capture–recapture estimation techniques for reliable RFID tag set resolution

group 09gr855

rasmus melchior jacobsen
karsten fyhn nielsen
{raller,kfyhn}@es.aau.dk

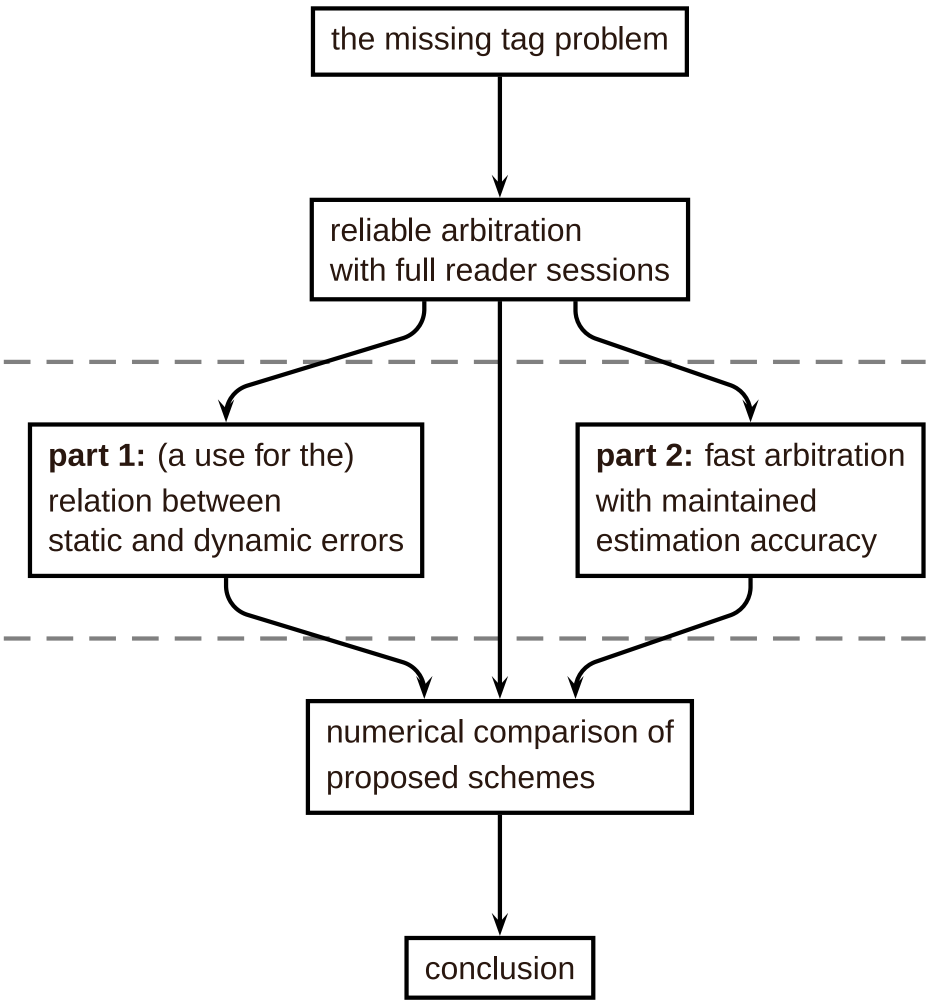
aalborg university, aalborg, denmark

june 18, 2009

introduction to reliable arbitration

time efficient arbitration schemes

talk roundup



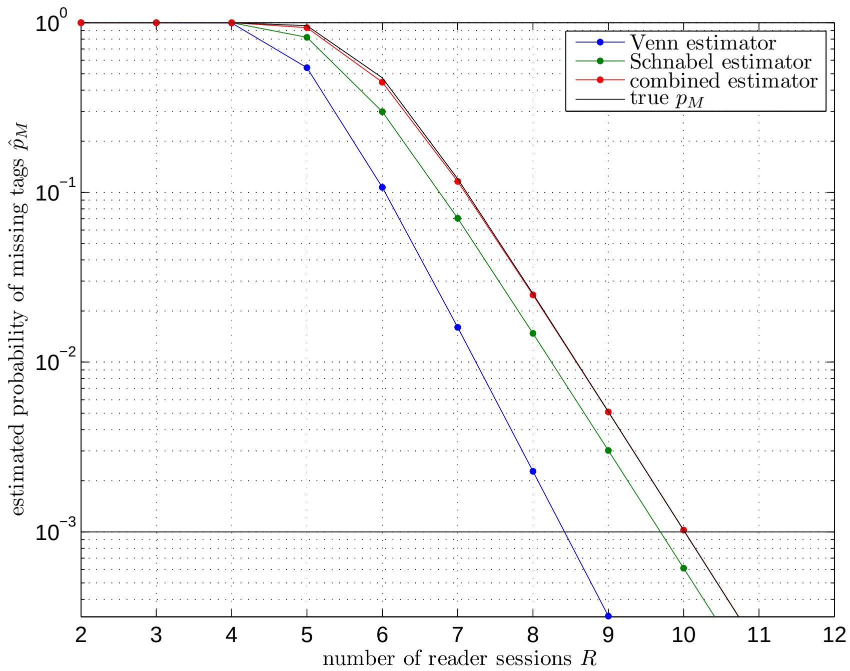
- **problem:** a generic problem in RFID systems is to ensure that the readers can *reliably read a set of tags* in their proximity, that is, ensure with some probability that tags are not missed.
- **why:** the reliability is challenged because there is a non–zero probability that a tag is not read.
    - error–prone wireless link (dynamic errors).
    - tags are in a blind spot (static errors).
    - MAC protocols are originally designed not to account for errors.

- UHF tags, yogurt, portal, success rate: 37.5%
- UHF tags, sunscreen, turning table, success rate: 60.4%
- UHF tags, sunscreen, manual, success rate: 99.6%

test results and image are from: EPCglobal France Lab: RFID for logistic applications – Tests results, test number 28, 57, and 8

1. perform a reader session
2. perform an additional (independent!) reader session
3. relations between sets are used to:
   - estimate the probability of missing a tag in a reader session p,
   - estimate the tag set cardinality N, and
   - (most importantly) estimate the probability of missing tags p_M.
