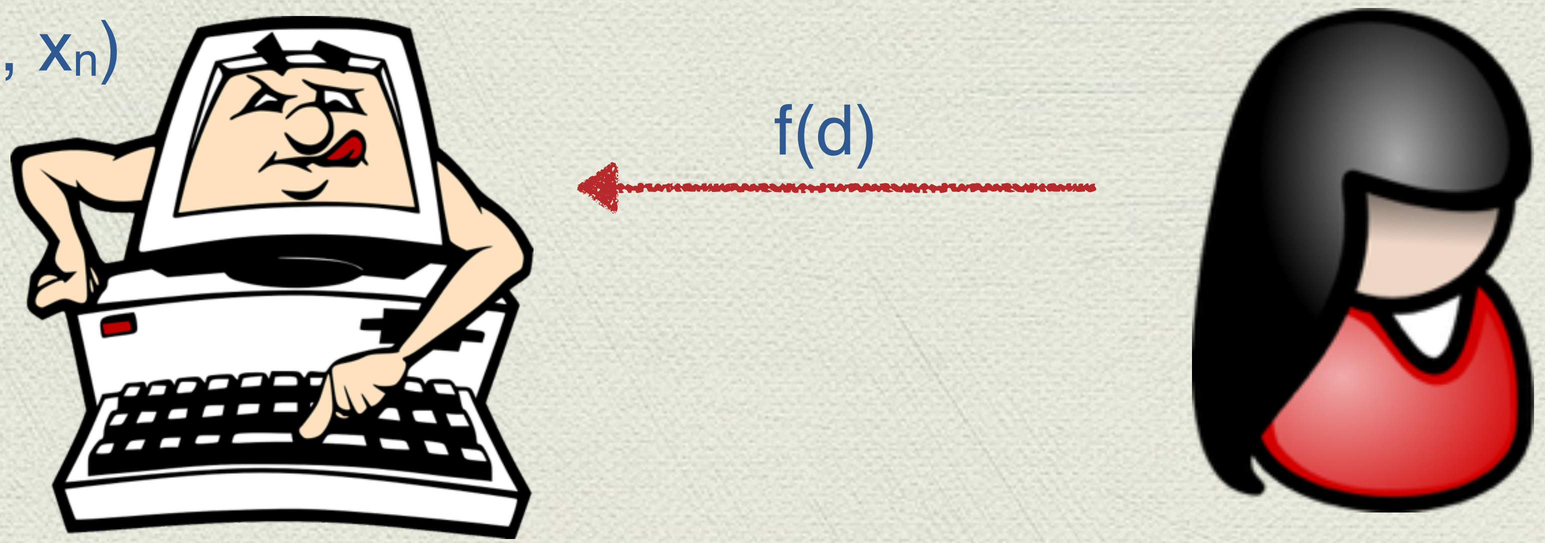
f(d)

[DN'04, BDMN '05, DMNS '06]

Differential Privacy

[Dwork '11]

$d = (x_1, x_2, \ldots, x_n)$

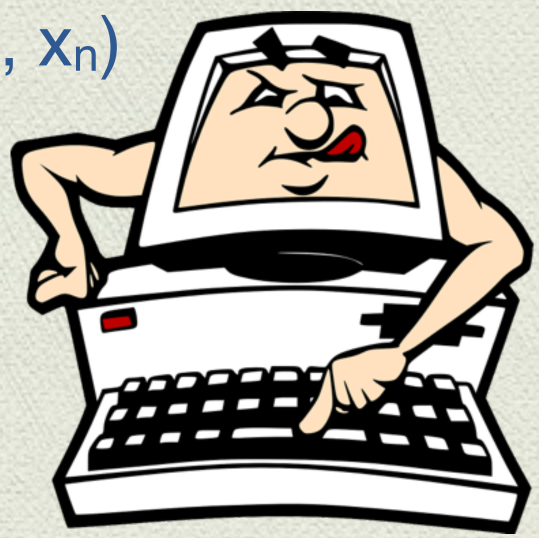


[DN'04, BDMN '05, DMNS '06]

Differential Privacy

[Dwork '11]

$d = (x_1, x_2, \ldots, x_n)$



f(d)

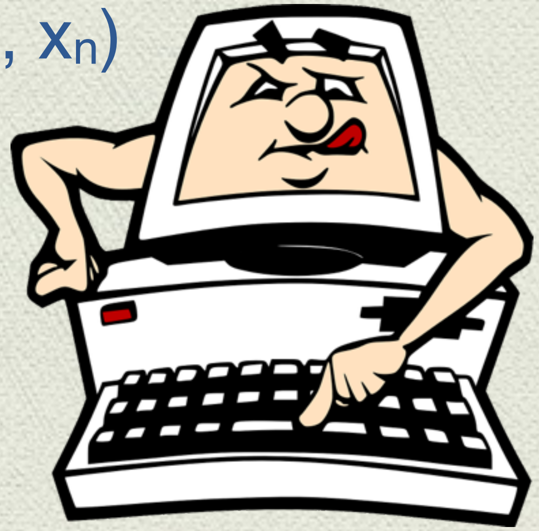
f'(d) = f(d) + noise

[DN'04, BDMN '05, DMNS '06]

Differential Privacy

[Dwork '11]

$d = (x_1, x_2, \ldots, x_n)$



f(d)

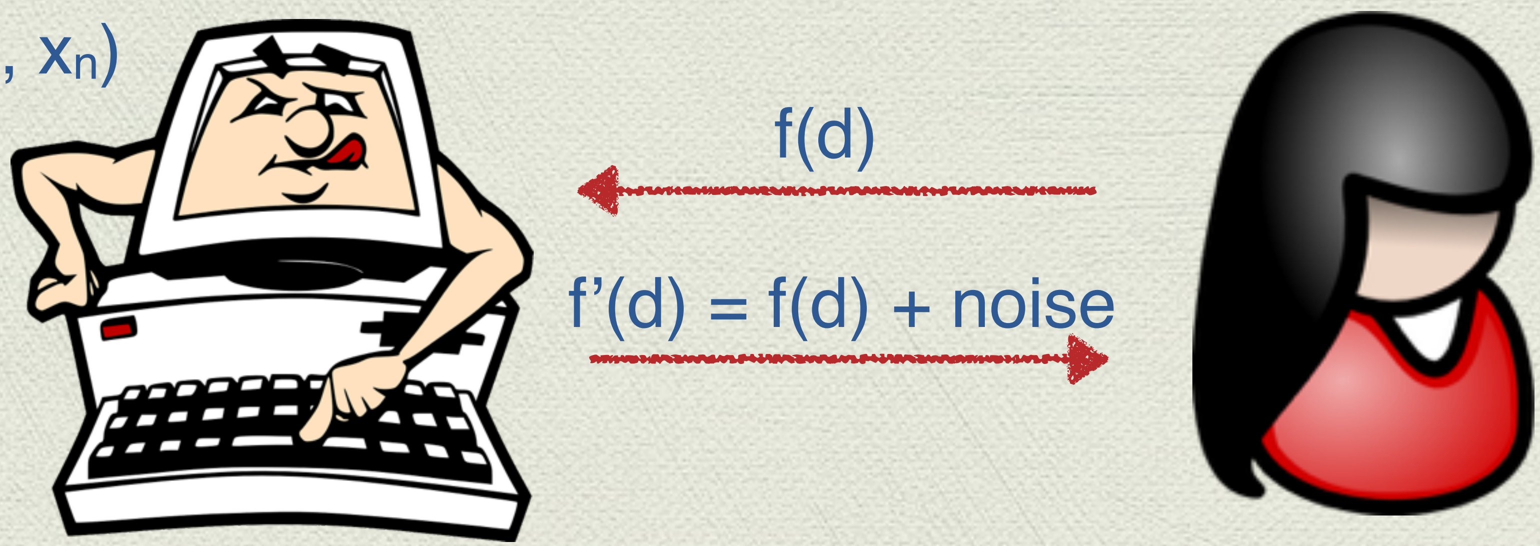
f'(d) = f(d) + noise

Differential Privacy

[DN'04, BDMN '05, DMNS '06]

[Dwork '11]

$d = (x_1, x_2, \ldots, x_n)$



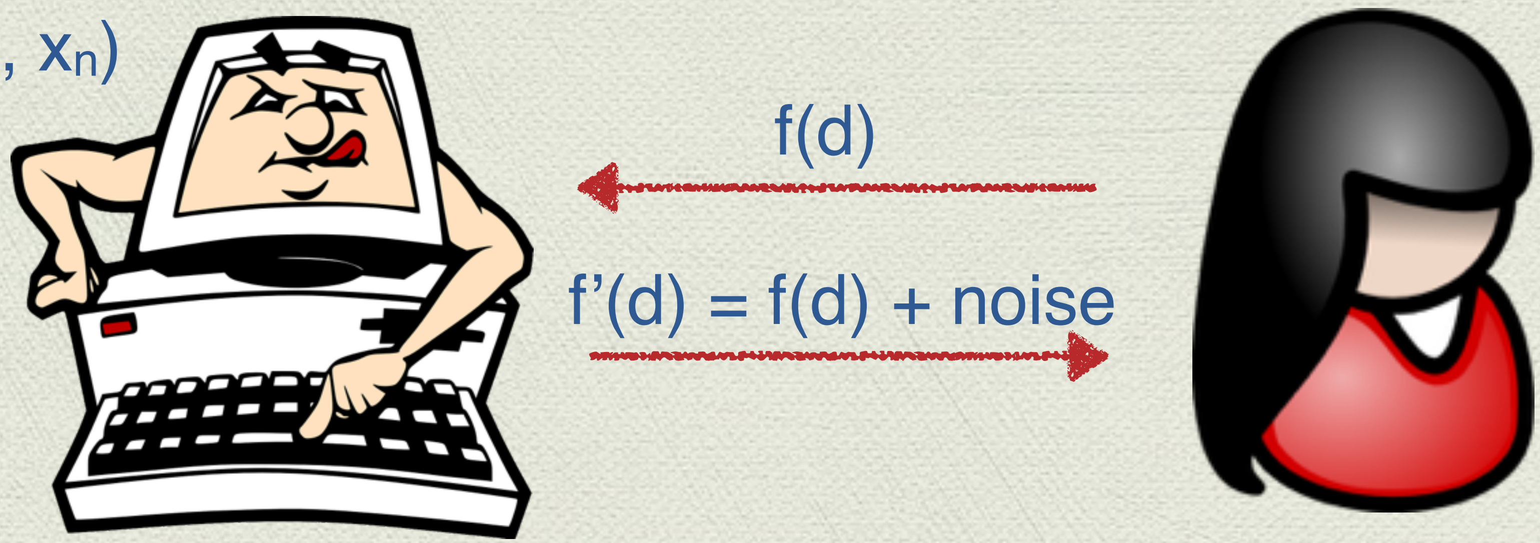
- Privacy ε
  - Outcome f'(d) not significantly influenced by exclusion of one record
  - Distribution (noise) is such that $\exp^{-1}(\varepsilon) \leq \dfrac{\Pr[f'(d) = z^*]}{\Pr[f'(d - x_i) = z^*]} \leq \exp(\varepsilon)$

[DN’04, BDMN ’05, DMNS ’06]

Differential Privacy

[Dwork ’11]

$d = (x_1, x_2, \ldots, x_n)$



- Privacy ε
  - Outcome f’(d) not significantly influenced by exclusion of one record
  - Distribution (noise) is such that $\exp^{-1}(\varepsilon) \leq \dfrac{\Pr[f'(d) = z^*]}{\Pr[f'(d - x_i) = z^*]} \leq \exp(\varepsilon)$
- Accuracy α
  - $f'(d)$ is a good estimate of $f(d)$. For f, ε, optimal accuracy $\alpha_{\varepsilon,f}^{(opt)}$

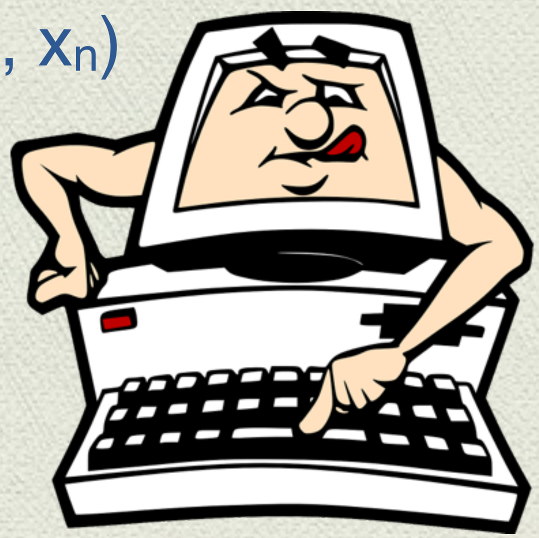
Differential Privacy

[DN'04, BDMN '05, DMNS '06]

[Dwork '11]

(ε, α) DP Protocol in Client-Server Setting

$d = (x_1, x_2, \ldots, x_n)$



f(d) ←

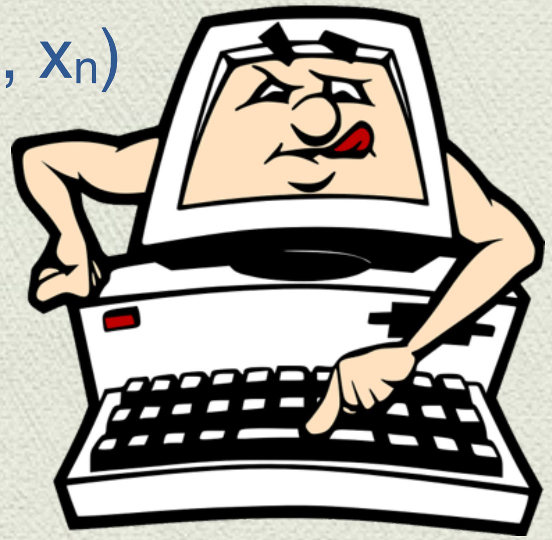
f'(d) = f(d) + noise →

- Privacy ε
  - Outcome f'(d) not significantly influenced by exclusion of one record
  - Distribution (noise) is such that $\exp^{-1}(\varepsilon) \leq \dfrac{\Pr[f'(d) = z^*]}{\Pr[f'(d - x_i) = z^*]} \leq \exp(\varepsilon)$
- Accuracy α
  - f′(d) is a good estimate of f(d). For f,ε, optimal accuracy $\alpha_{\varepsilon,f}^{(opt)}$

