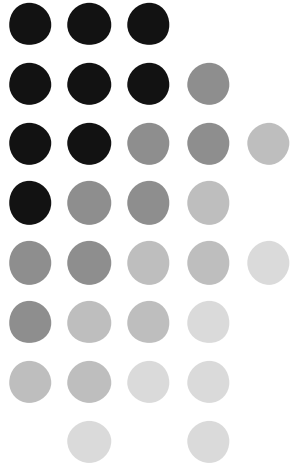
Hidden Markov Model Cryptanalysis

Chris Karlof and David Wagner
University of California-Berkeley

CHES 2003

September 8, 2003

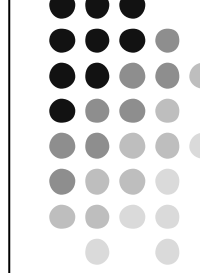


Randomization: A good defense against side-channel attacks?

- Many randomized countermeasures
- Few cryptanalysis techniques
- Our goal:

Evaluate randomization as a defense against side-channel attacks.

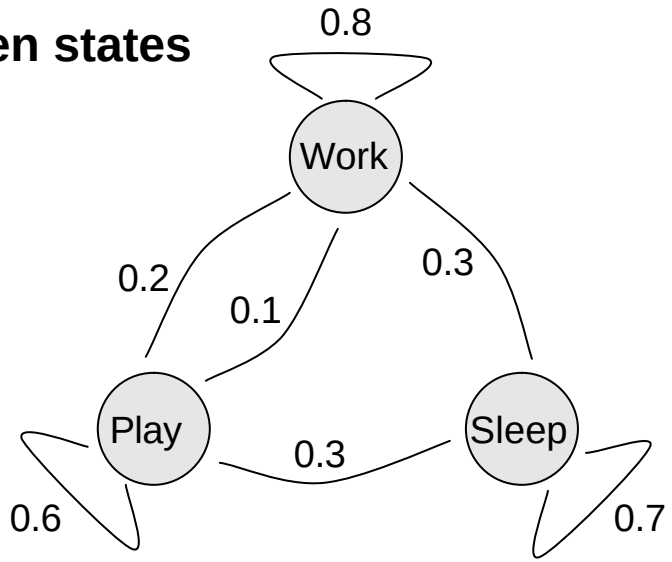
Our contribution: Hidden Markov Model Cryptanalysis



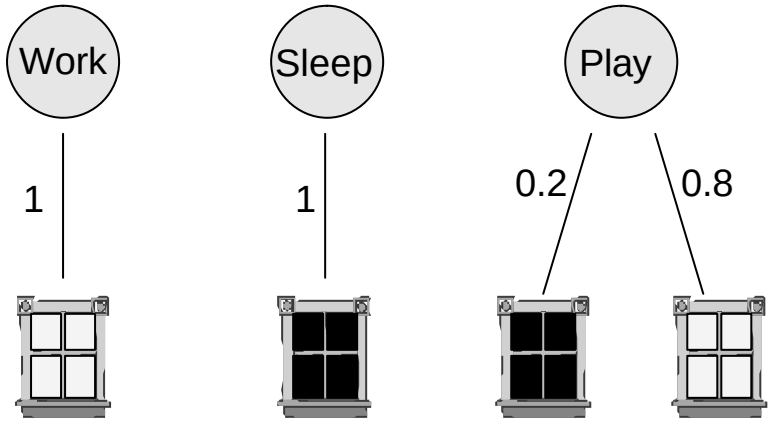
- Introduce *Input Driven Hidden Markov Models*
  - Variation on Hidden Markov Models
  - Model input (i.e., keys)
- Model randomized countermeasures as IDHMM's
- Inference algorithms for IDHMM's
  - Single trace
  - Multiple traces
- Experimental results