4. if $p_M >$ threshold, go to step 2; exit otherwise

- **previous work:**
  for correlated reader sessions, the sequential decision process may terminate prematurely.
- **continued work:**
  yield the Combined estimator.



(setup: $N = 500$, $\rho = 0.3$)

let $\mathbf{O}$ be an $R \times N$ matrix describing all outcomes (read/not read) of N tags in R reader sessions, and let each element $\mathbf{O}_{ij}$ be defined as

$$\mathbf{O}_{ij} = \begin{cases} 1 & \text{if tag } j \text{ is read in reader session } i, \\ 0 & \text{otherwise.} \end{cases}$$

an example: $\mathbf{O}$ could be found to be, after $R = 4$ reader sessions:

$$\mathbf{O} = \begin{bmatrix} 1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \qquad \mathbf{n} = \sum_{\text{cols}} = \begin{bmatrix} 8 \\ 9 \\ 9 \\ 9 \end{bmatrix} \qquad \mathbf{k} = \begin{bmatrix} 6 \\ 3 \\ 1 \\ 0 \end{bmatrix}$$

$$\sum_{\text{rows}} = \begin{bmatrix} 4 & 3 & 2 & 4 & 4 & 3 & 4 & 4 & 3 & 4 \end{bmatrix}$$

k_1=# of 4s, k_2=# of 3s, etc.

(note that the number of columns, N, is not a priori known, and that the number of rows R is sequentially increased.)

- the Schnabel estimator uses the information in the rows $\mathbf{n}$.
- the Venn estimator uses the information in the columns $\mathbf{k}$.
- the Combined estimator uses both $\mathbf{n}$ and $\mathbf{k}$.
  - numerical results indicate reliable estimates.