Distributed Differential Privacy

Distributed Differential Privacy

$d_1 = (x_1, x_2, \ldots, x_n)$

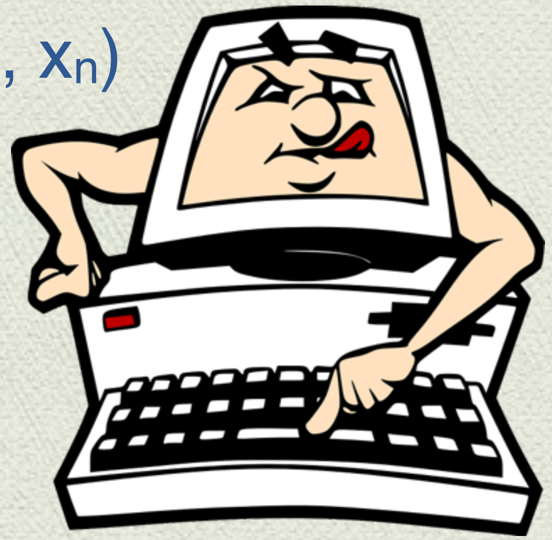


$d_2 = (y_1, y_2, \ldots, y_n)$

Distributed Differential Privacy

f

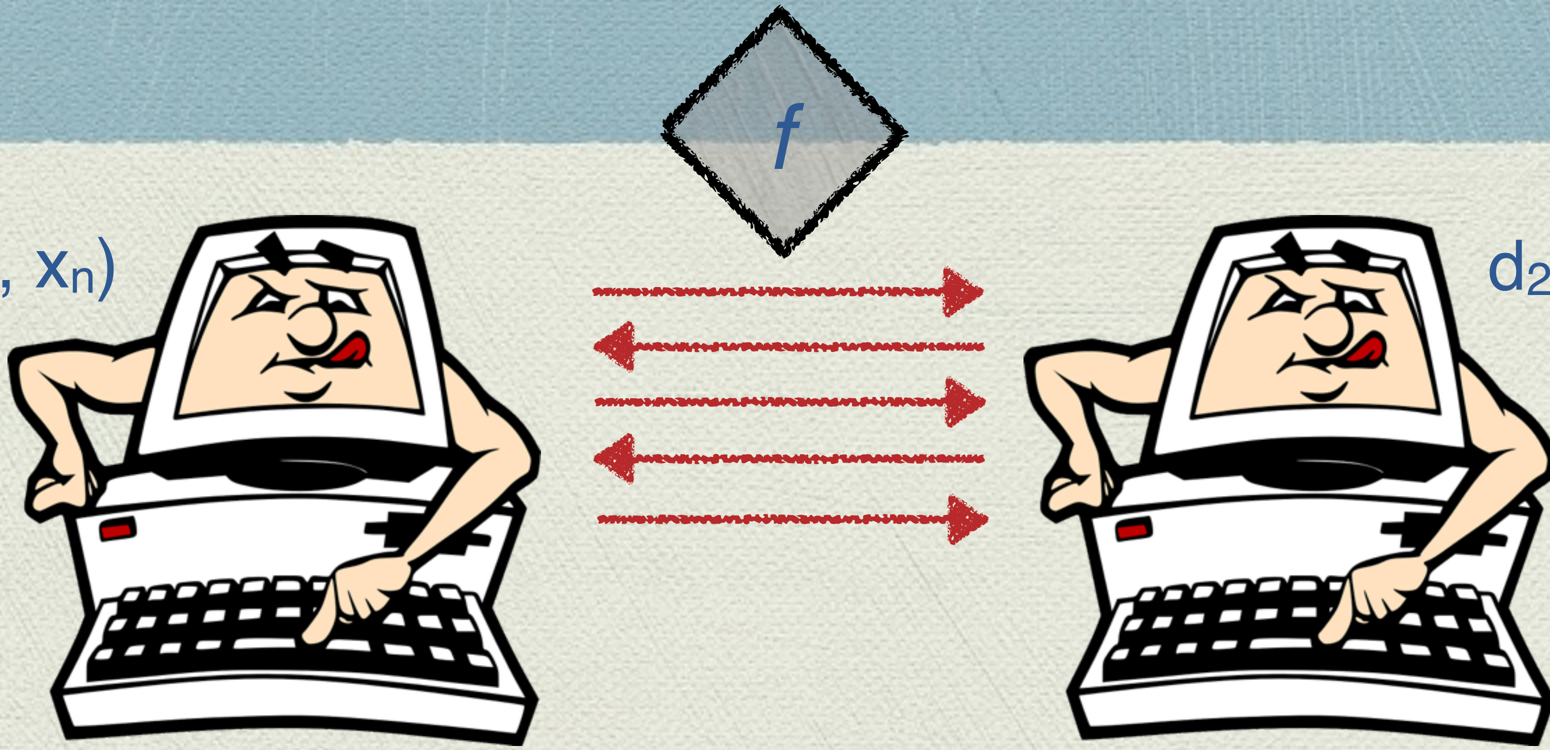
$d_1 = (x_1, x_2, \ldots, x_n)$



$d_2 = (y_1, y_2, \ldots, y_n)$

Distributed Differential Privacy

$d_1 = (x_1, x_2, \ldots, x_n)$

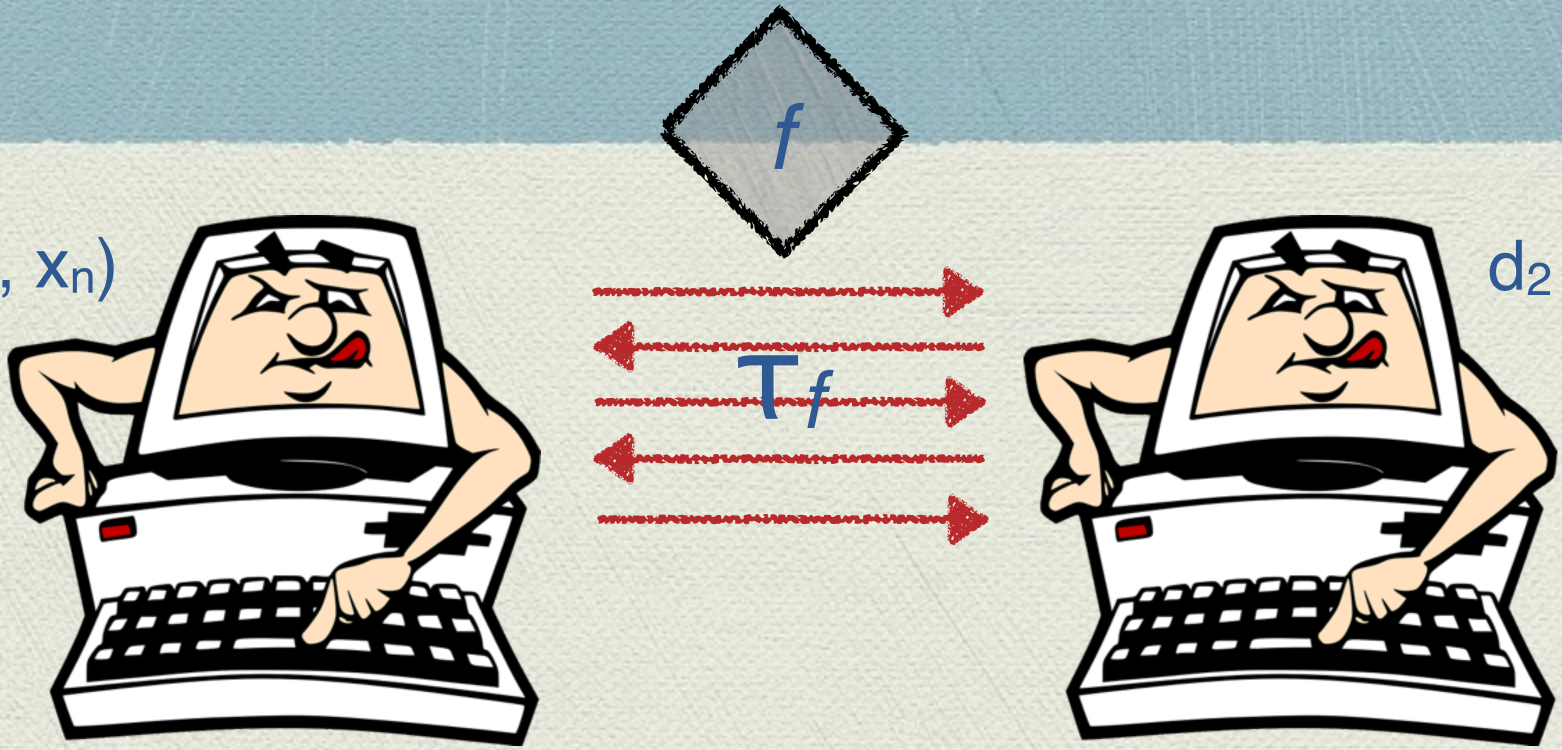


$d_2 = (y_1, y_2, \ldots, y_n)$

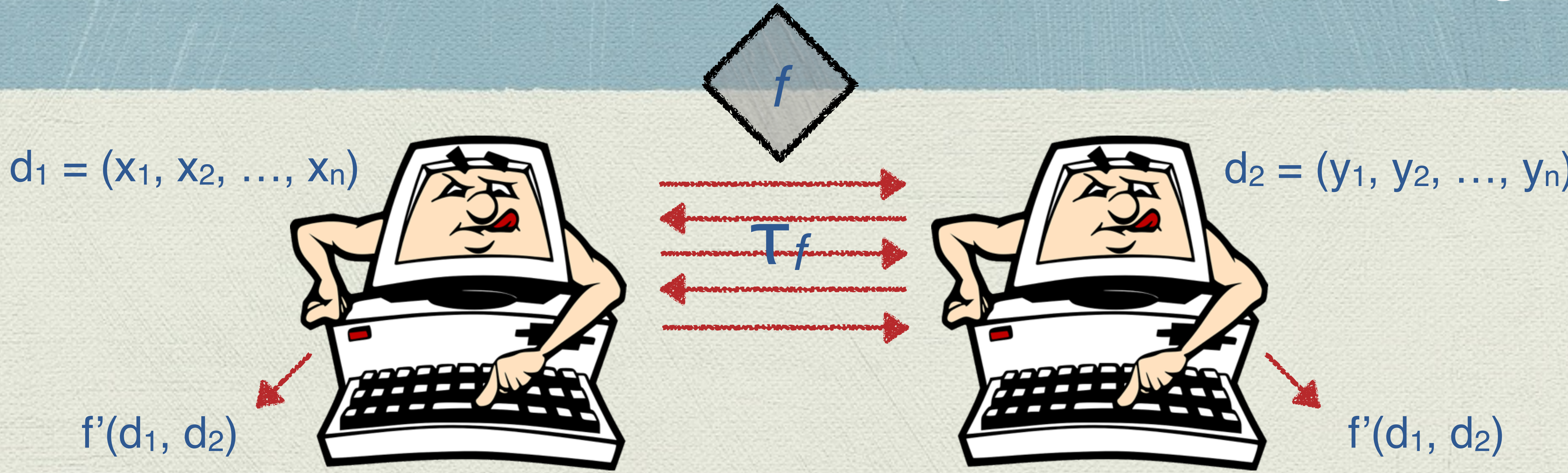
Distributed Differential Privacy

$d_1 = (x_1, x_2, \ldots, x_n)$

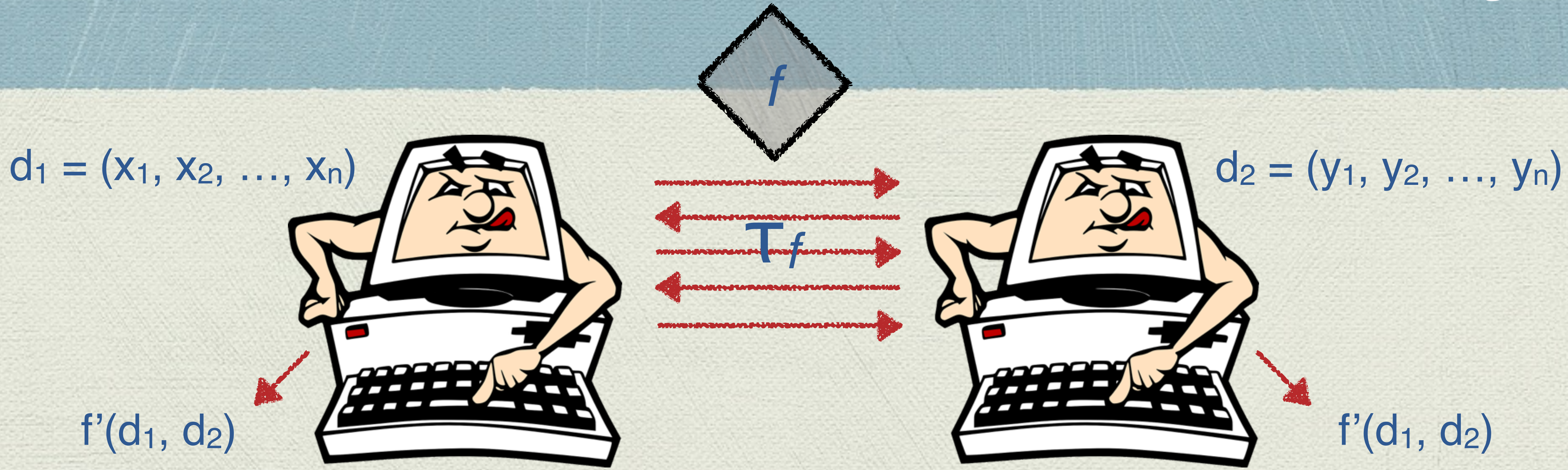
$d_2 = (y_1, y_2, \ldots, y_n)$



Distributed Differential Privacy

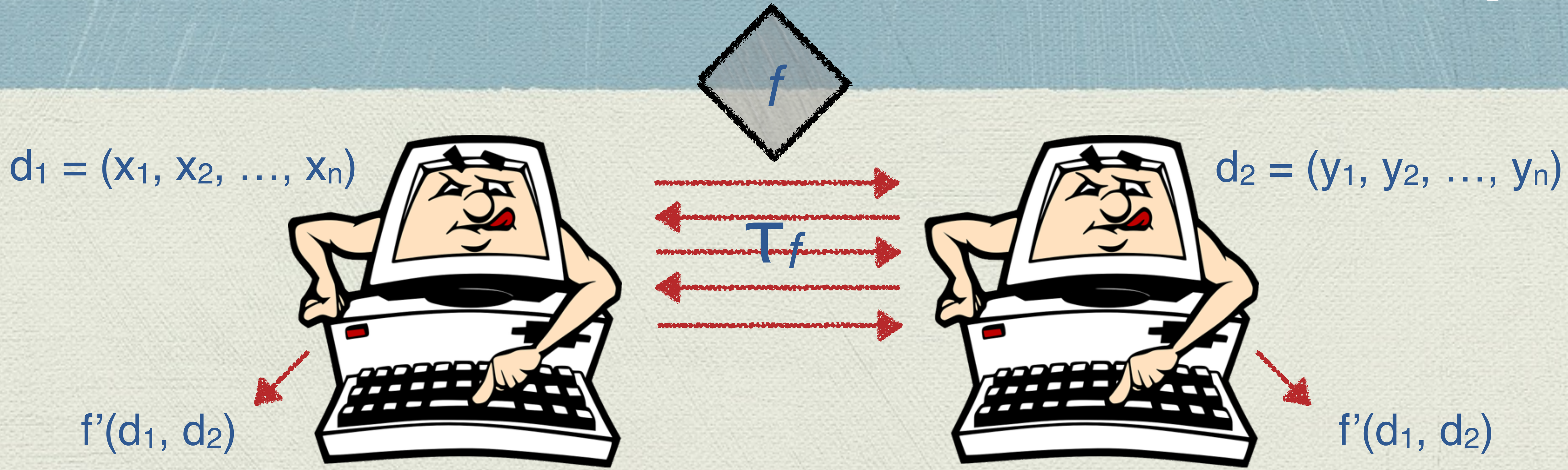


Distributed Differential Privacy



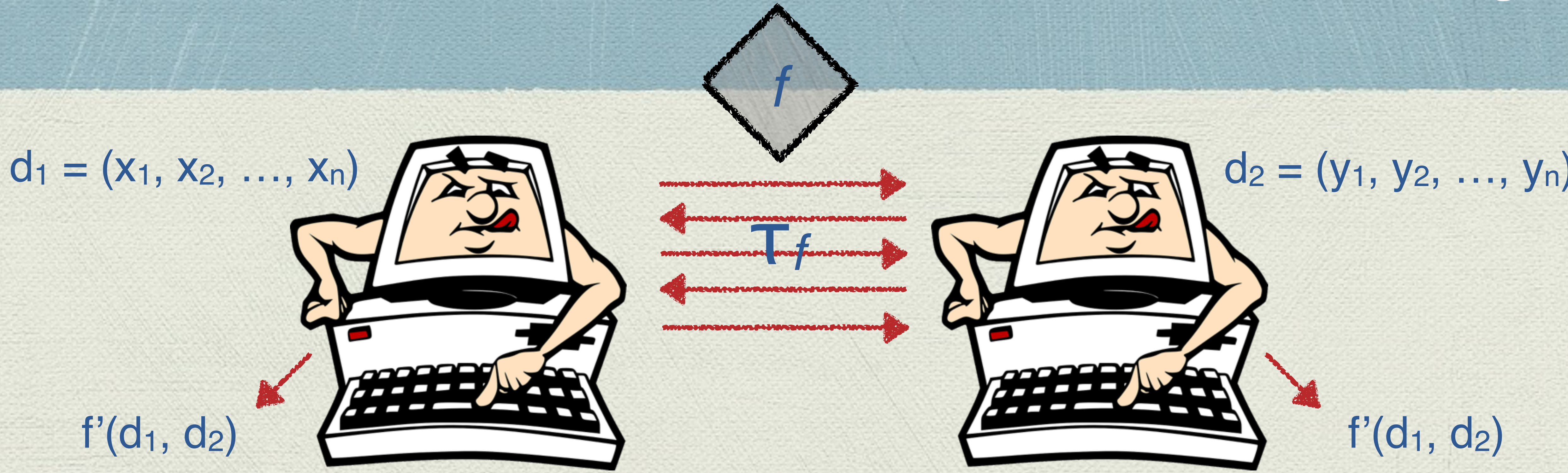
- Given function f, parties should agree on output f'

Distributed Differential Privacy



- Given function f, parties should agree on output f'
- Privacy ε
  - Distribution (τ_f) is such that $\exp^{-1}(\varepsilon) \leq \dfrac{\Pr[\tau_f = \tau_f^* \mid d_1, d_2]}{\Pr[\tau_f = \tau_f^* \mid d_1 - x_i, d_2]} \leq \exp(\varepsilon)$

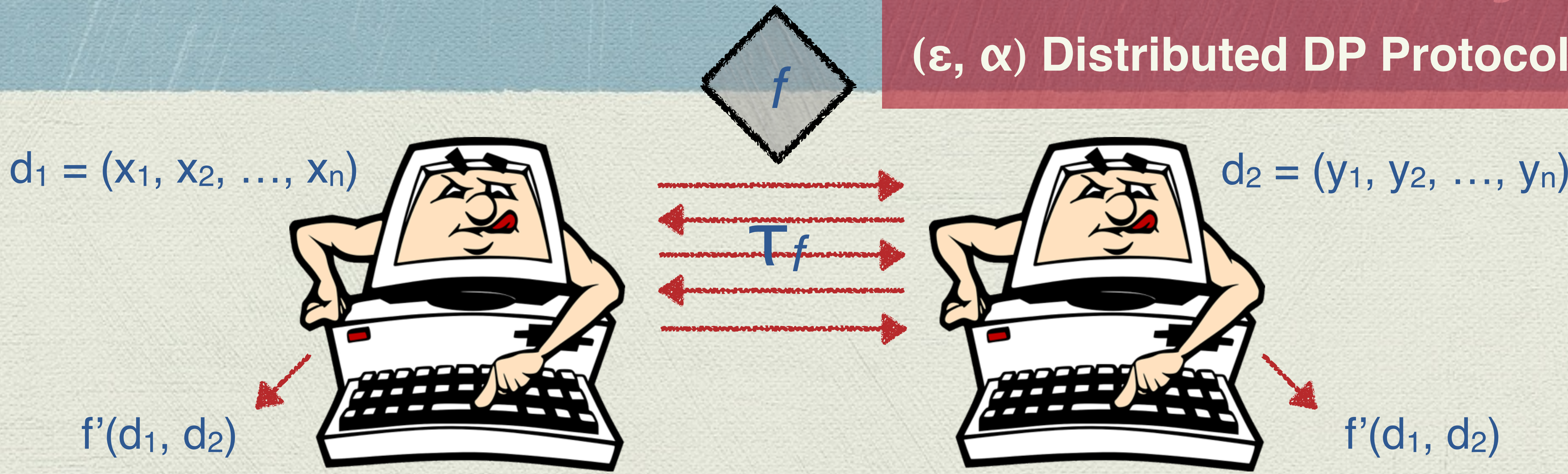
Distributed Differential Privacy



- Given function f, parties should agree on output f'
- Privacy ε
  - Distribution (τ_f) is such that $\exp^{-1}(\varepsilon) \leq \frac{\Pr[\tau_f = \tau_f^* \mid d_1, d_2]}{\Pr[\tau_f = \tau_f^* \mid d_1 - x_i, d_2]} \leq \exp(\varepsilon)$
- Accuracy α
  - $f'(d_1,d_2)$ good estimate of $f(d_1,d_2)$. For f, ε, best accuracy $\alpha_{\varepsilon,f}^{(max)}$

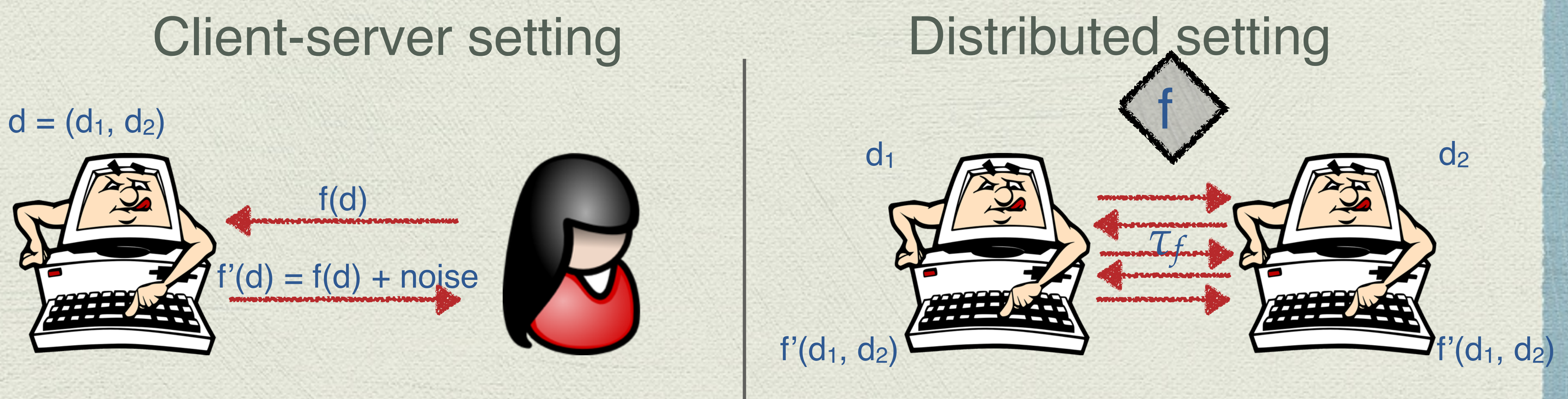
Distributed Differential Privacy

(ε, α) Distributed DP Protocol



- Given function f, parties should agree on output f'
- Privacy ε
  - Distribution (τ_f) is such that $\exp^{-1}(\varepsilon) \leq \dfrac{\Pr[\tau_f = \tau_f^* \mid d_1, d_2]}{\Pr[\tau_f = \tau_f^* \mid d_1 - x_i, d_2]} \leq \exp(\varepsilon)$
- Accuracy α
  - $f'(d_1,d_2)$ good estimate of $f(d_1,d_2)$. For f, ε, best accuracy $\alpha_{\varepsilon,f}^{(max)}$