Hidden Markov Models

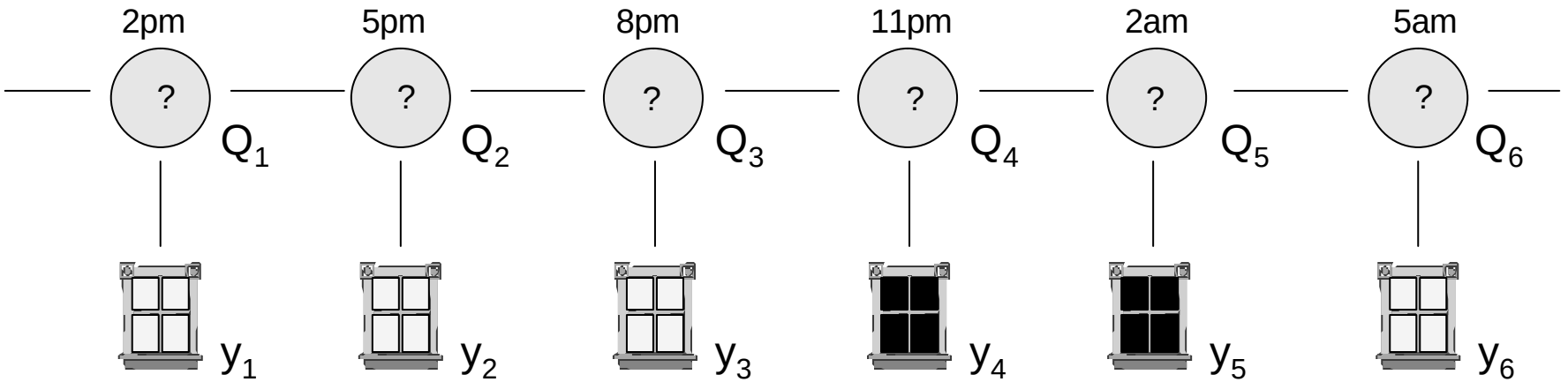
1. Hidden states



2. Observable outputs



3. Inference of hidden states given the observables



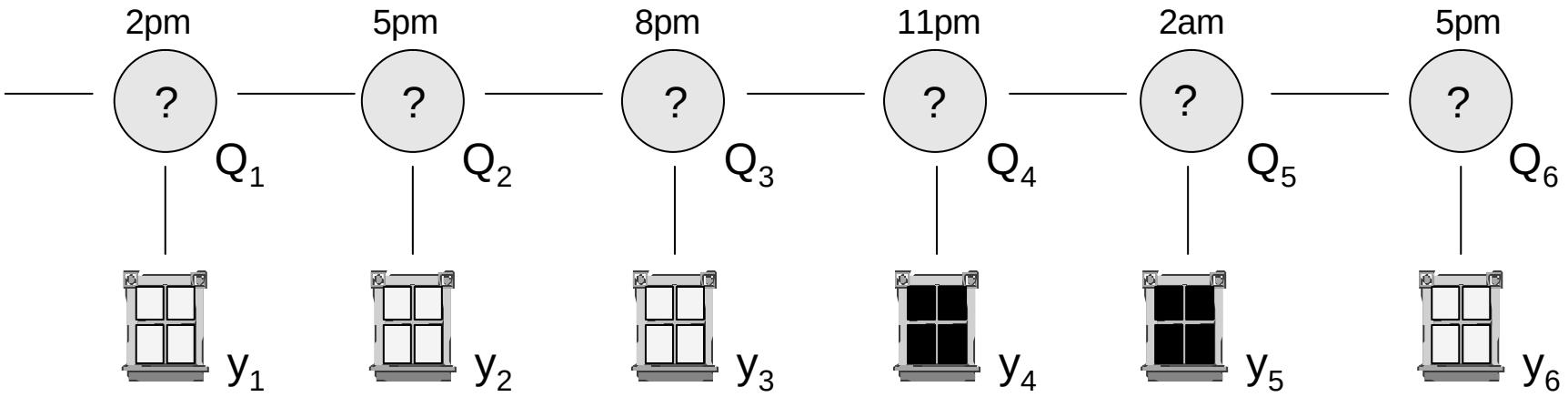
HMM “questions”

- Find most likely sequence of hidden states

$$\underset{Q}{\operatorname{argmax}}\{\Pr[Q \mid y_1, y_2, y_3, y_4, y_5, y_6]\}$$