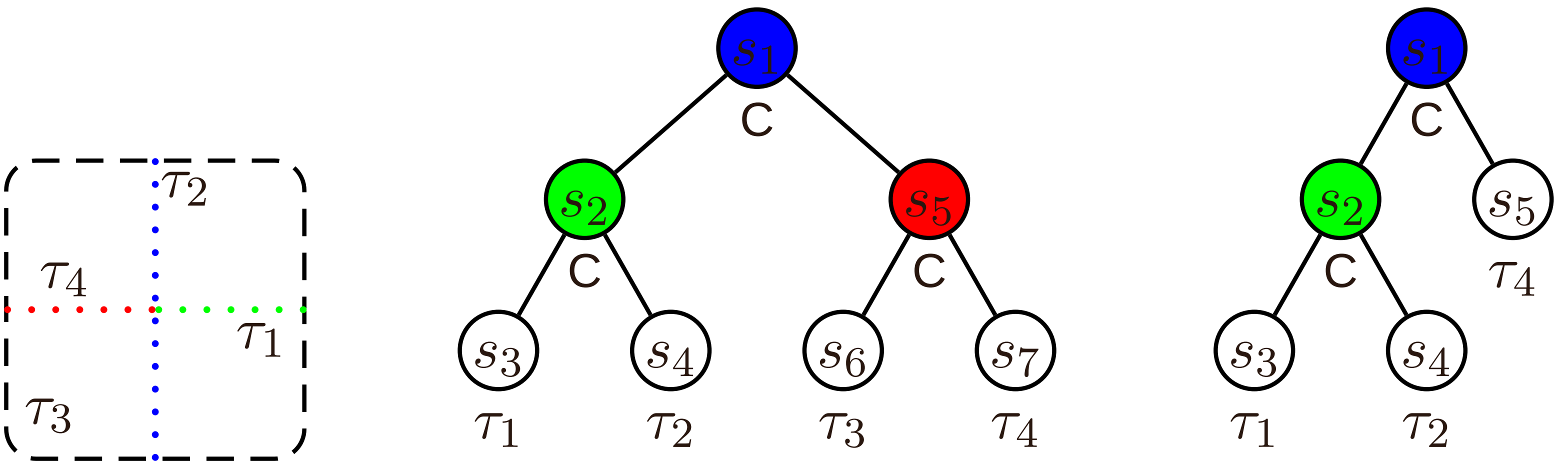
time efficient arbitration schemes, part 1

static and dynamic errors (exemplified by the binary tree algorithm)

- static errors – a tag is missed after a reader session.
- dynamic errors – a tag is missed in a query in a reader session.