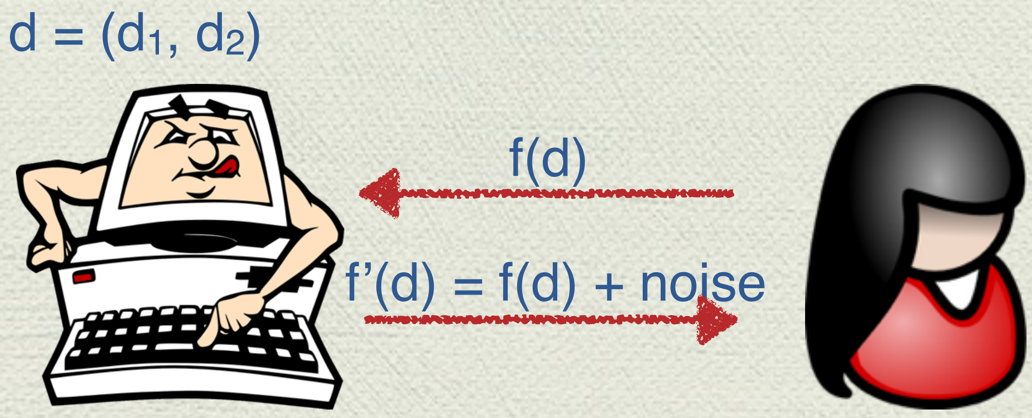
Comparing the two settings



- Distributed protocol with accuracy $\alpha_{\varepsilon,f}^{(max)}$ equal to optimal $\alpha_{\varepsilon,f}^{(opt)}$ in client-server

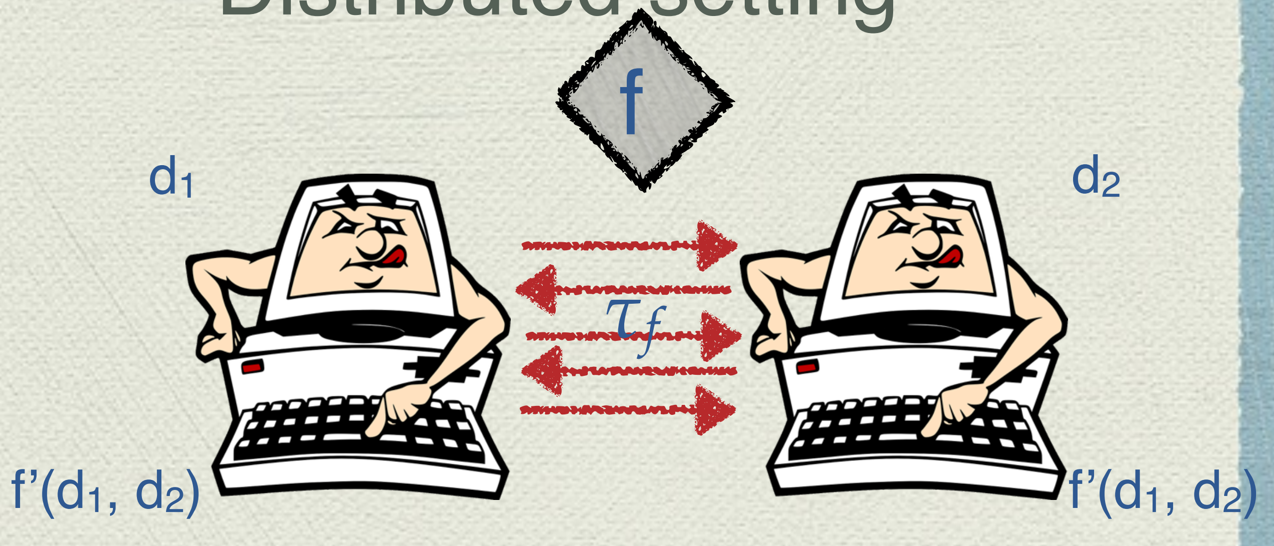
Comparing the two settings

Client-server setting



Privacy!

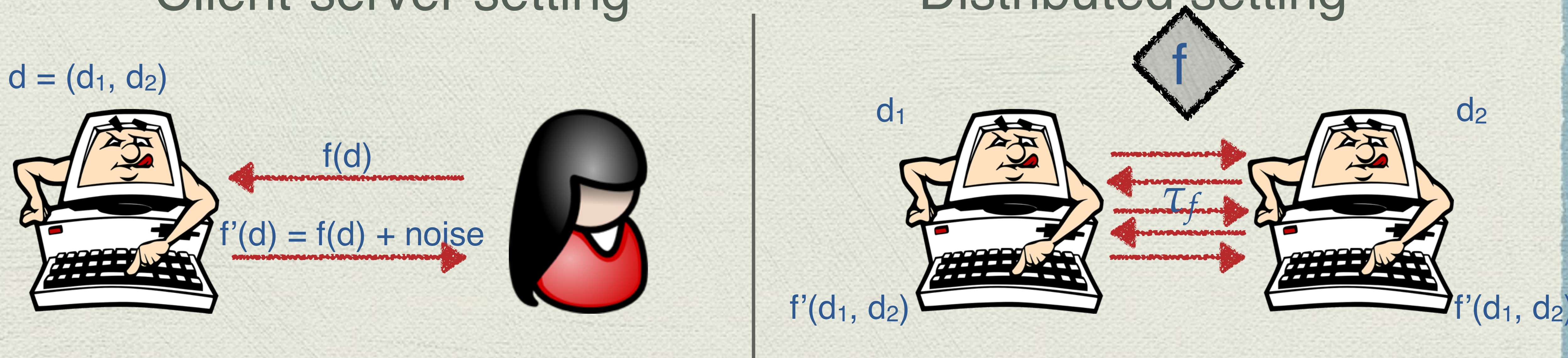
Distributed setting



- Distributed protocol with accuracy $\alpha_{\varepsilon,f}^{(max)}$ equal to optimal $\alpha_{\varepsilon,f}^{(opt)}$ in client-server

Comparing the two settings

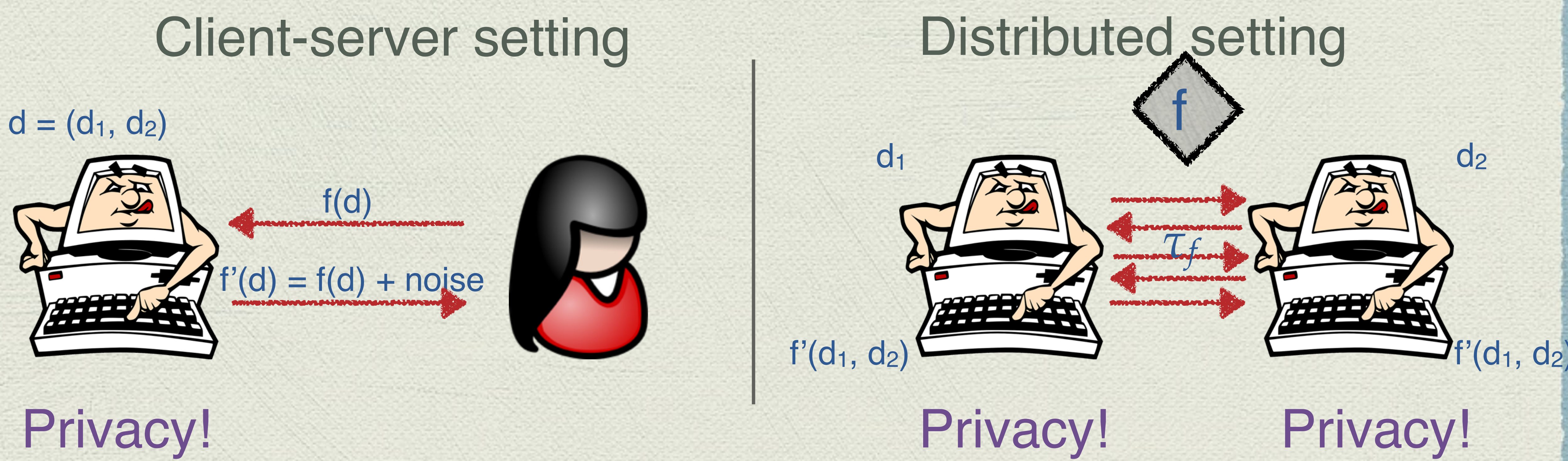
Client-server setting | Distributed setting



Privacy! | Privacy!

- Distributed protocol with accuracy $\alpha_{\varepsilon,f}^{(max)}$ equal to optimal $\alpha_{\varepsilon,f}^{(opt)}$ in client-server

Comparing the two settings



- Distributed protocol with accuracy $\alpha_{\varepsilon,f}^{(max)}$ equal to optimal $\alpha_{\varepsilon,f}^{(opt)}$ in client-server

Problem Statement

Few Observations

Few Observations

- [Goyal-Mironov-Pandey-Sahai/GMPS13]

Few Observations

- [Goyal-Mironov-Pandey-Sahai/GMPS13]
- Information-theoretically impossible to achieve optimal accuracy in the distributed setting

Few Observations

- [Goyal-Mironov-Pandey-Sahai/GMPS13]
- Information-theoretically impossible to achieve optimal accuracy in the distributed setting

	IT
Client-server	
Distributed	

Few Observations

- [Goyal-Mironov-Pandey-Sahai/GMPS13]
- Information-theoretically impossible to achieve optimal accuracy in the distributed setting

	IT
Client-server	∦
Distributed	

Few Observations

- [Goyal-Mironov-Pandey-Sahai/GMPS13]
- Information-theoretically impossible to achieve optimal accuracy in the distributed setting

	IT
Client-server	
Distributed	

$\not\parallel$

Accuracy gap exists!

Few Observations

Few Observations

- In the computational setting, existence of sh-OT suffices
- sh-OT gives MPC

Few Observations

- In the computational setting, existence of sh-OT suffices
- sh-OT gives MPC

	IT	OT
Client-server	$\nparallel$	
Distributed		

Few Observations

- In the computational setting, existence of sh-OT suffices
- sh-OT gives MPC

<table>
<tr><th></th><th>IT</th><th>OT</th></tr>
<tr><td>Client-server</td><td rowspan="2">∦</td><td rowspan="2">∥</td></tr>
<tr><td>Distributed</td></tr>
</table>

Few Observations

- In the computational setting, existence of sh-OT suffices
- sh-OT gives MPC

	IT	OT
Client-server	$\neq$	$=$
Distributed		

No accuracy gap!

A Natural Question

Minimal assumption

A Natural Question

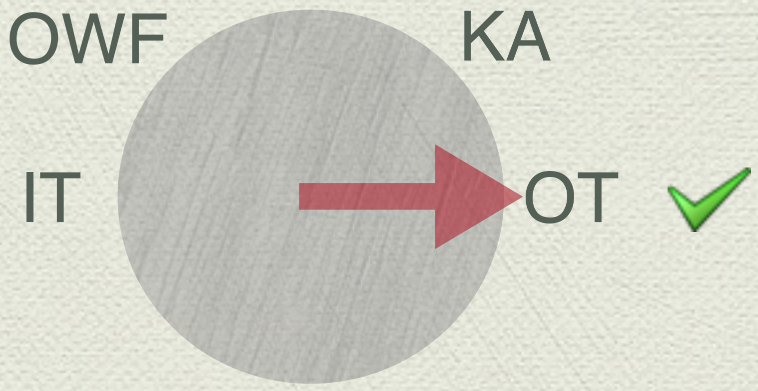
Minimal assumption

A Natural Question

Minimal assumption

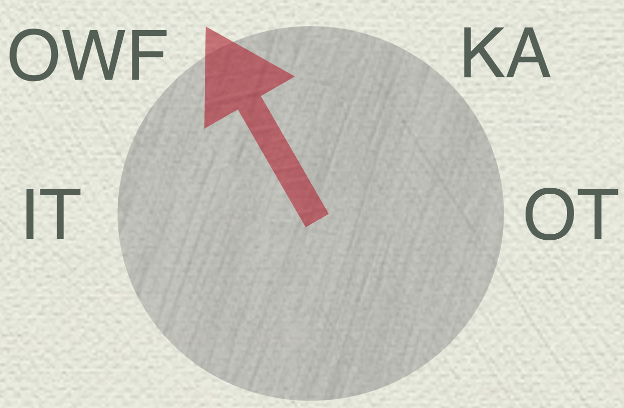
A Natural Question

Minimal assumption



A Natural Question

Minimal assumption



A Natural Question

Minimal assumption

Main Theorem

Minimal assumption

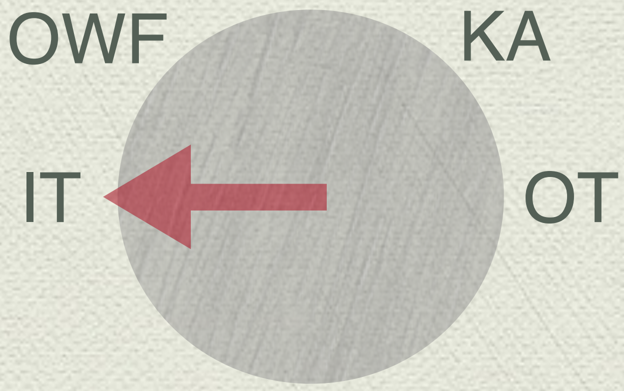
Main Theorem

Minimal assumption

Key-agreement cannot be used in a black-box manner to obtain optimally accurate distributed DP

Main Theorem

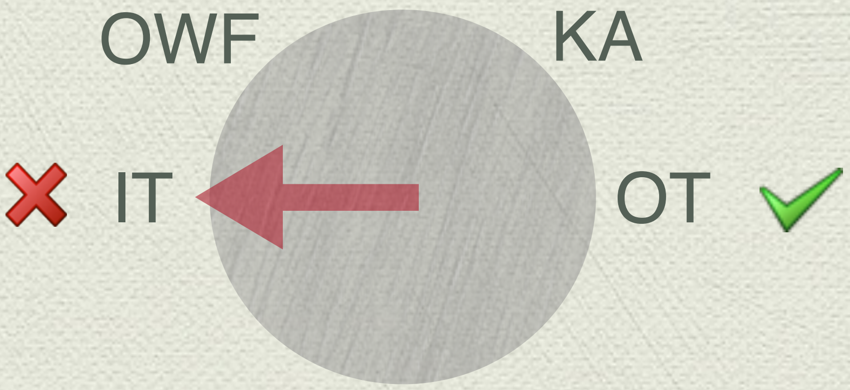
Minimal assumption



Key-agreement cannot be used in a black-box manner to obtain optimally accurate distributed DP

Main Theorem

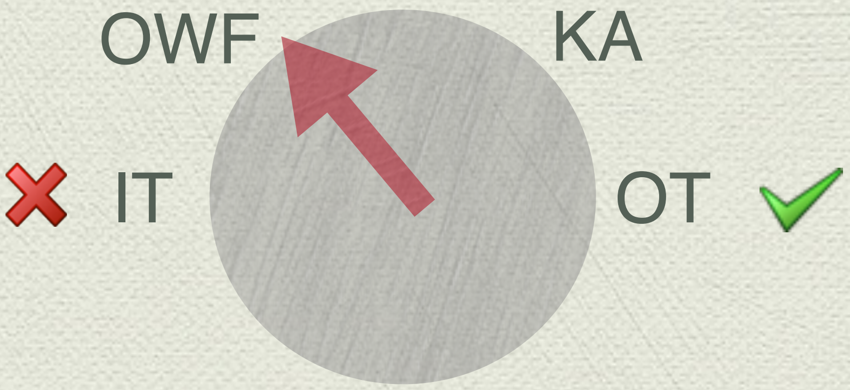
Minimal assumption



Key-agreement cannot be used in a black-box manner to obtain optimally accurate distributed DP

Main Theorem

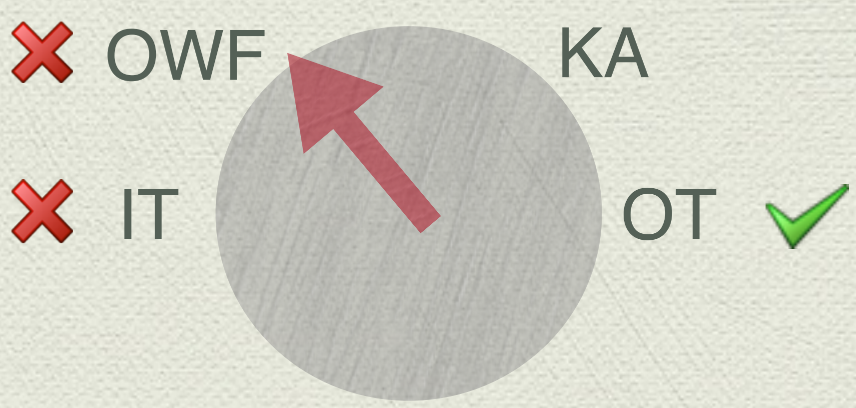
Minimal assumption



Key-agreement cannot be used in a black-box manner to obtain optimally accurate distributed DP

Main Theorem

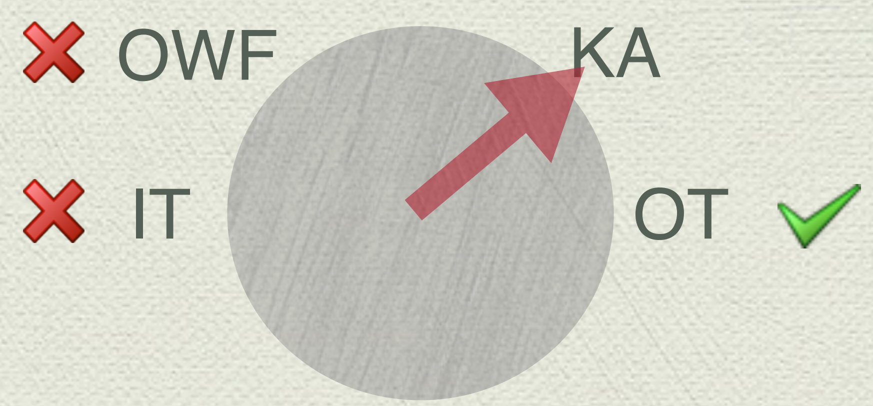
Minimal assumption



Key-agreement cannot be used in a black-box manner to obtain optimally accurate distributed DP

Main Theorem

Minimal assumption



Key-agreement cannot be used in a black-box manner to obtain optimally accurate distributed DP

Main Theorem

Minimal assumption

Key-agreement cannot be used in a black-box manner to obtain optimally accurate distributed DP