- Find posterior distribution of hidden states

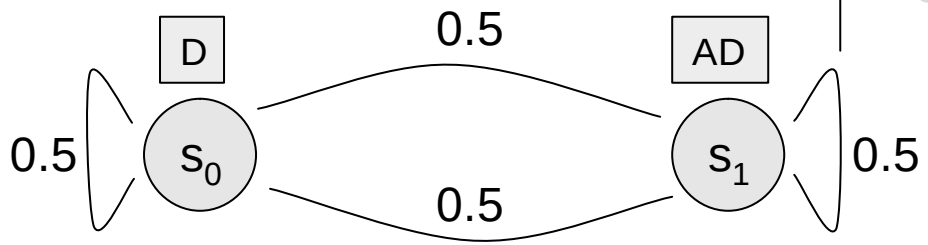
$$\Pr[Q_4 = \text{play} \mid y_1, y_2, y_3, y_4, y_5, y_6]$$

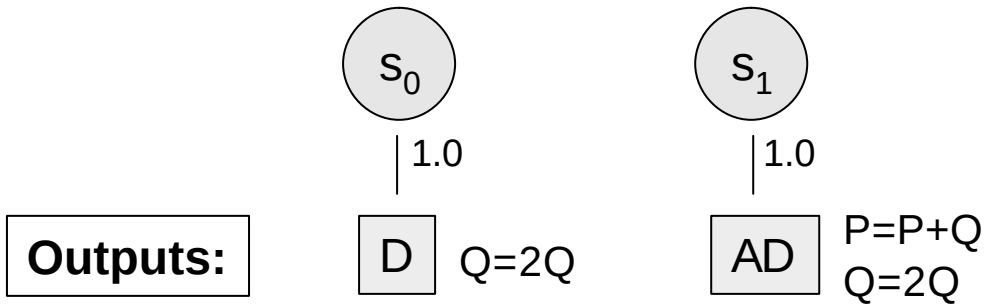


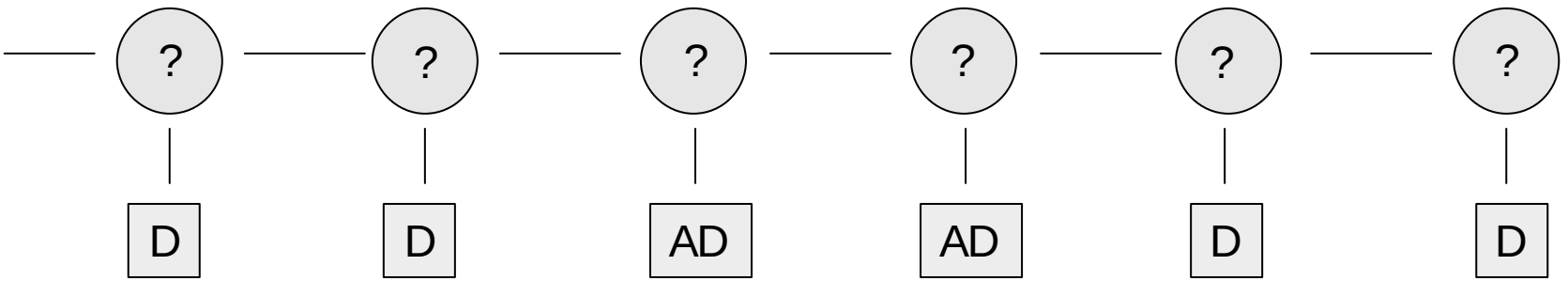
Modeling the binary algorithm for ECC scalar multiplication

Input: k,M Output: kM

```
Q=M
P=0
for i=1 to N
    if (k_i == 1) then P=P+Q
    Q = 2Q
return P
```