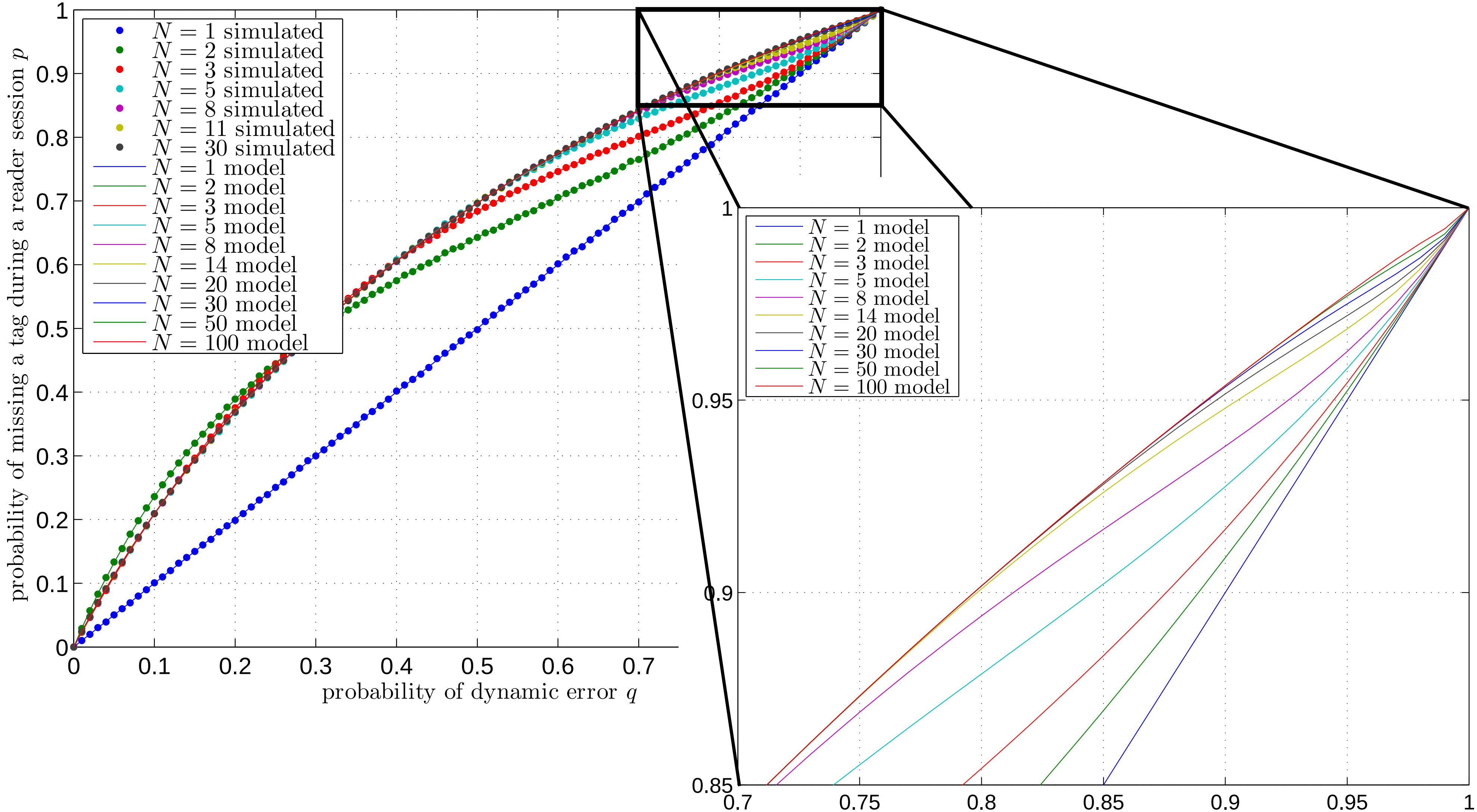
no errors,
$q = 0 \Rightarrow p = 0$

example with errors,
$q = 0.125 \Rightarrow p = 0.25$



- τ_3 is missed because it is in error in s_5.
  - this does not necessarily mean that the tag is in a blind spot.
- however, static errors absorb dynamic errors
  - the estimators work in the presence of dynamic errors.
- = everything is ok, but can we utilize the presence of dynamic errors?

relation between error types for the binary tree arbitration protocol



- **conjecture:** as N increases, the relation becomes independent of N.

we want to find the smallest value of R where (recall: M_L is the expected number of missed tags when L tags participate, i.e. $\hat{p}_L = \frac{M_L}{L}$):

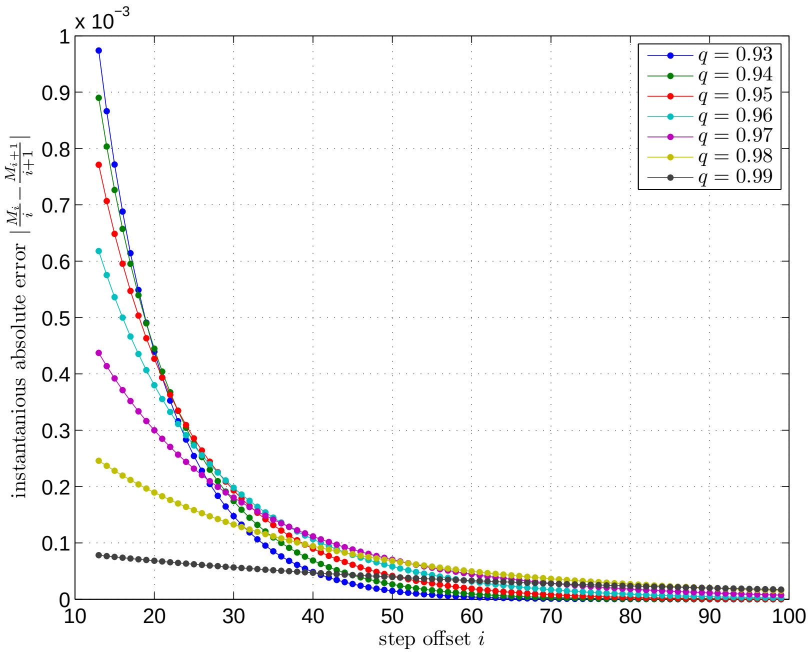
$$\lim_{L\to\infty}\left(\frac{M_R}{R} - \frac{M_L}{L}\right) \leq 10^{-2}$$