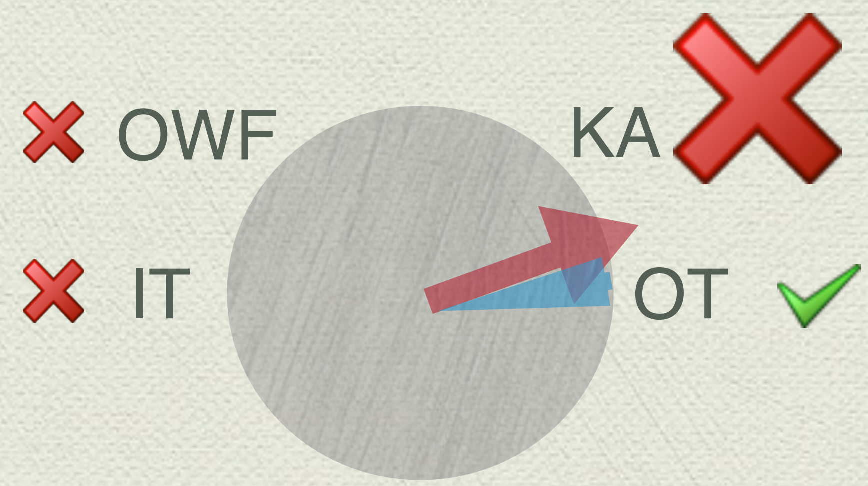
Main Theorem

Minimal assumption

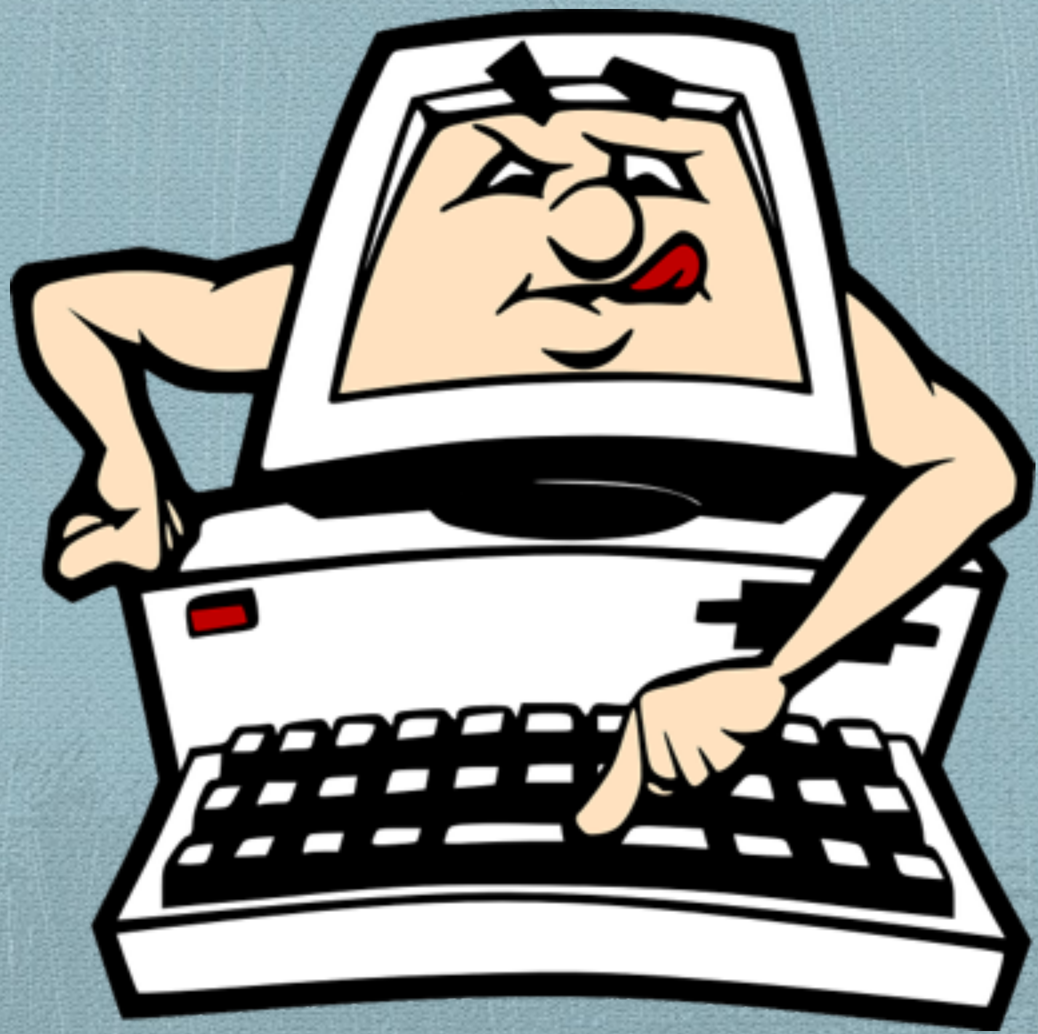
Key-agreement cannot be used in a black-box manner to obtain optimally accurate distributed DP

Main Theorem

Minimal assumption



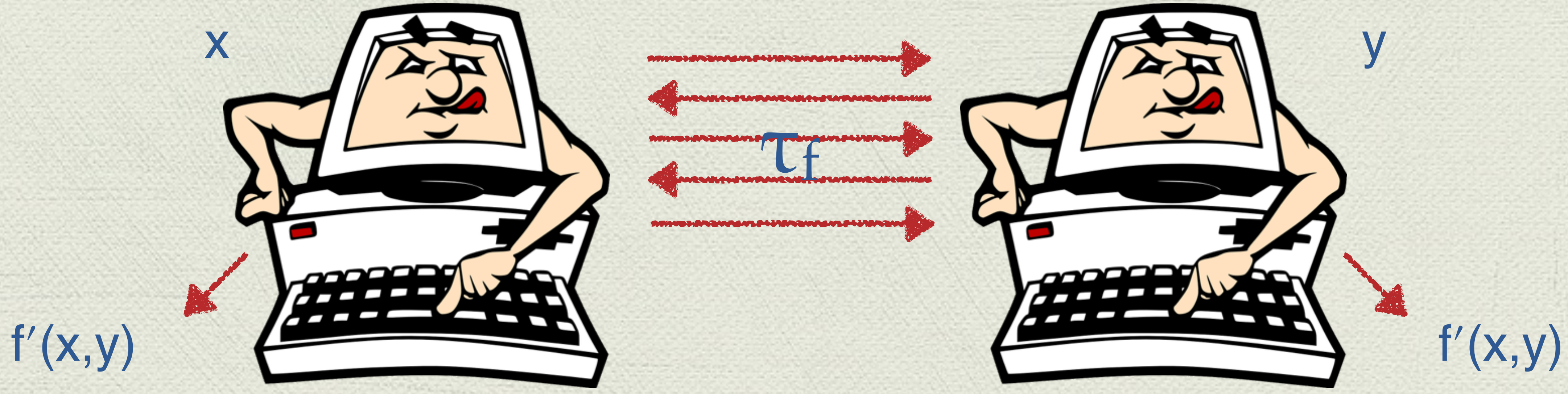
Key-agreement cannot be used in a black-box manner to obtain optimally accurate distributed DP



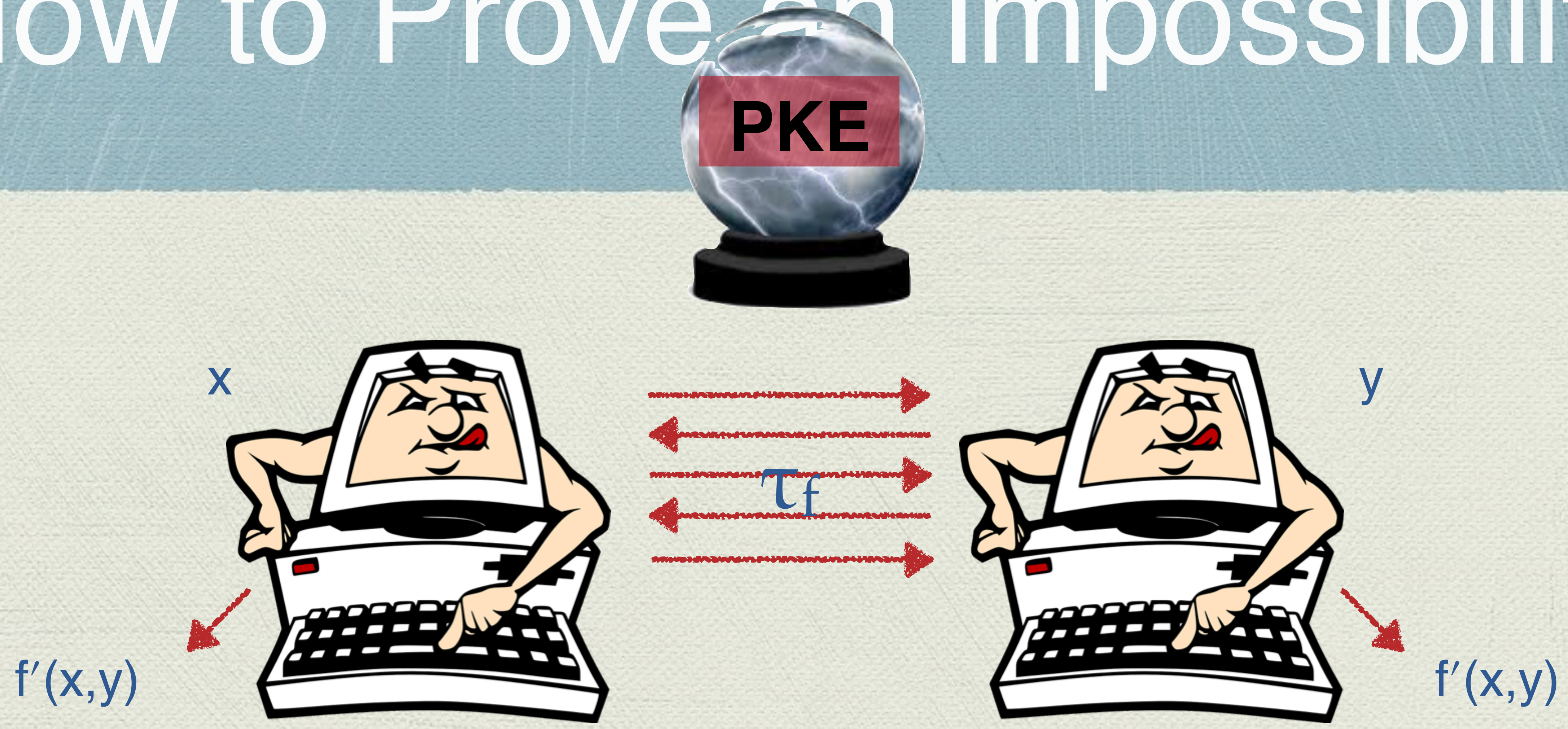
Proving the Theorem

How to Prove an Impossibility

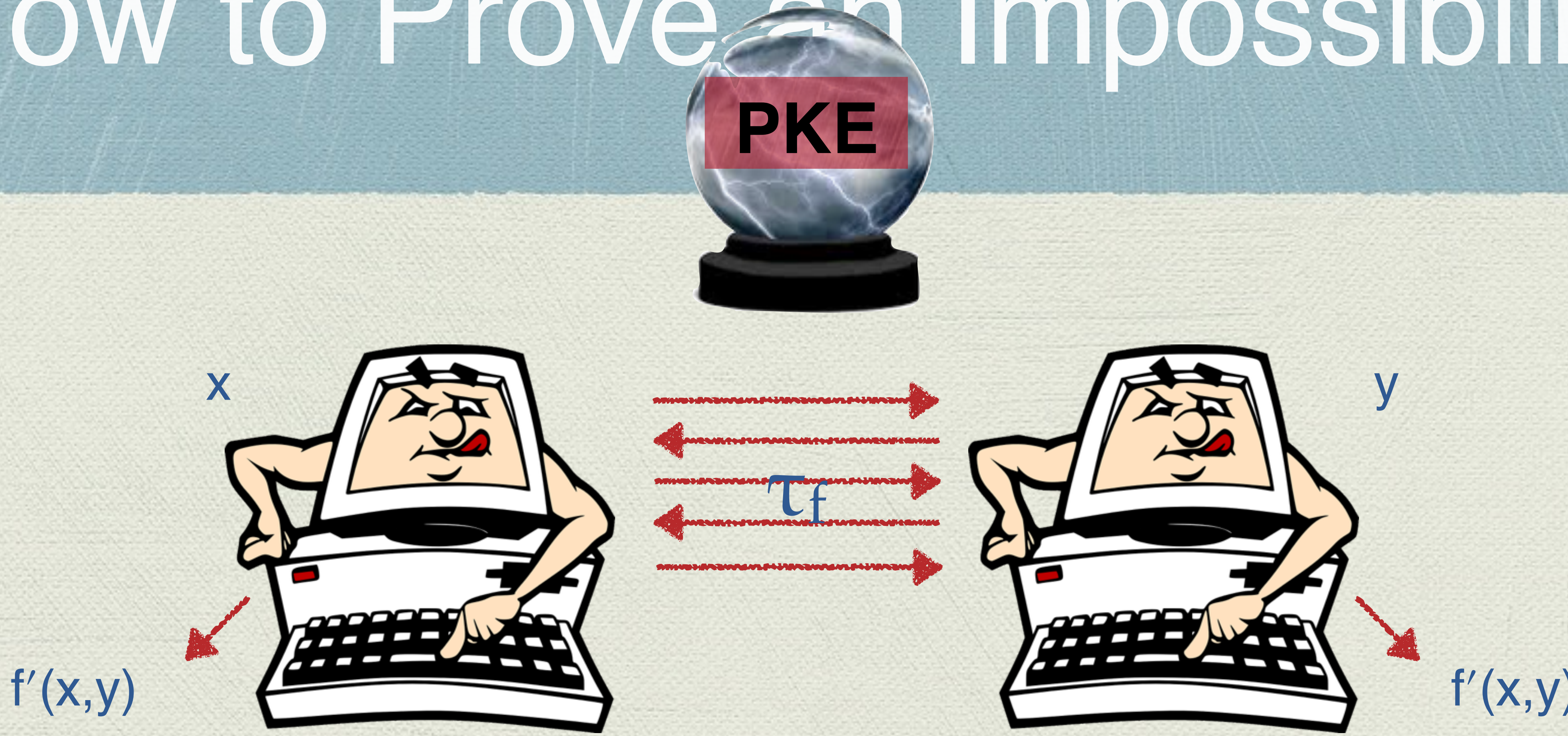
How to Prove an Impossibility



How to Prove an Impossibility



How to Prove an Impossibility



PKE oracle allows key agreement but useless for optimally accurate distributed DP.

How to Prove an impossibility?

- Black-box separation techniques [Impagliazzo-Rudich ’89, Barak-Mahmoody ’09]
- Information-theoretic impossibility in PKE oracle world
- Use [Reingold-Trevisan-Vadhan ’04] to convert oracle impossibility into a separation

What is known information theoretically?

What is known information theoretically?

- All two-party non-trivial Boolean functions
  - Accuracy. $\alpha = \min_{x,y} (\Pr[f'(x,y) = f(x,y)])$
  - Optimal. $\alpha_{f,\varepsilon}^{(opt)} = \lambda/(1+\lambda)$, for $\lambda = \exp(\varepsilon)$.

What is known information theoretically?

- All two-party non-trivial Boolean functions
  - Accuracy. $\alpha = \min_{x,y} (\Pr[f'(x,y) = f(x,y)])$
  - Optimal. $\alpha_{f,\varepsilon}^{(opt)} = \lambda/(1+\lambda)$, for $\lambda = \exp(\varepsilon)$.
- Consider representative AND and XOR

What is known information theoretically?

- All two-party non-trivial Boolean functions
  - Accuracy. $\alpha = \min_{x,y} (\Pr[f'(x,y) = f(x,y)])$
  - Optimal. $\alpha_{f,\varepsilon}^{(opt)} = \lambda/(1+\lambda)$, for $\lambda = \exp(\varepsilon)$.
- Consider representative AND and XOR
- Maximal achievable distributed information-theoretic accuracy [GMPS13]
  - AND. $\alpha_{IT,AND,\varepsilon}^{(dist)} = \lambda(\lambda^2 + \lambda + 2)/(\lambda+1)^3$, for $\lambda = \exp(\varepsilon)$.
  - XOR. $\alpha_{IT,XOR,\varepsilon}^{(dist)} = (\lambda^2 + 1)/(\lambda+1)^2$, for $\lambda = \exp(\varepsilon)$.

In the PKE Oracle World

- We will show that maximal achievable distributed accuracy in PKE oracle world
  - AND. $\alpha_{PKE,AND,\varepsilon}^{(dist)} \approx \lambda(\lambda^2 + \lambda + 2)/(\lambda+1)^3$, for $\lambda = \exp(\varepsilon)$.
  - XOR. $\alpha_{PKE,XOR,\varepsilon}^{(dist)} \approx (\lambda^2 + 1)/(\lambda+1)^2$, for $\lambda = \exp(\varepsilon)$.