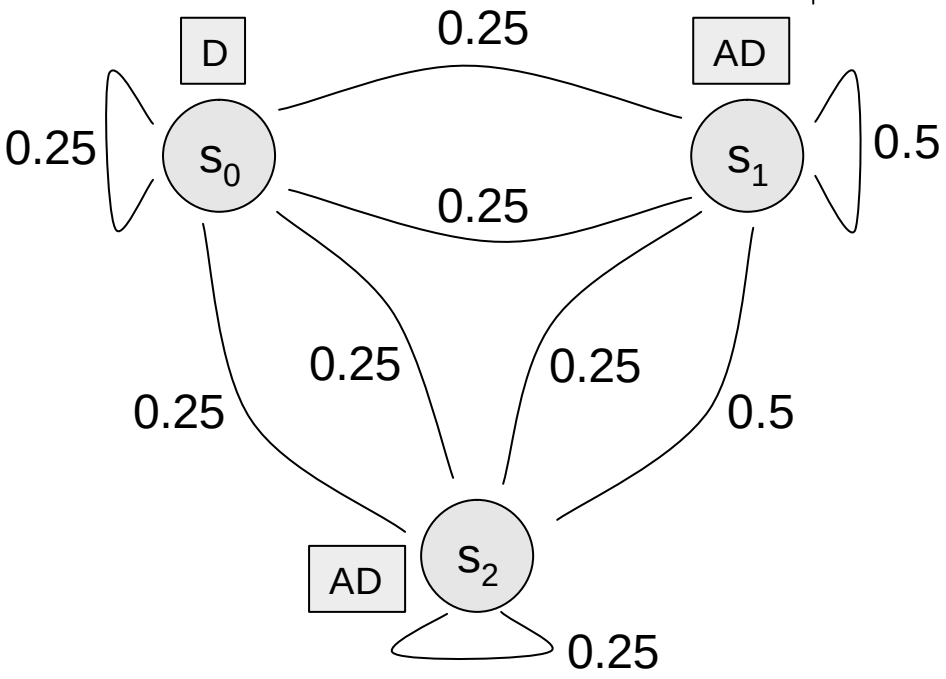


hidden state sequence: $s_0s_0s_1s_1s_0s_0$ → **k = 001100**

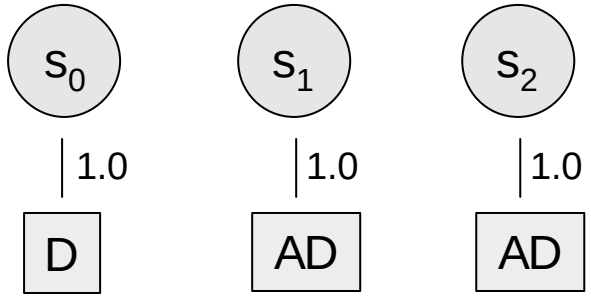
Straw man randomized binary algorithm

Input: k,M Output: kM

```
Q=M
P=0
for i=1 to N
    if (k_i == 1) then P=P+Q
    R = P
    b = rand_bit()
    if (k_i == 0) then
        if (b == 1) then
            R = R+Q
    Q = 2Q
return P
```