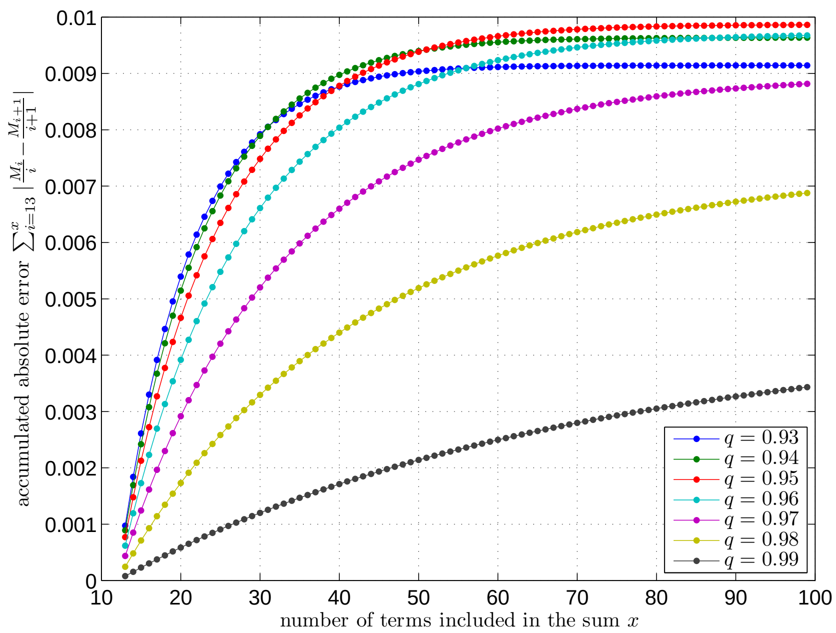
by operating on incremental errors we have:

$$\begin{aligned}
\lim_{L\to\infty}\left(\frac{M_R}{R} - \frac{M_L}{L}\right) &= \sum_{i=R}^{\infty}\left(\frac{M_i}{i} - \frac{M_{i+1}}{i+1}\right) \\
&\leq \sum_{i=R}^{\infty}\left|\frac{M_i}{i} - \frac{M_{i+1}}{i+1}\right| \\
&= \sum_{i=R}^{x}\left|\frac{M_i}{i} - \frac{M_{i+1}}{i+1}\right| + \sum_{i=x+1}^{\infty}\left|\frac{M_i}{i} - \frac{M_{i+1}}{i+1}\right| \leq 10^{-2}
\end{aligned}$$

note that the error is smaller for all values of R greater than the smallest R satisfying the requirement.

$$10^{-2} \geq \sum_{i=R}^{x} \left| \frac{M_i}{i} - \frac{M_{i+1}}{i+1} \right| + \sum_{i=x+1}^{\infty} \left| \frac{M_i}{i} - \frac{M_{i+1}}{i+1} \right| \approx \sum_{i=R}^{x} \left| \frac{M_i}{i} - \frac{M_{i+1}}{i+1} \right|$$