Oracle Separation

Information Theoretic Plain Model

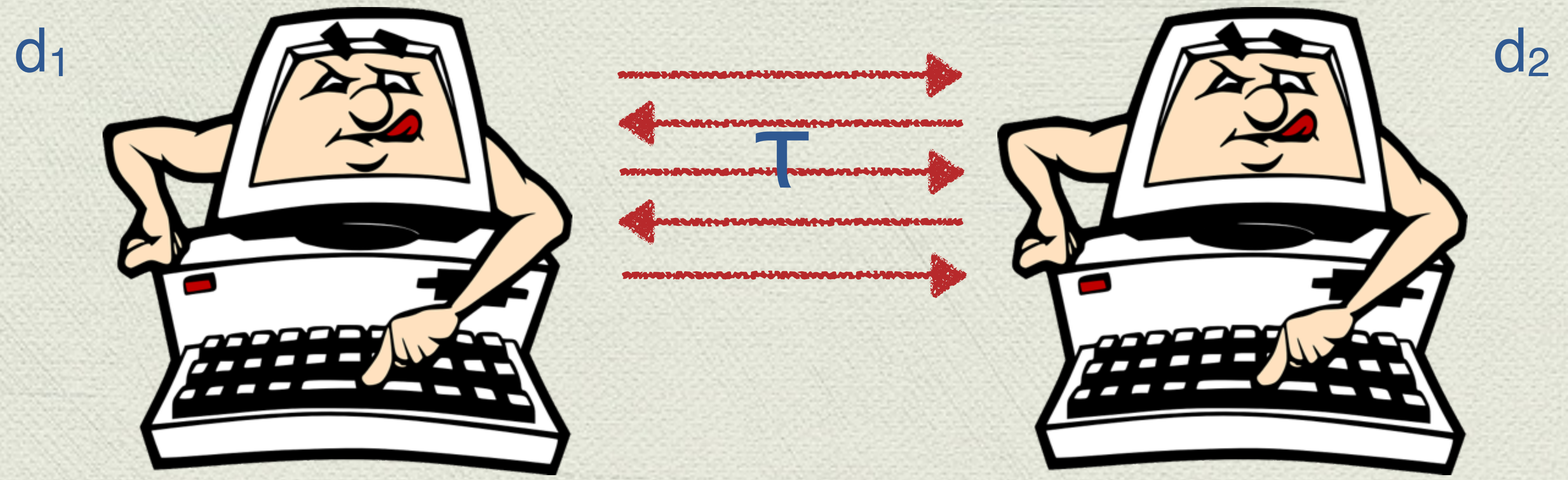
Information Theoretic Plain Model

Information Theoretic Plain Model

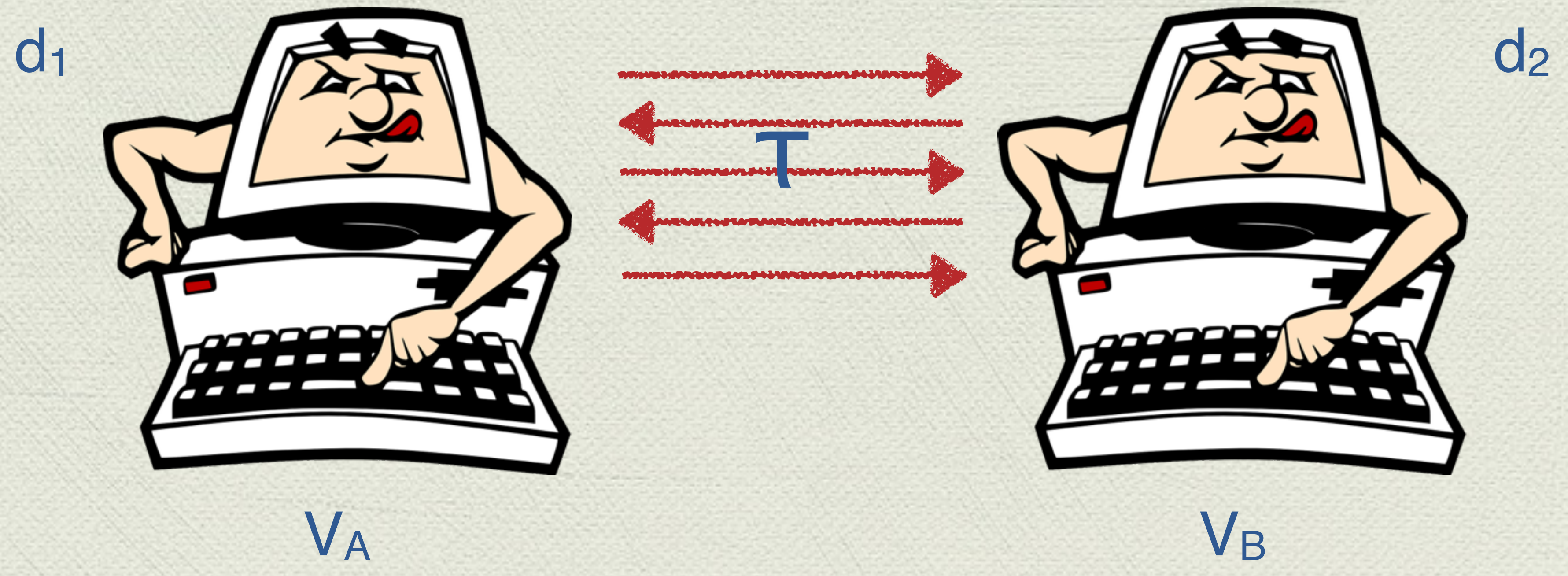
d_1

d_2

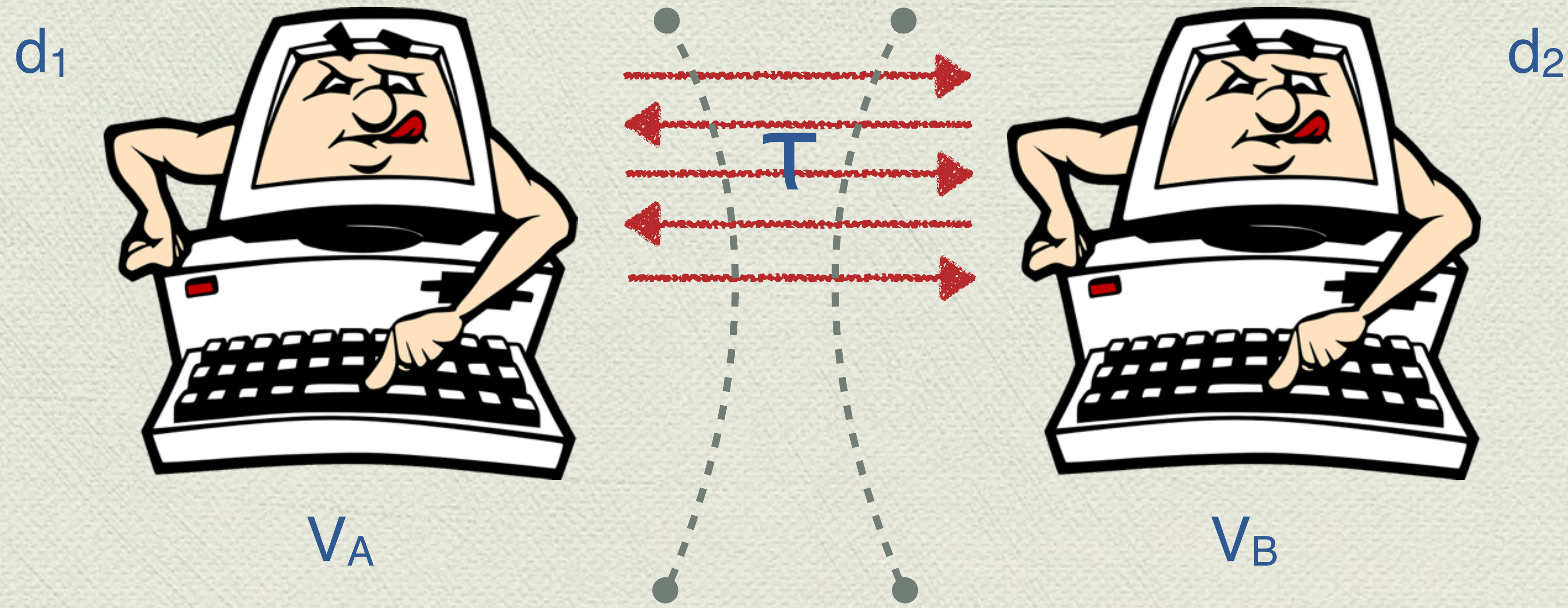
Information Theoretic Plain Model



Information Theoretic Plain Model

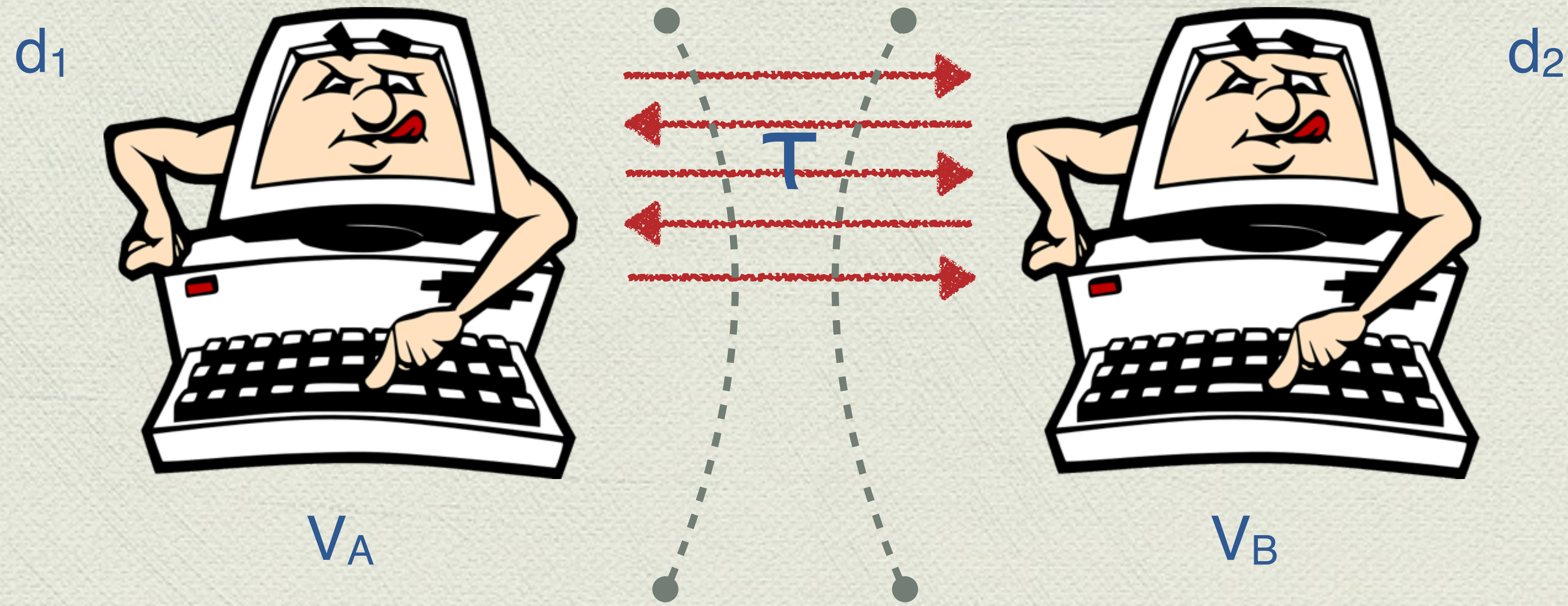


Information Theoretic Plain Model



- Independent views.
  $(V_A \times V_B | \tau, d_1, d_2) = (V_A | \tau, d_1) \times (V_B | \tau, d_2)$

Information Theoretic Plain Model



- Independent views.

  $(V_A \times V_B \mid \tau, d_1, d_2) = (V_A \mid \tau, d_1) \times (V_B \mid \tau, d_2)$

  ⇒ optimal accuracy cannot be achieved [GMPS13]