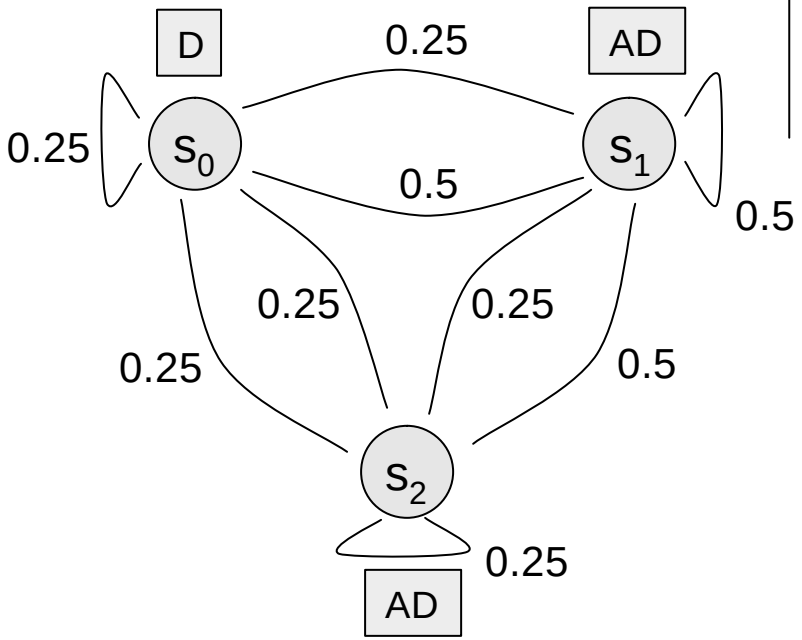


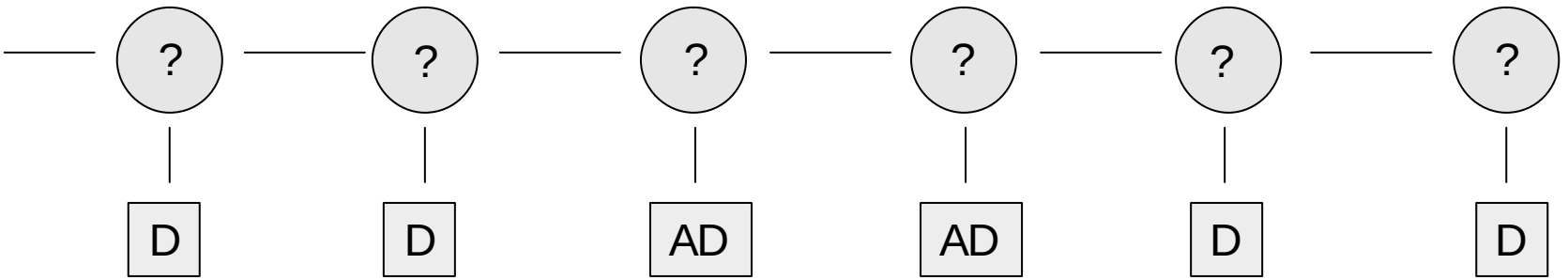
Outputs:



Straw man continued

```
Input: k,M      Output: kM
Q=M
P=0
for i=1 to N
    if (k_i == 1) then P = P+Q
    R = P
    b = rand_bit()
    if (k_0 == 0) then
        if (b == 1) then
            R = R+Q
    Q = 2Q
return P
```





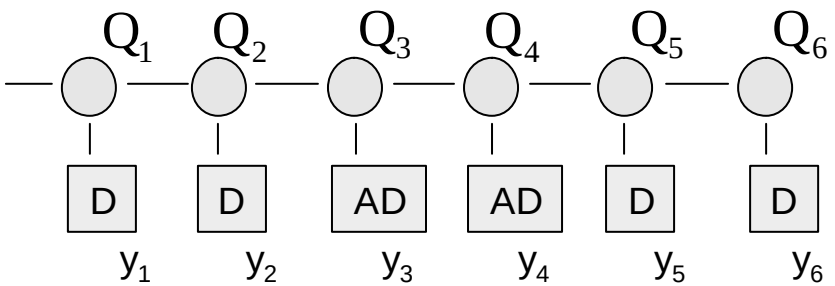
sequences: $s_0s_0s_1s_1s_0s_0$ or $s_0s_0s_1s_2s_0s_0$ or $s_0s_0s_2s_1s_0s_0$ or $s_0s_0s_2s_2s_0s_0$