Information Theoretic PKE World

Information Theoretic PKE World

Information Theoretic PKE World

d_1

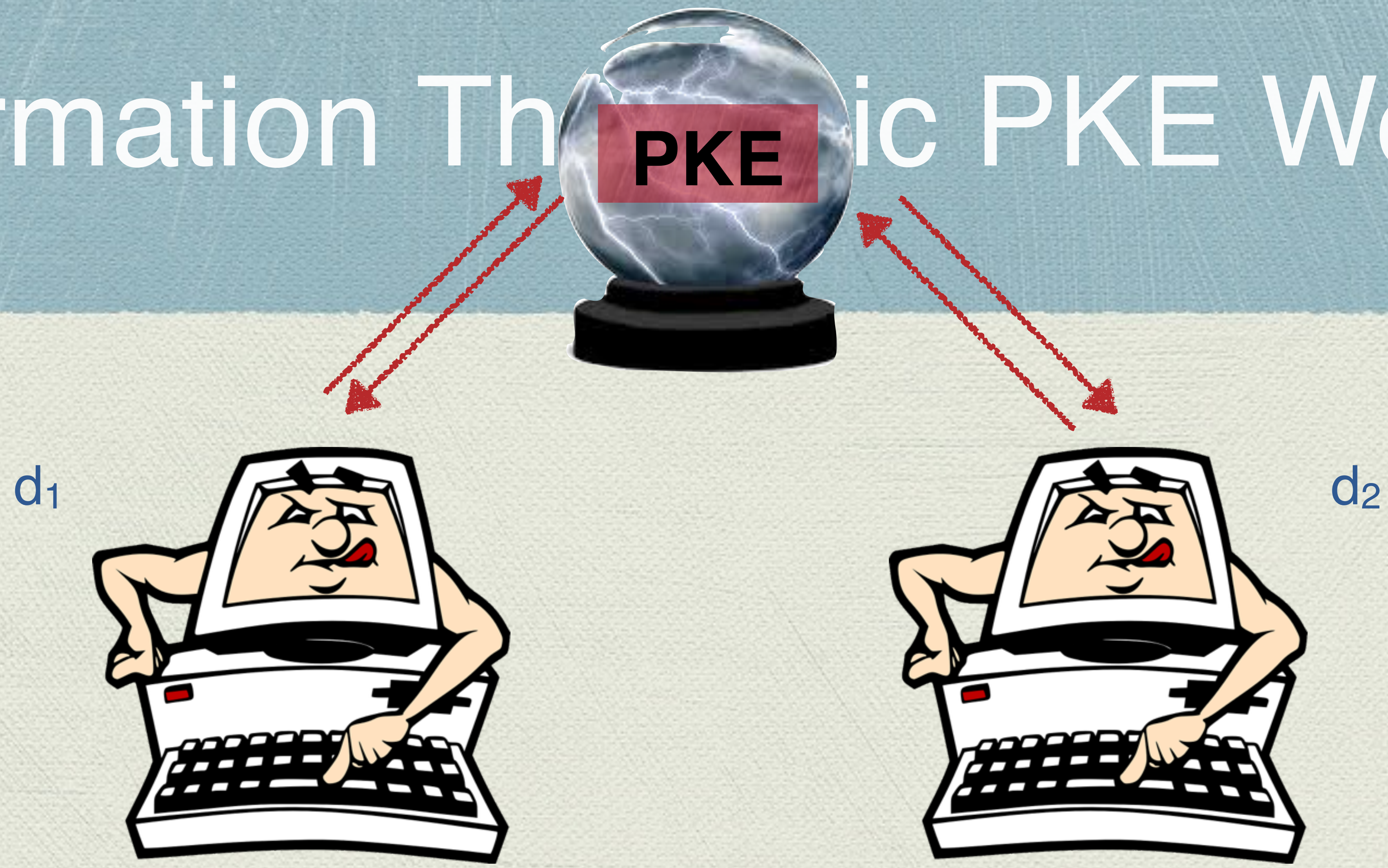
d_2

Information Th ic PKE World

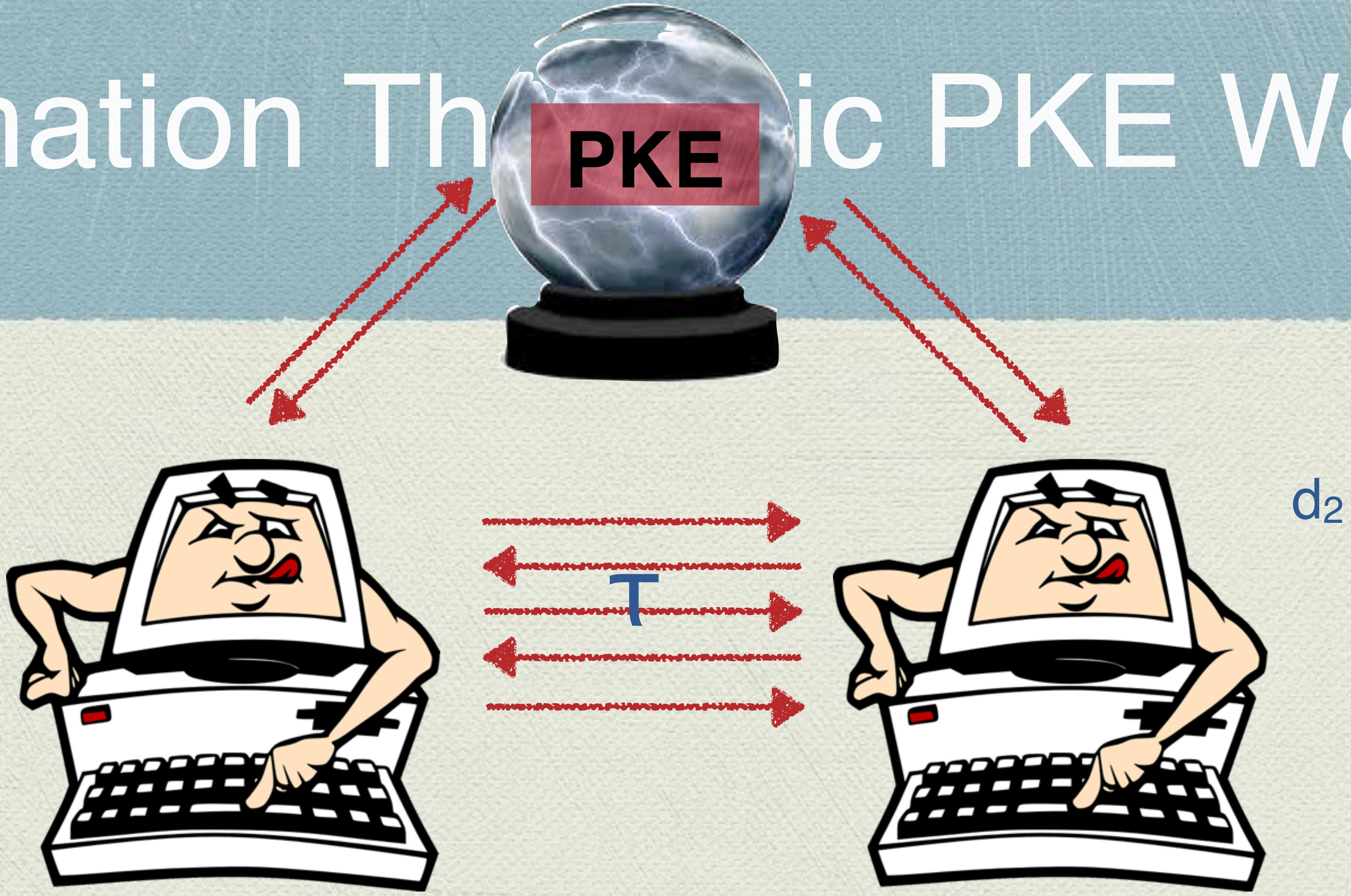
d_1

d_2

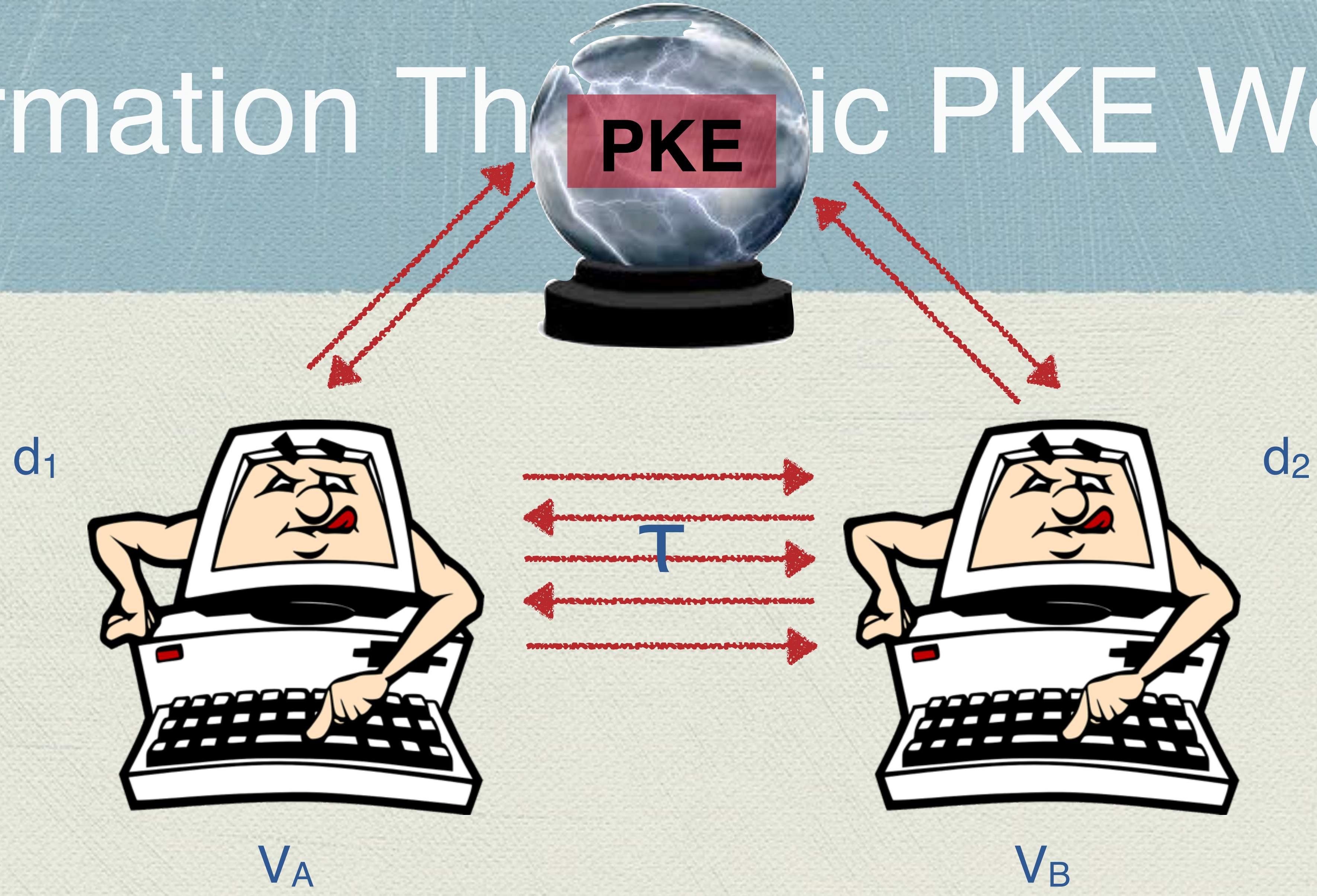
Information Th ic PKE World



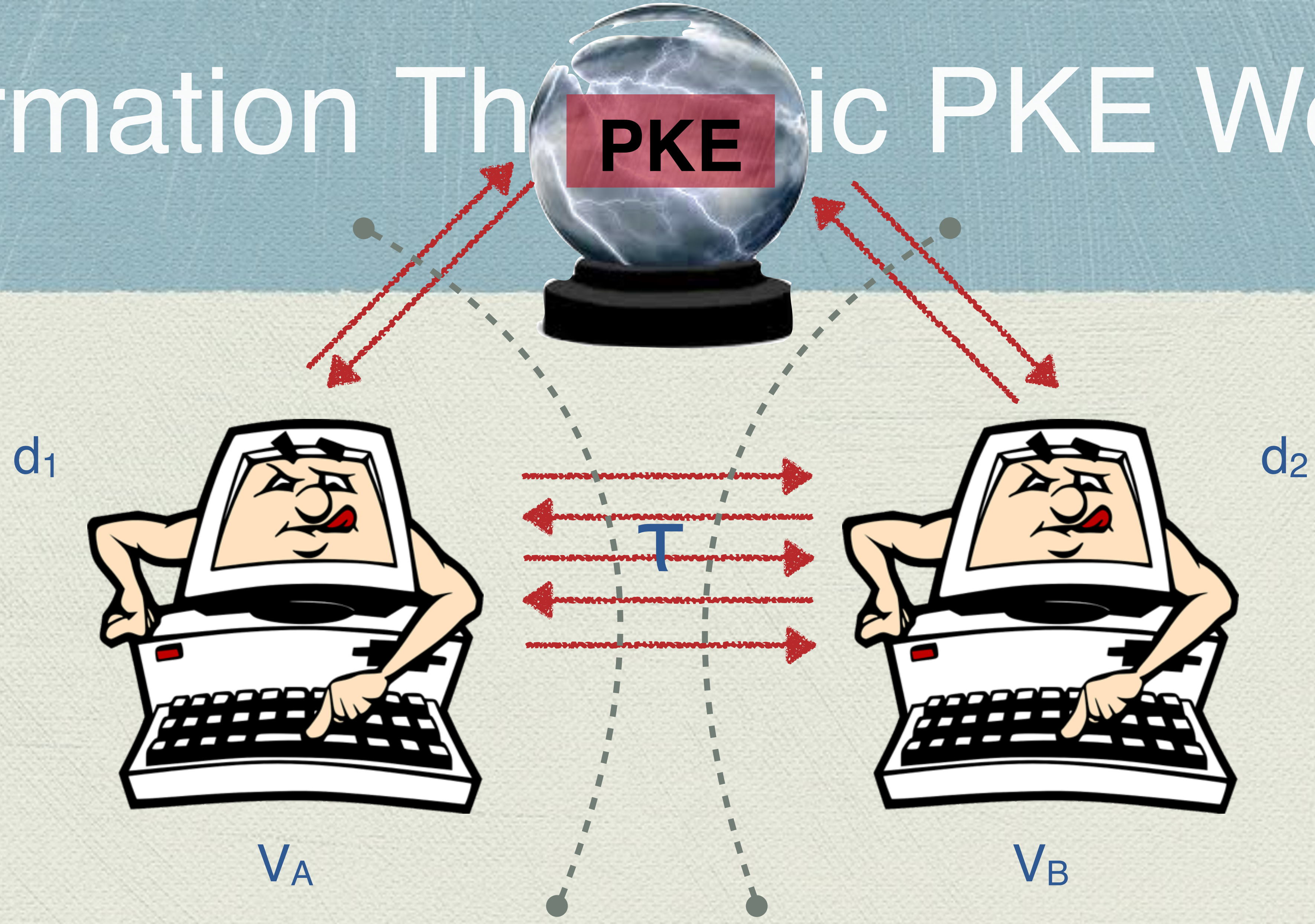
Information Th ic PKE World



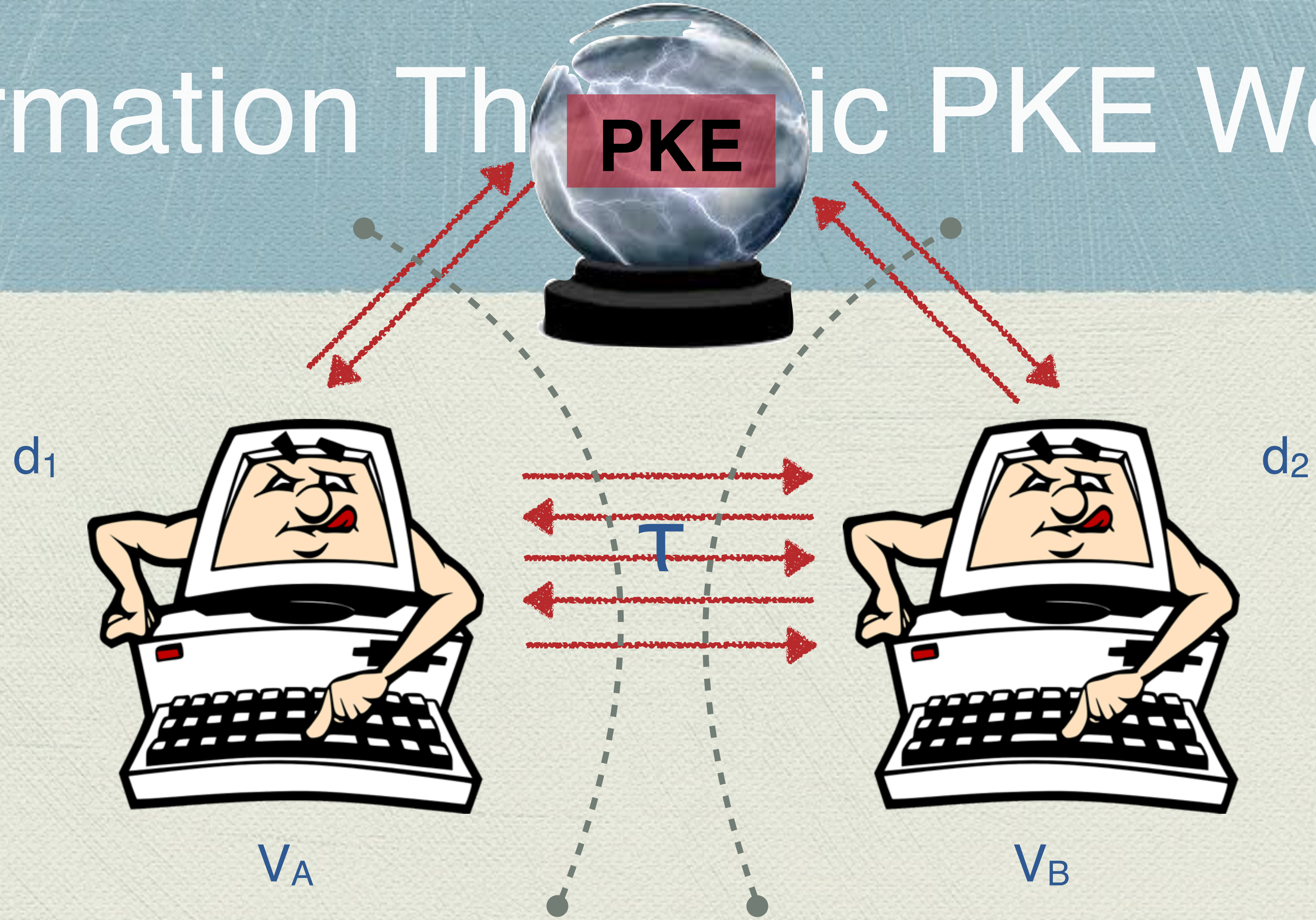
Information Th ic PKE World



Information Th ic PKE World



Information Th ic PKE World



- Views no longer independent

⇒ optimal accuracy could possibly be achieved

PKE Oracle

PKE Oracle

- **PKE = (Gen, Enc, Dec)**
- Gen (sk) → pk. Length-tripling Random Oracle
- Enc $^{(pk)}$(m) → c. (Collection of keyed) Length-tripling Random Oracles
- Dec $^{(sk)}$(c) → m. (Smallest) m: pk=Gen(sk), c=Enc$^{(pk)}$(m)

PKE Oracle

- **PKE = (Gen, Enc, Dec)**
- Gen (sk) → pk.   Length-tripling Random Oracle
- Enc $^{(pk)}$(m) → c.   (Collection of keyed) Length-tripling Random Oracles
- Dec $^{(sk)}$(c) → m.   (Smallest) m: pk=Gen(sk), c=Enc$^{(pk)}$(m)

OT!!

PKE Oracle

- **PKE = (Gen, Enc, Dec)**
- Gen (sk) → pk. Length-tripling Random Oracle
- Enc $^{(pk)}$(m) → c. (Collection of keyed) Length-tripling Random Oracles
- Dec $^{(sk)}$(c) → m. (Smallest) m: pk=Gen(sk), c=Enc$^{(pk)}$(m)

PKE Oracle

- **PKE = (Gen, Enc, Dec) + ($Test_1$, $Test_2$)**
- Gen (sk) → pk. Length-tripling Random Oracle
- Enc $^{(pk)}$(m) → c. (Collection of keyed) Length-tripling Random Oracles
- Dec $^{(sk)}$(c) → m. (Smallest) m: pk=Gen(sk), c=Enc$^{(pk)}$(m)
- $Test_1$ (pk) = 0/1. Whether there exists sk such that Gen(sk) = pk
- $Test_2$ $^{(pk)}$(c) = 0/1. Whether there exists m such that Test$^{(pk)}$(m) = c

PKE World

PKE World

- Compile out Decryption Oracle following [Mahmoody-Maji-Prabhakaran 2014, TCC]

PKE World

- Compile out Decryption Oracle following [Mahmoody-Maji-Prabhakaran 2014, TCC]
- Both parties pre-emptively decrypt each other’s relevant queries

PKE World

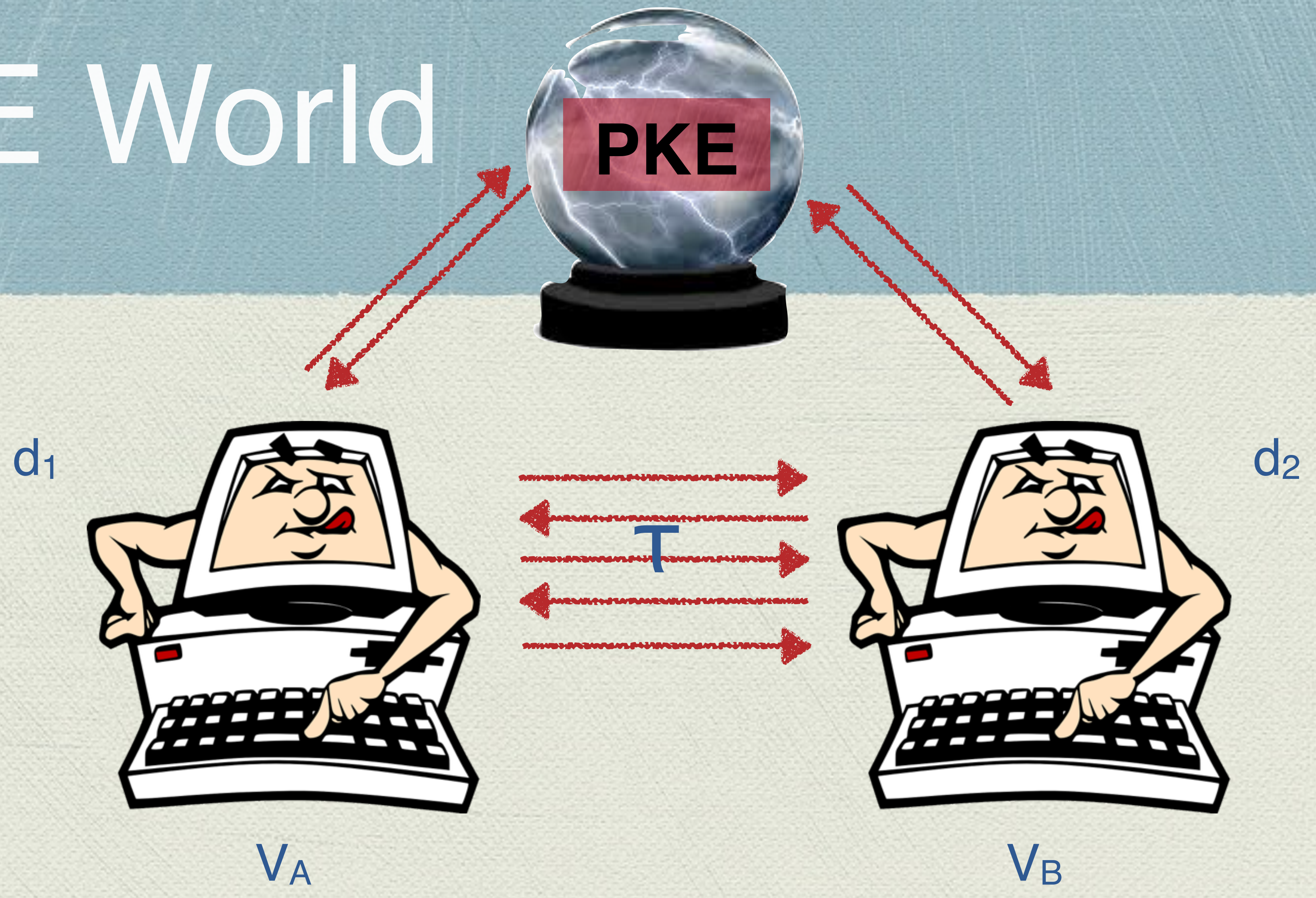
- Compile out Decryption Oracle following [Mahmoody-Maji-Prabhakaran 2014, TCC]
- Both parties pre-emptively decrypt each other's relevant queries
- (Gen, Enc, $Test_1$, $Test_2$) remain

PKE World

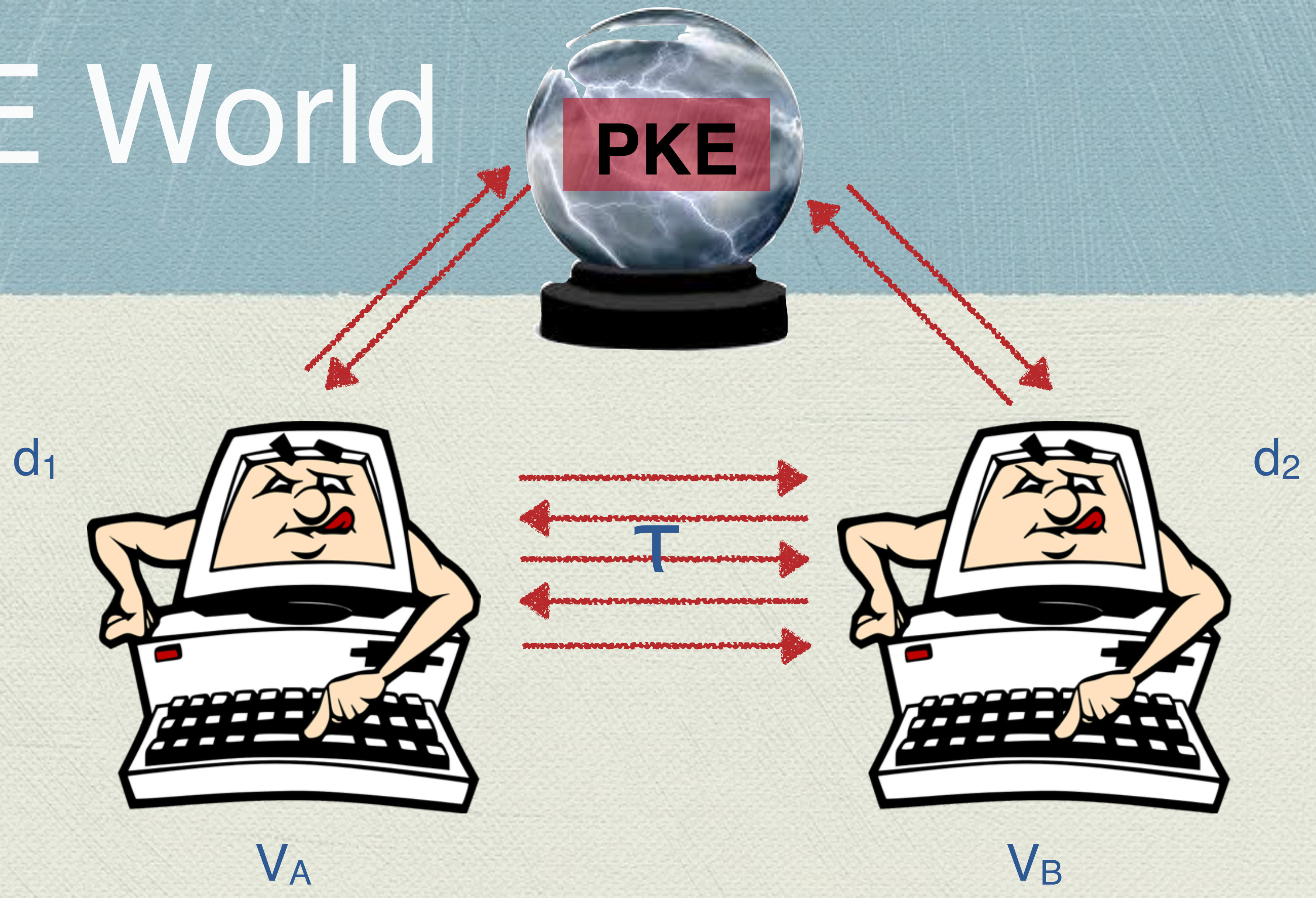
- Compile out Decryption Oracle following [Mahmoody-Maji-Prabhakaran 2014, TCC]
- Both parties pre-emptively decrypt each other's relevant queries
- (Gen, Enc, Test_1, Test_2) remain
- Compiled protocol has slightly lower accuracy

PKE World

PKE World



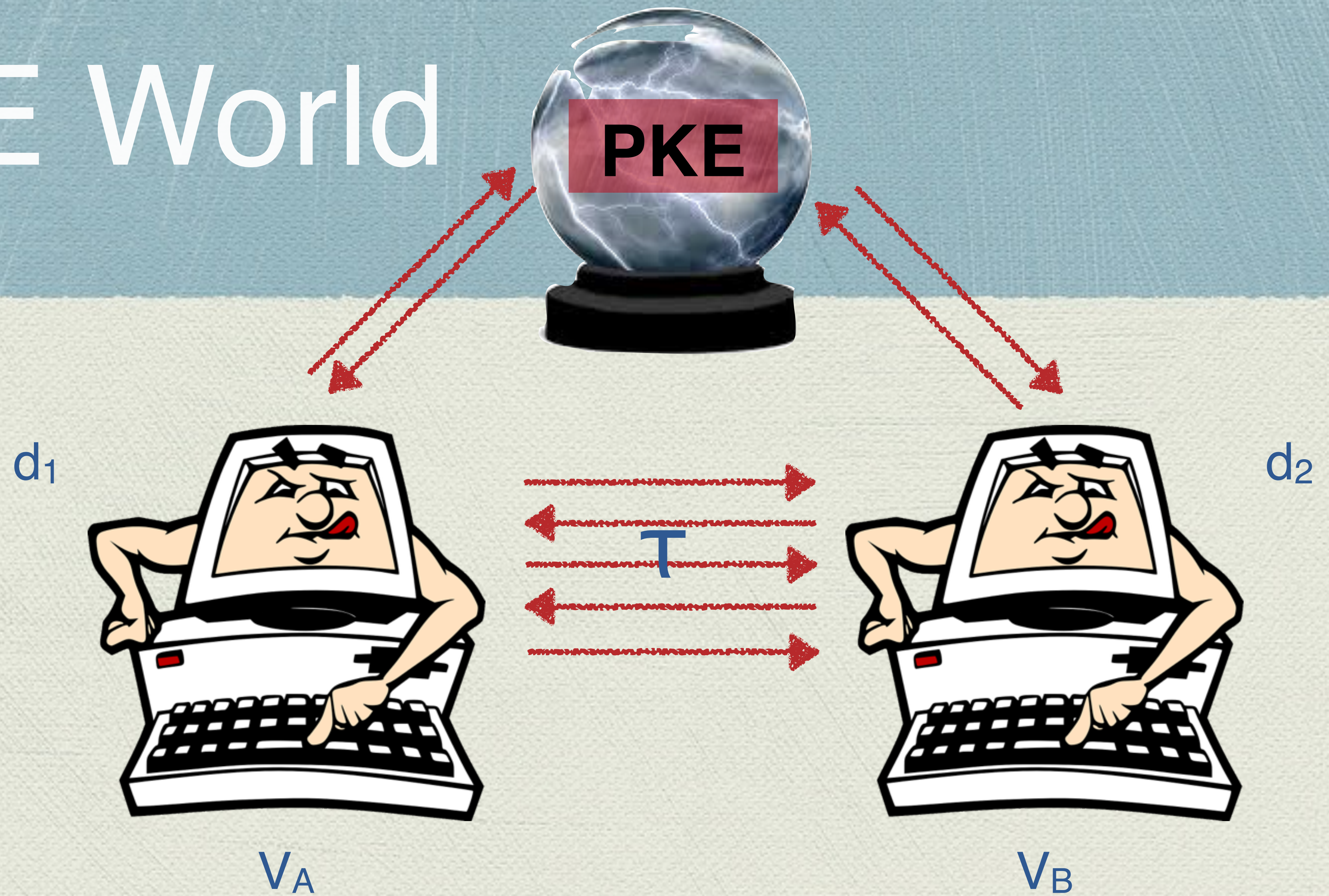
PKE World



- ∀ (ε, α) DP protocol in PKE World

⇒ ∃ (ε, α⁻) DP protocol in (PKE - Dec) World

PKE World

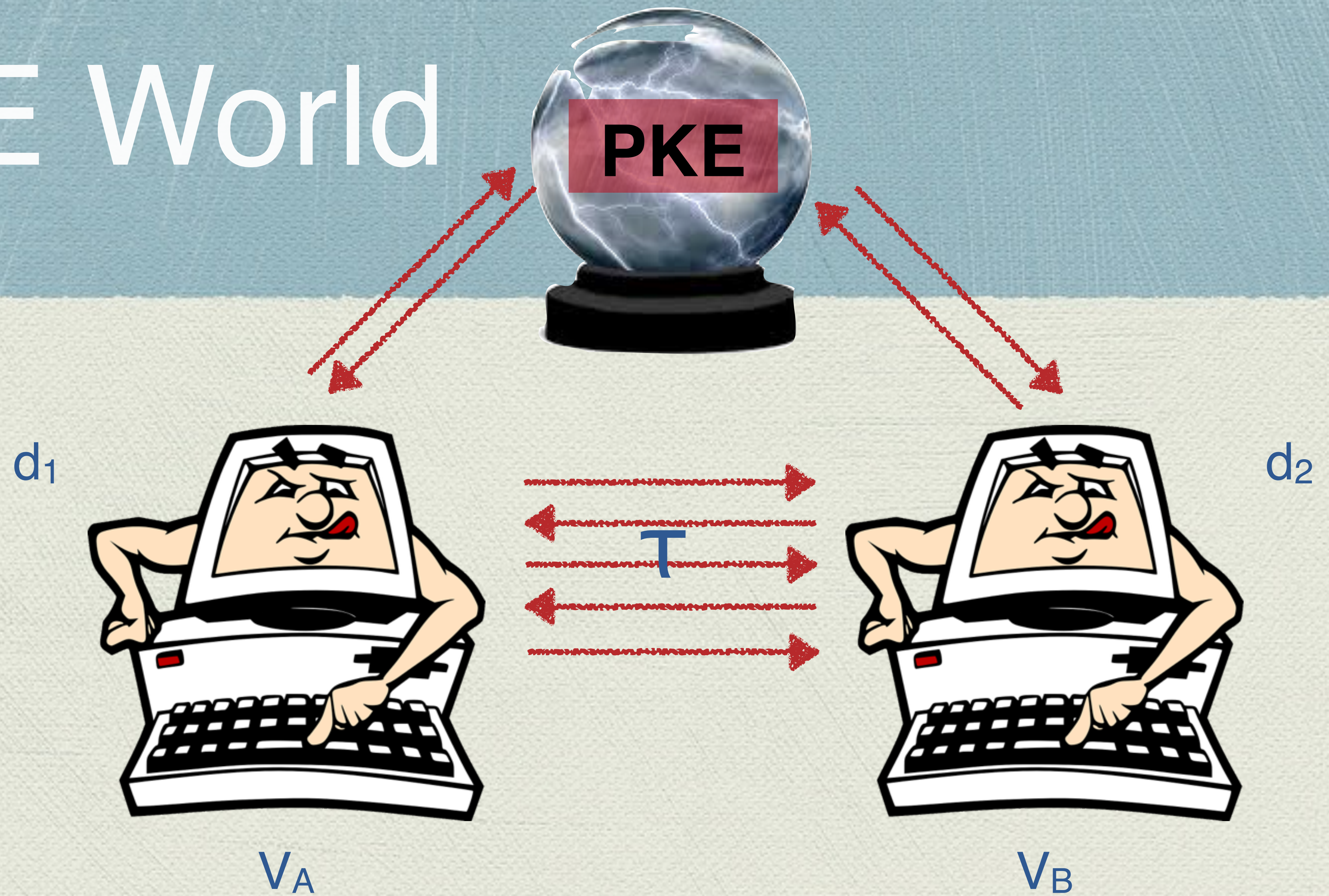


- ∀ (ε, α) DP protocol in PKE World

  [MMP14- TCC]

  ⇒ ∃ (ε, α^{-}) DP protocol in (PKE - Dec) World

PKE World



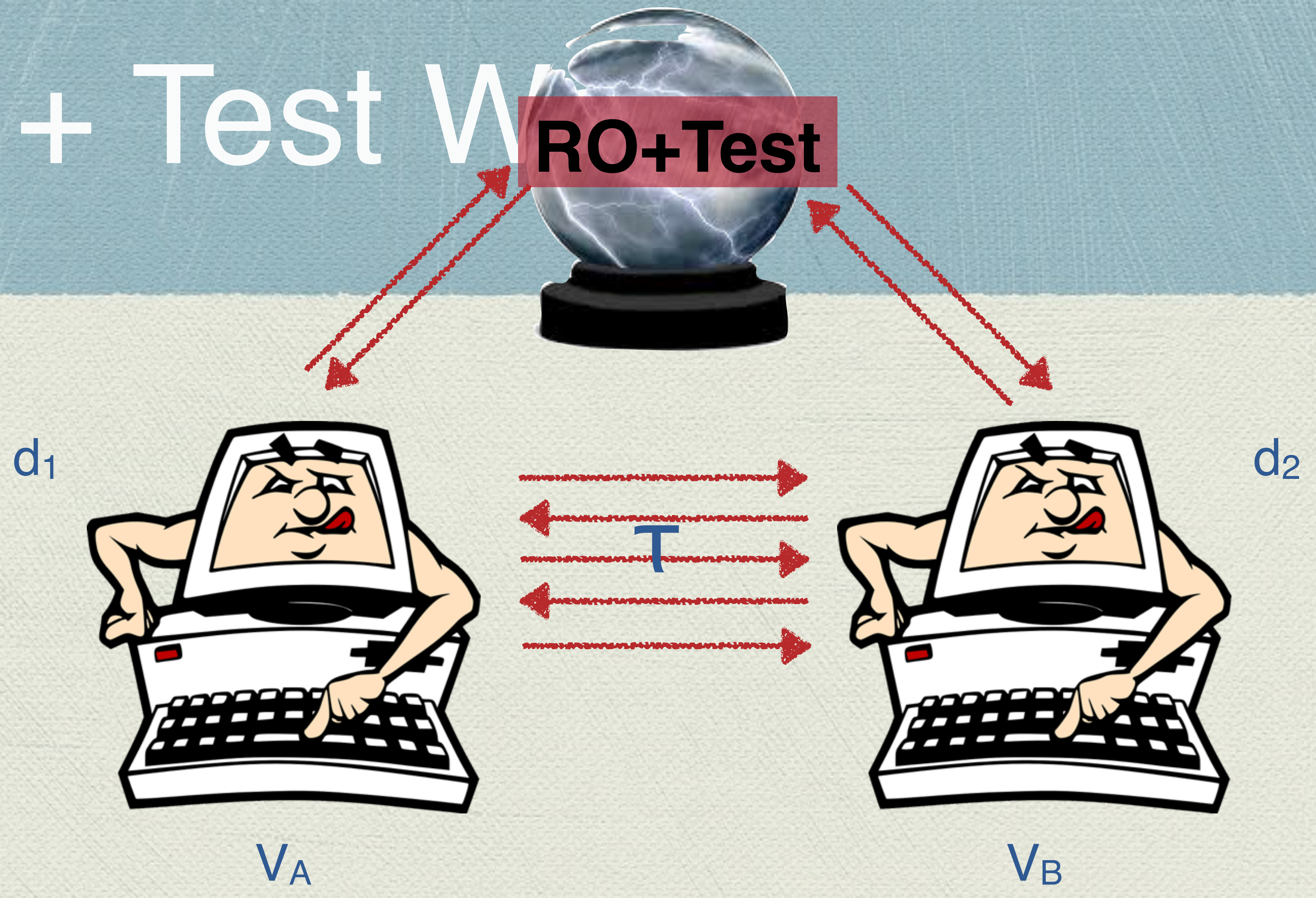
- $\forall$ (ε, α) DP protocol in PKE World

[MMP14- TCC]

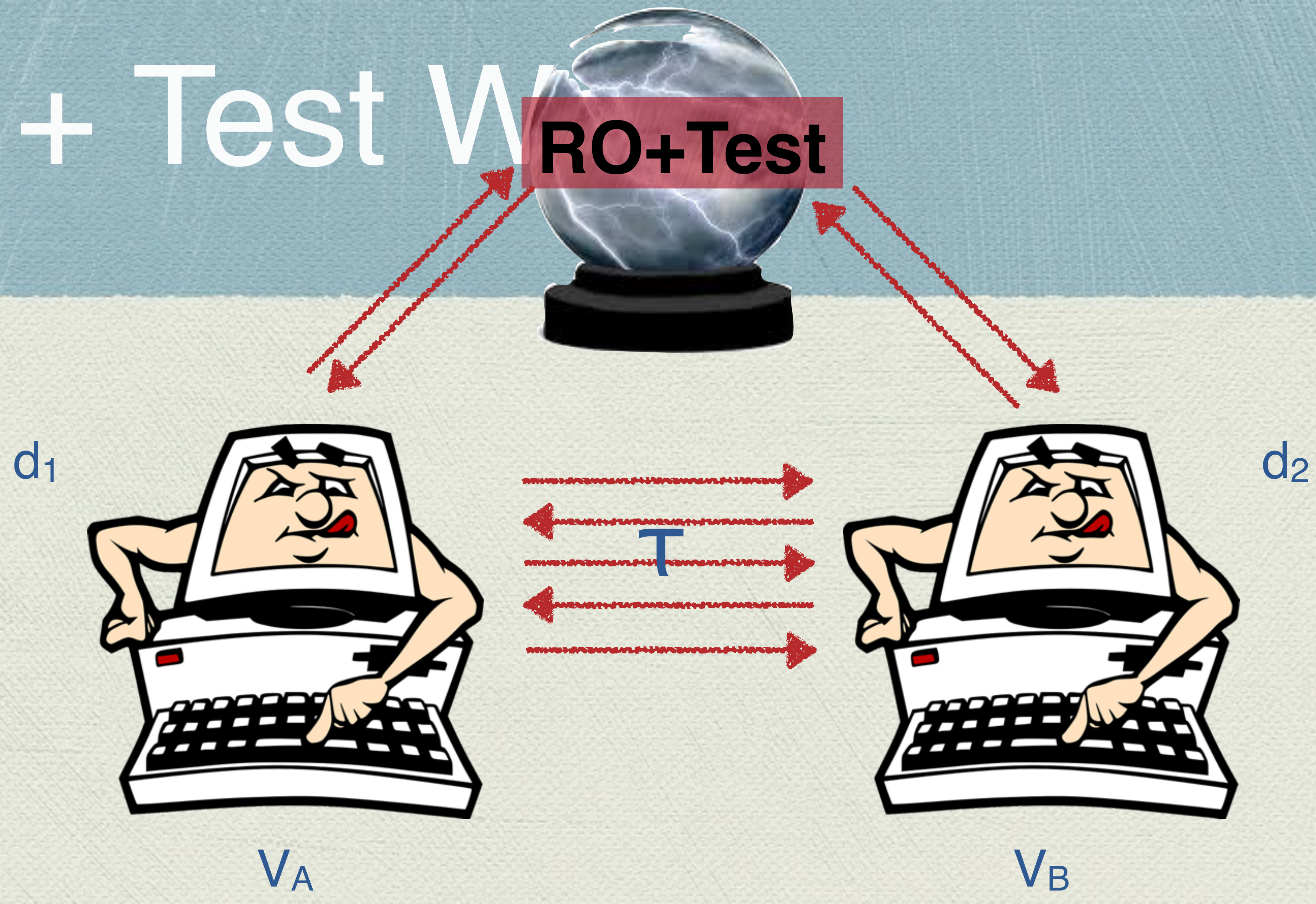
$\Rightarrow$ $\exists$ (ε, α^{-}) DP protocol in (RO + Test) World

RO + Test World

RO + Test W



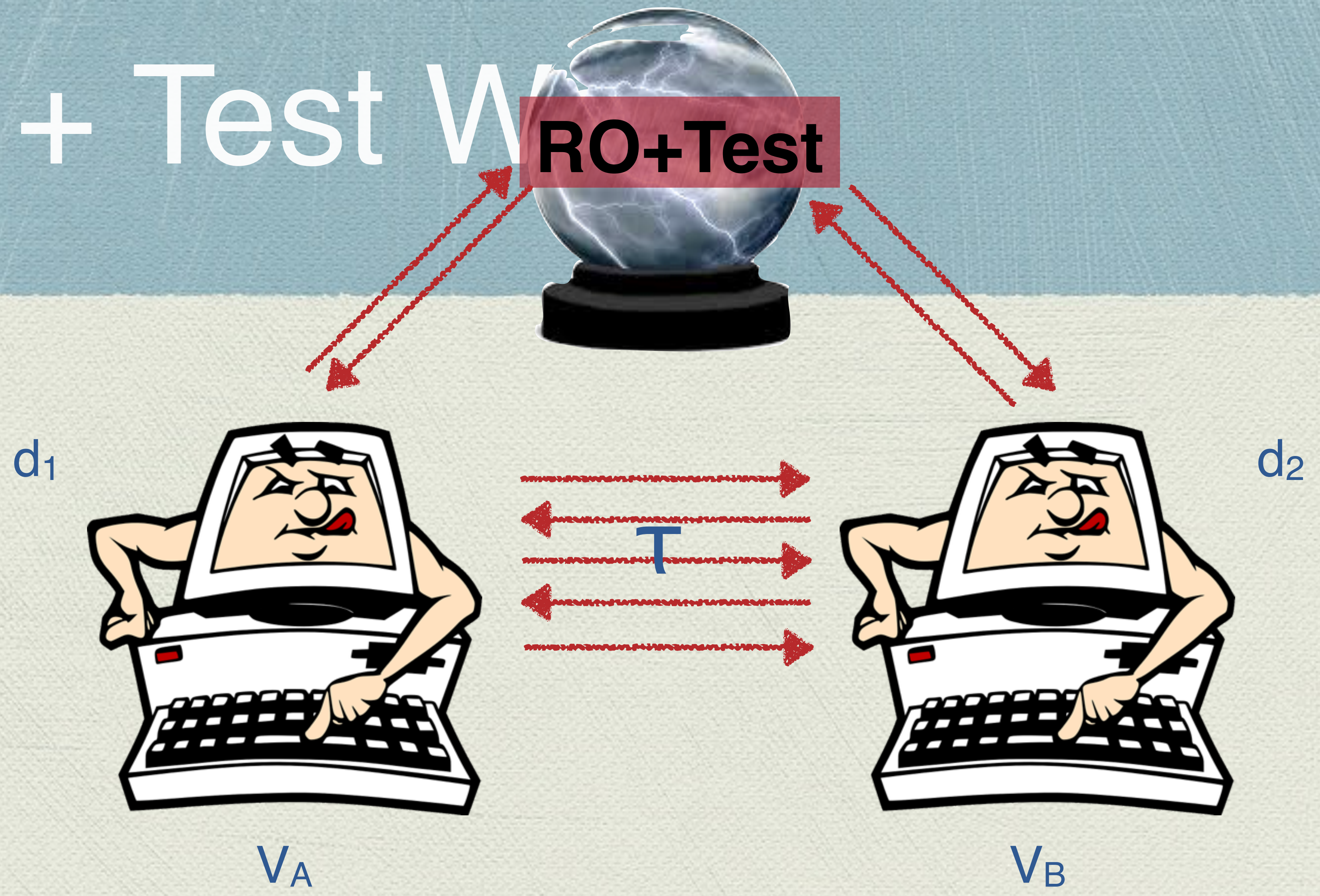
RO + Test W



- $\forall$ (ε, α^{-}) DP protocol in (RO + Test) World

$\Rightarrow \exists$ (ε, α^{--}) DP protocol in (RO) World

RO + Test W



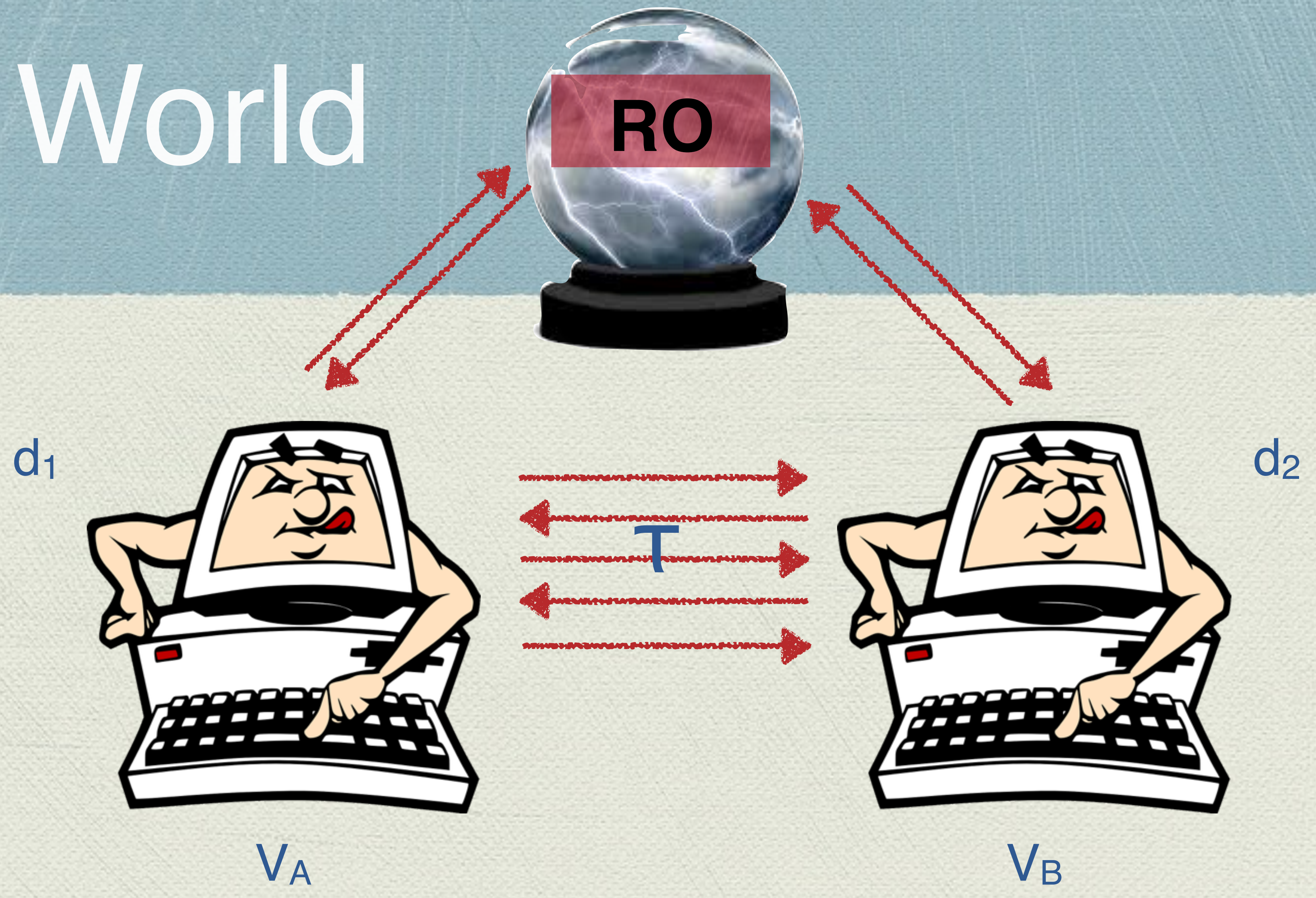
- ∀ (ε, α^{-}) DP protocol in (RO + Test) World