k:	001100	001000	000100	000000

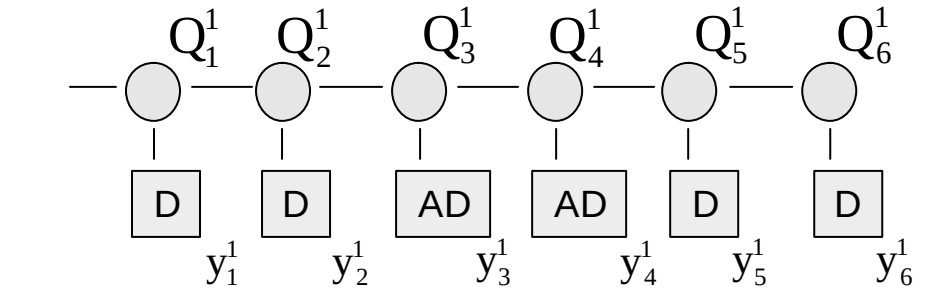
Two issues

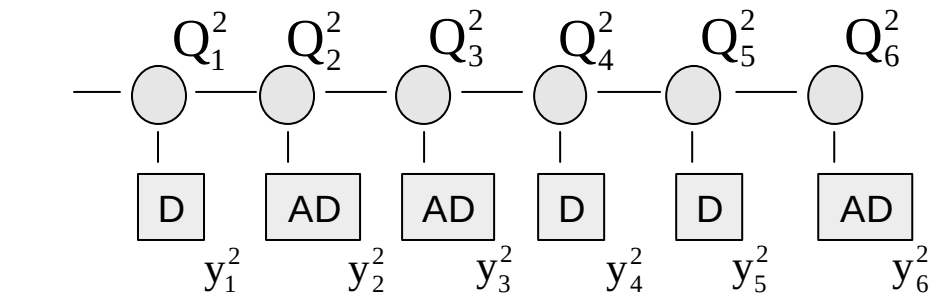
1.

$$\Pr[Q_t = s_i \mid y] \quad \text{vs.} \quad \Pr[K_t = 1 \mid y]$$

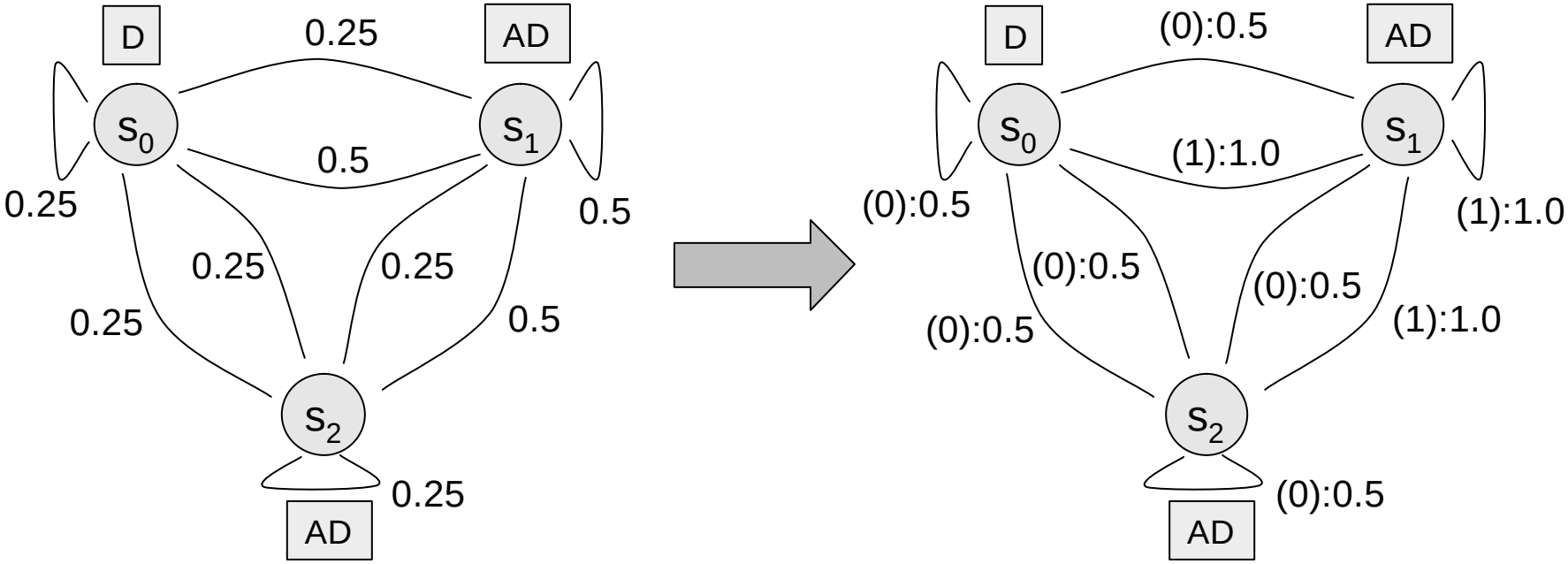


2. **How do we correlate multiple traces?**

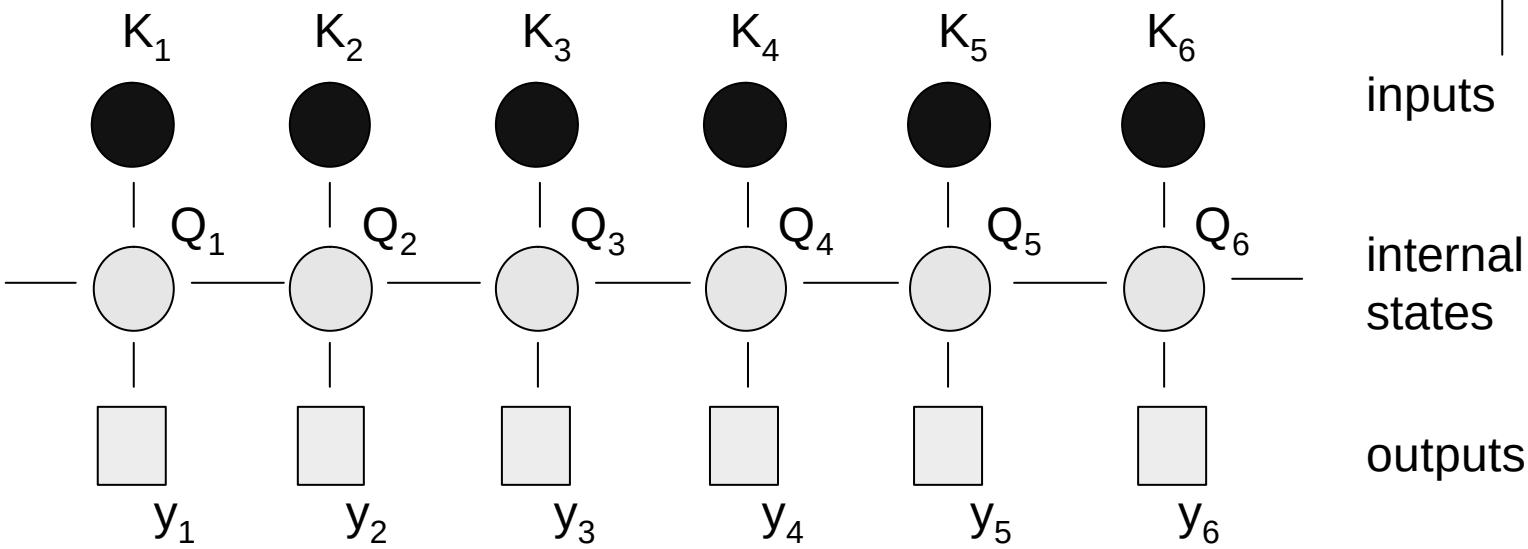




Input Driven HMM's



Input Driven HMM's: Single trace inference



- What is the most likely $K_1K_2\ldots K_n$ given the observed output?
  - Problem: multiple traces.
- Instead, update our "belief" about each key bit → $\Pr[K_t = 1 \mid y] \quad \text{for all } t$
  - We have an $O(|S|^2 \cdot N)$ inference algorithm for a single trace.
    (S is set of internal states and N is the length of the execution)

Multiple trace inference

- Problem: Calculating $\Pr[K_i = 1 \mid y^1, y^2, \ldots, y^L]$ exactly is inefficient.
- Solution: Use a variational inference algorithm.