  ⇒ ∃ (ε, α^{--}) DP protocol in (RO) World

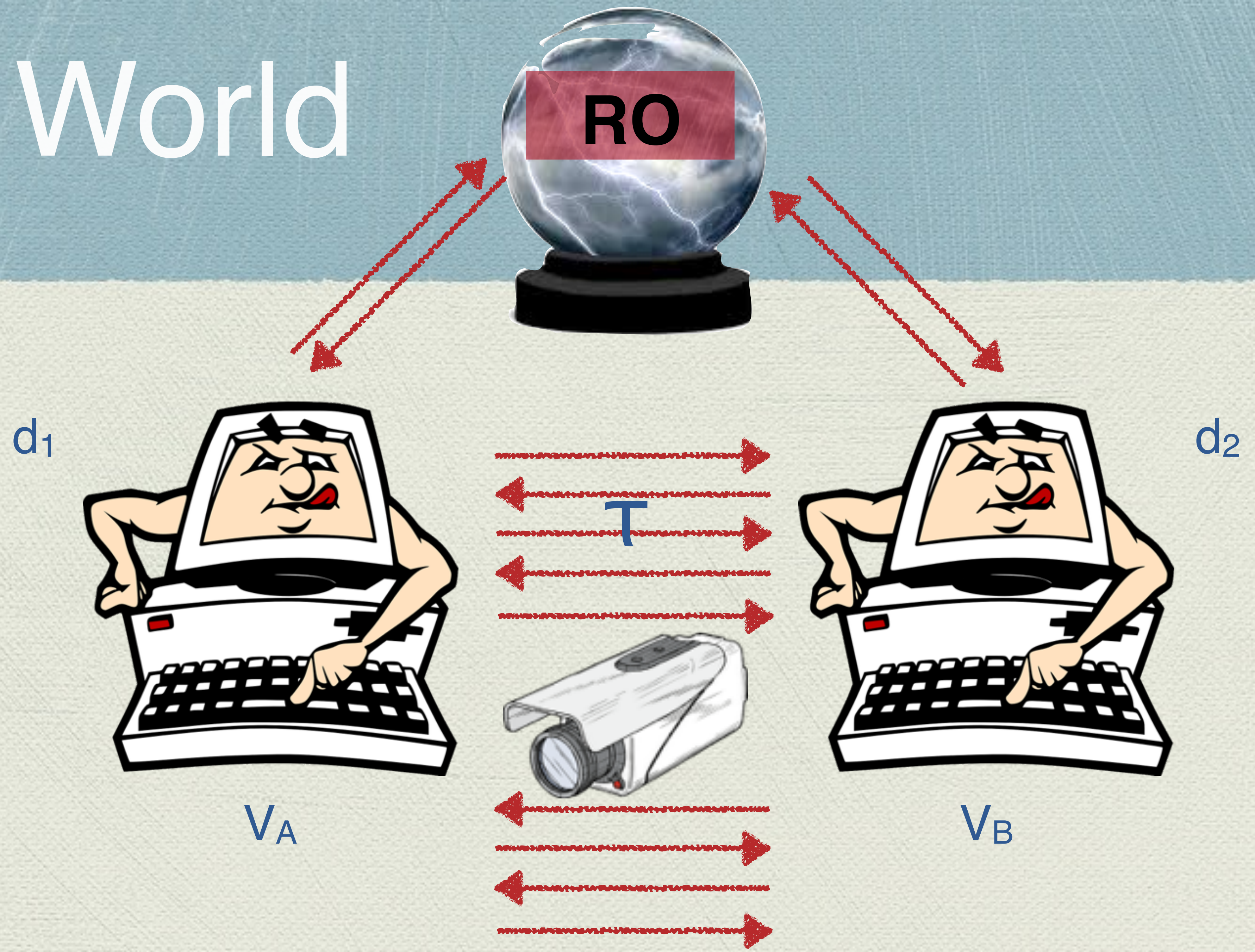
[MMP14- TCC]

RO World

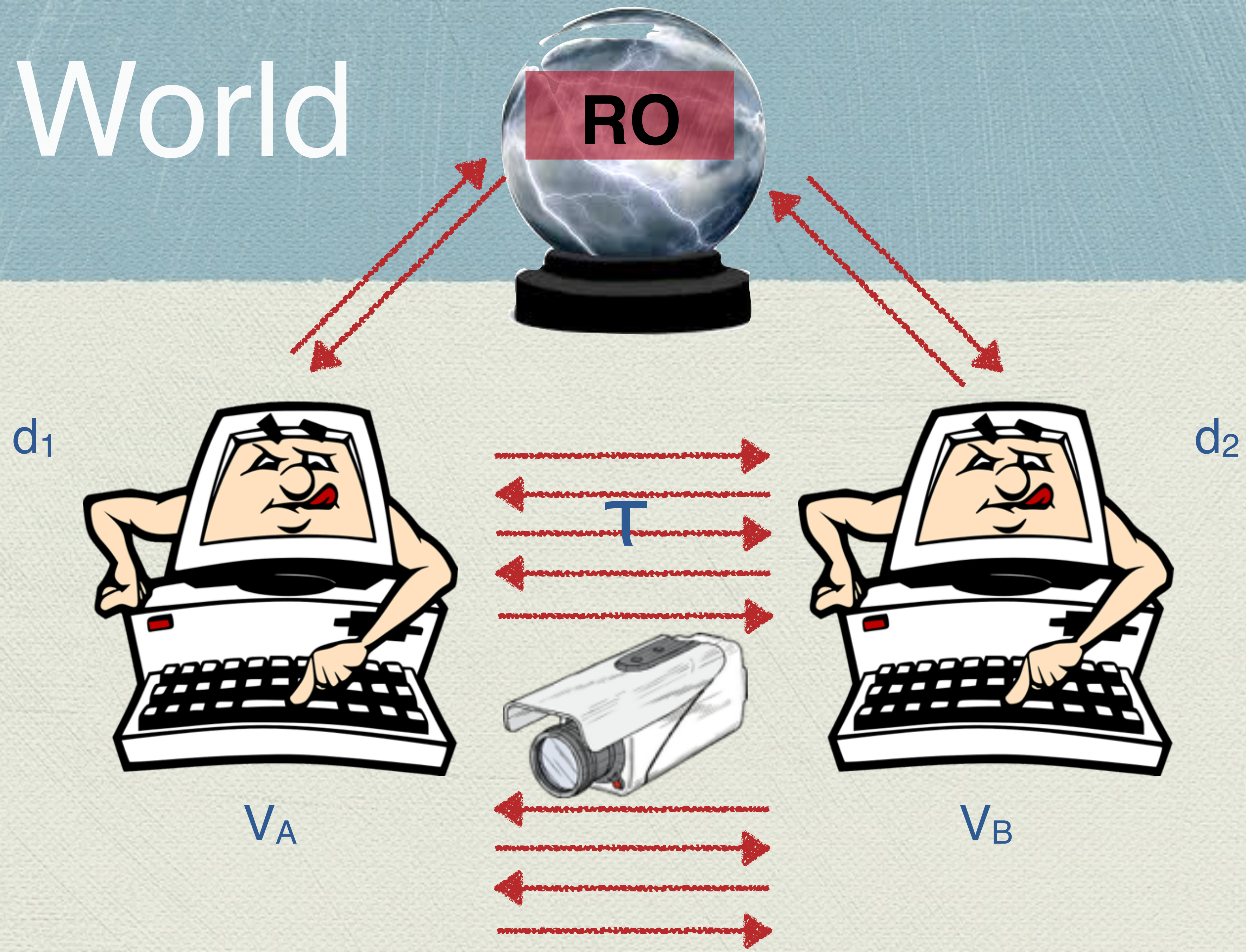
RO World



RO World

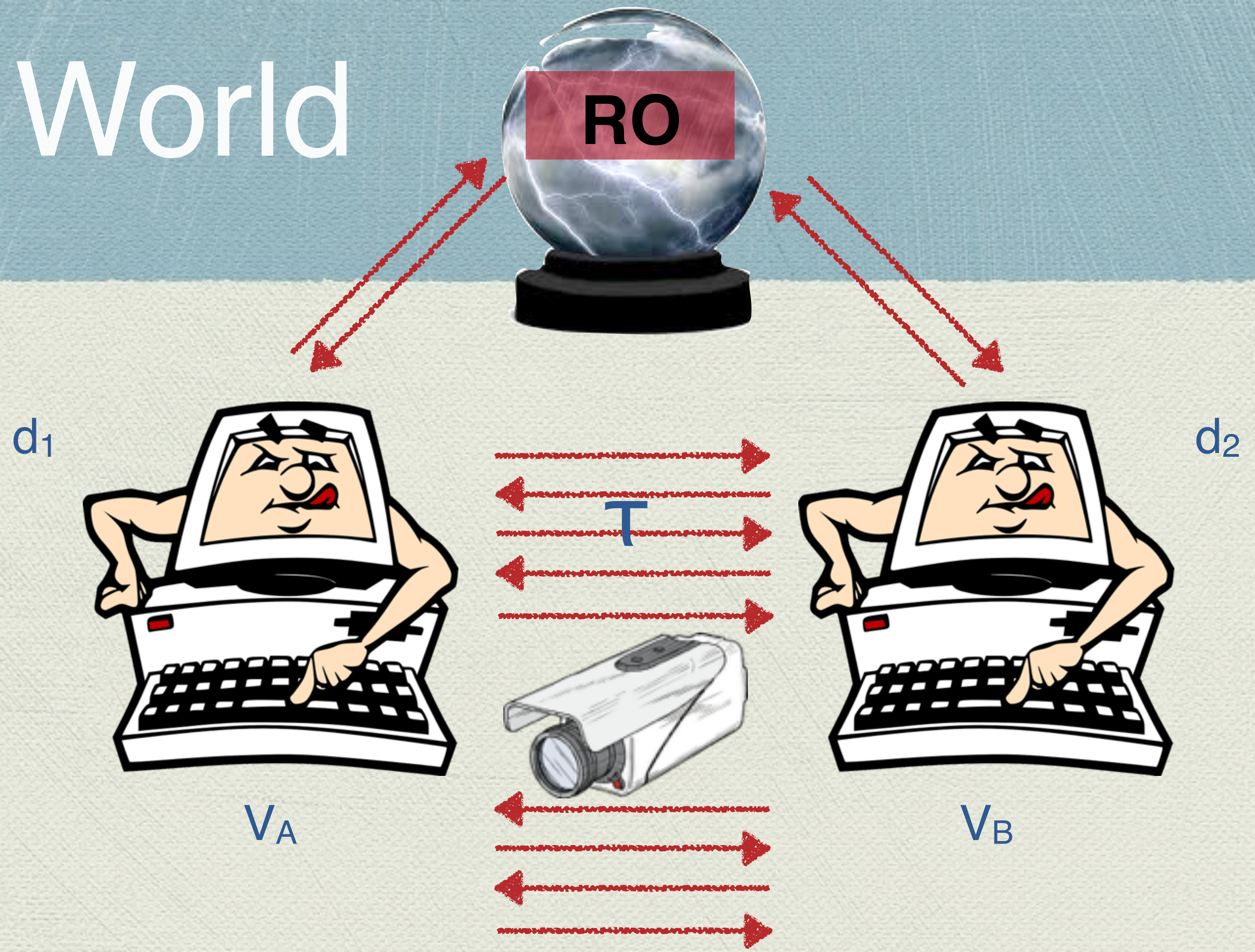


RO World



[MMP14- ITCS]

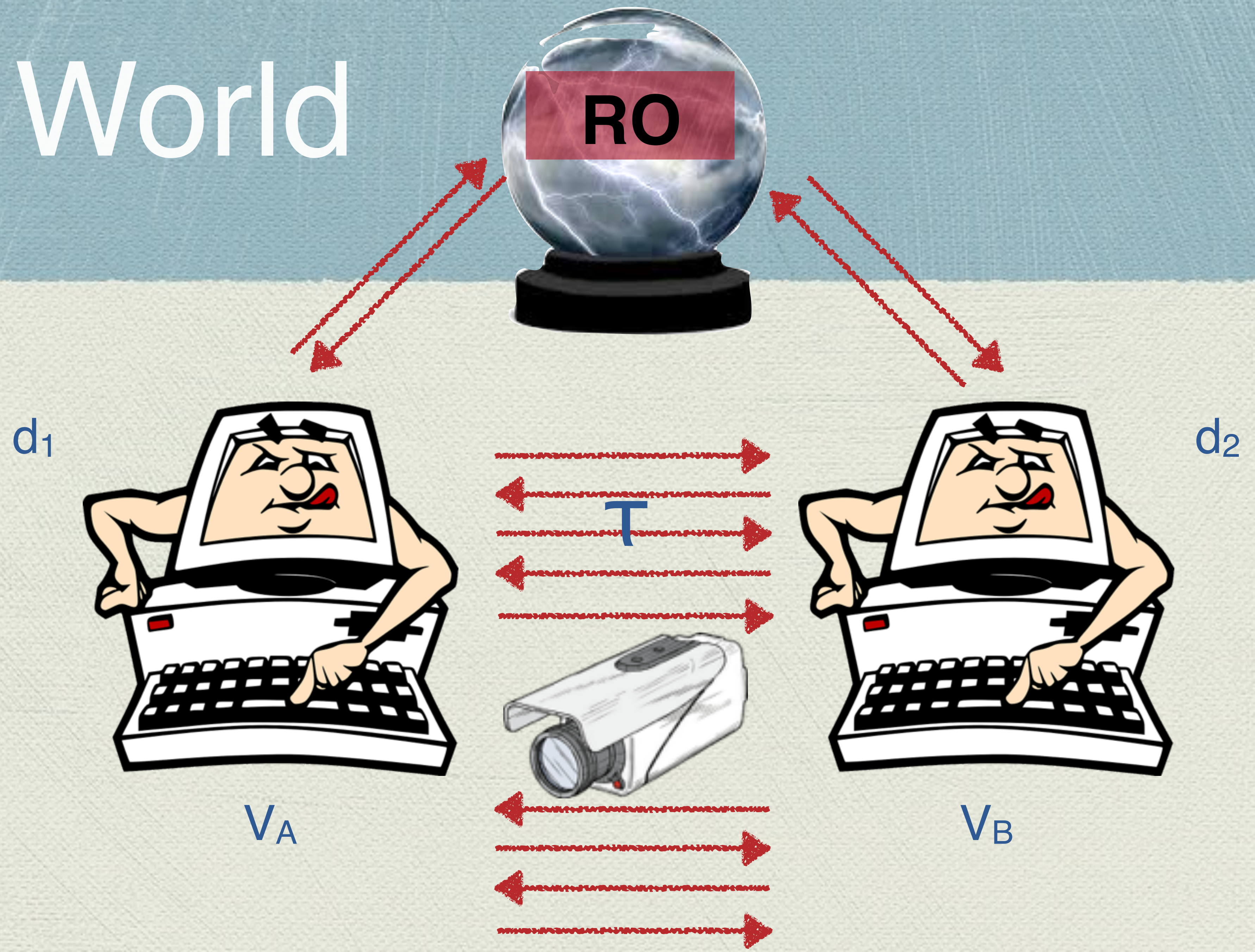
RO World



- Near-independence in RO World [MMP14- ITCS]

$\Rightarrow \alpha_{RO,f,\varepsilon}^{(max)} = (\alpha_{IT,f,\varepsilon}^{(max)})^{+}$

RO World



- Near-independence in RO World [MMP14- ITCS]

$\Rightarrow \alpha_{RO,f,\varepsilon}^{(max)} = (\alpha_{IT,f,\varepsilon}^{(max)})^{+}$ [GMPS13]

Putting it all together,

Putting it all together,

- (ε, α) DP in PKE World ⇒ (ε, α^{--}) DP in RO World

$\Rightarrow \alpha_{PKE,f,\varepsilon}^{(max)} = (\alpha_{RO,f,\varepsilon}^{(max)})^{+}$

Putting it all together,

- (ε, α) DP in PKE World $\Rightarrow$ $(\varepsilon, \alpha^{--})$ DP in RO World

  $\Rightarrow \alpha_{PKE,f,\varepsilon}^{(max)} = (\alpha_{RO,f,\varepsilon}^{(max)})^{+}$

- Near independence in RO world

  $\Rightarrow \alpha_{RO,f,\varepsilon}^{(max)} = (\alpha_{IT,f,\varepsilon}^{(max)})^{+} << \alpha_{f,\varepsilon}^{(opt)}$

Putting it all together,

- (ε, α) DP in PKE World $\Rightarrow$ $(\varepsilon, \alpha^{--})$ DP in RO World

  $\Rightarrow \alpha_{PKE,f,\varepsilon}^{(max)} = (\alpha_{RO,f,\varepsilon}^{(max)})^{+}$

- Near independence in RO world

  $\Rightarrow \alpha_{RO,f,\varepsilon}^{(max)} = (\alpha_{IT,f,\varepsilon}^{(max)})^{+} << \alpha_{f,\varepsilon}^{(opt)}$

- $\Rightarrow \alpha_{PKE,f,\varepsilon}^{(max)} = (\alpha_{IT,f,\varepsilon}^{(max)})^{++} << \alpha_{f,\varepsilon}^{(opt)}$

Conclusion

Technical Recap

- PKE Oracle = (Gen, Enc, Dec, Test_1, Test_2)
- PKE = (RO + Test + Dec) $\approx$ RO + Test $\approx$ RO [MMP14- TCC]
- (Nearly) Independent views in RO world [MMP14- ITCS]
- (Mimic) IT impossibility [GMPS13]

Open Questions

- Does optimally accurate distributed DP ⇒ OT?
  - Use Key Agreement in non black-box way
  - New intermediate computational assumptions equivalent to optimal distributed DP
  - New techniques to obtain OT from properties
    - Similar problems in optimal fair coin tossing

Thank You