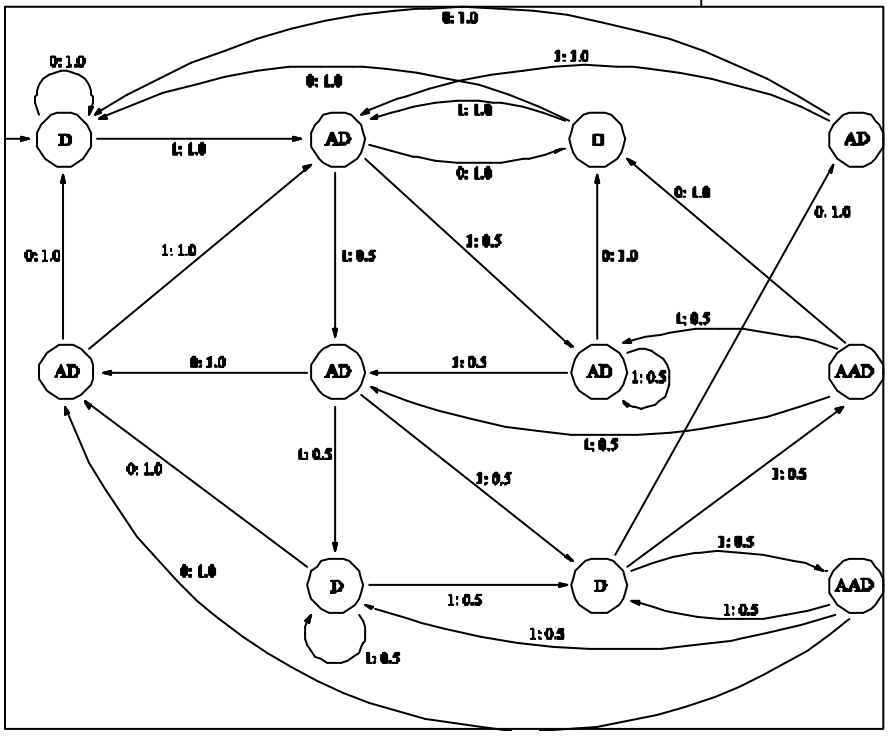
Input: traces $y^1, y^2, \ldots, y^L$
Initialize: $prior_o := \Pr[K_i = 1] = 0.5$ for all i.
for j=1 to L
Calculate: $posterior_j := \Pr[K_i = 1 \mid y^j, prior_{j-1}]$.
Set $prior_j := posterior_j$.
end
if $\Pr[K_i = 1] > 0.5$ then
guess K_i=1, otherwise guess K_i=0.

Results: Oswald-Aigner randomized exponentiation

- Two randomized exponentiation algorithms (OA1, OA2) for ECC
- OA1, OA2 randomize two transformations:

  OA1: $01^{a} \rightarrow 10^{a-1}(-1)$

  OA2: $01^{a}01^{b} \rightarrow 10^{a}(-1)10^{b-1}(-1)$
- Previous attacks: detailed analysis of algorithms
  - Okeya-Sakurai: OA1, ~292 traces, perfect side channel
  - C.D. Walter: OA1,OA2: 2-10 traces, perfect side channel
- Our attack:
  - No special analysis except modeling as an IDHMM
  - Noisy side channel



OA1 as an IDHMM. The set of observables is {D, AD, AAD}.

Results: Oswald-Aigner randomized exponentiation

Observation error

Number of traces used

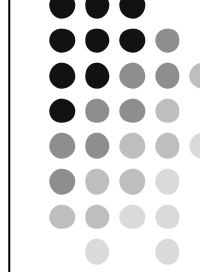
Countermeasure	p_e	1	5	10	25	50	100	500
OA1	0	170	187	192	192	192	192	192
OA1	0.1	157	178	184	185	187	192	192
OA1	0.25	143	163	173	180	182	183	184
OA1	0.4	120	147	159	168	172	173	174
OA2	0	165	188	192	192	192	192	192
OA2	0.1	156	174	184	187	189	192	192
OA2	0.25	135	161	174	177	180	181	182
OA2	0.4	126	146	154	168	171	172	173

Number of key bits correctly recovered (192 bit keys)

Conclusions

- IDHMM’s are useful for modeling randomized side channel measures
  - Model key bits explicitly
  - Handle noisy observations
  - Have efficient inference algorithms
- Future work
  - Tuning attacks
  - Analyze more countermeasures
  - Explore different variational inference techniques
  - More powerful models

Questions?



Extra slides

Comparison to previous work

Attack	Countermeasure	Observation error	Traces to recover key	Work
Okeya-Sakurai	OA1	0	292	minimal
C.D.Walter	OA1,OA2	0	2-10	minimal
IDHMM's	OA1,OA2	0	10	minimal
IDHMM's	OA1,OA2	0	5	2^{38}
IDHMM's	OA1,OA2	0.1	10	2^{38}
IDHMM's	OA1,OA2	0.25	50-100	2^{38}