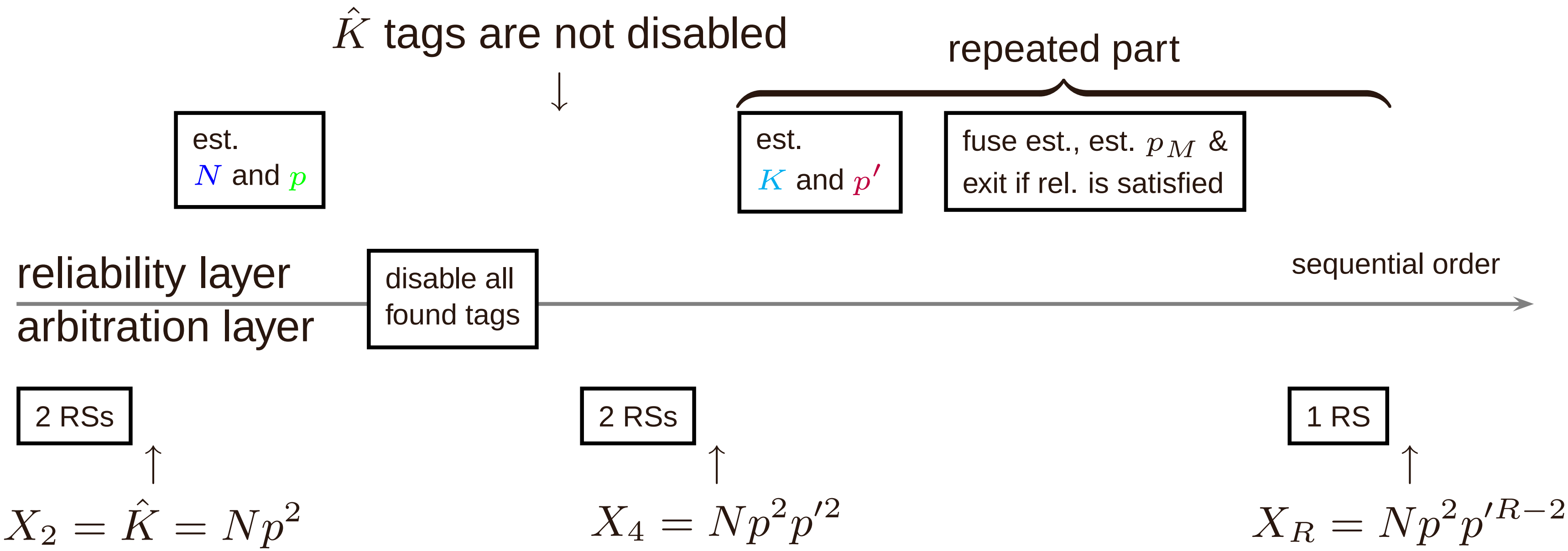




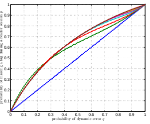
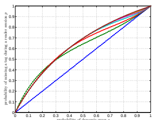
(numerical experiments indicate that the critical region lies in the region $q \in [0.93, 0.97]$.)

- **conjecture:** for $N \geq 13$, the absolute error is less than 10^{-2}.

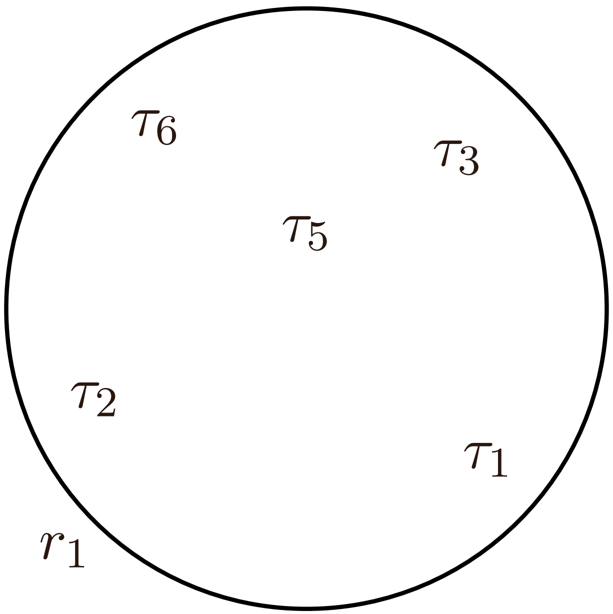
let X_i denote the exp. number of missed tags after i reader sessions

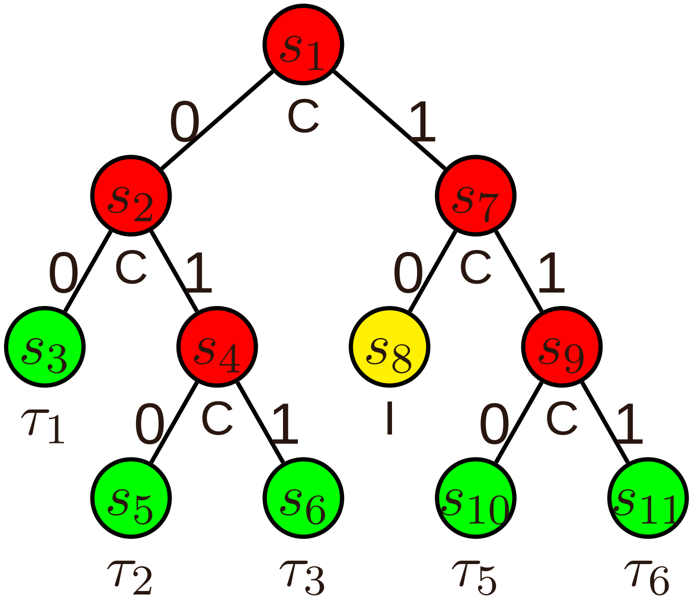


$p \neq p'$ as $K \geq 13$ can't be assumed; q is assumed to remain the same.

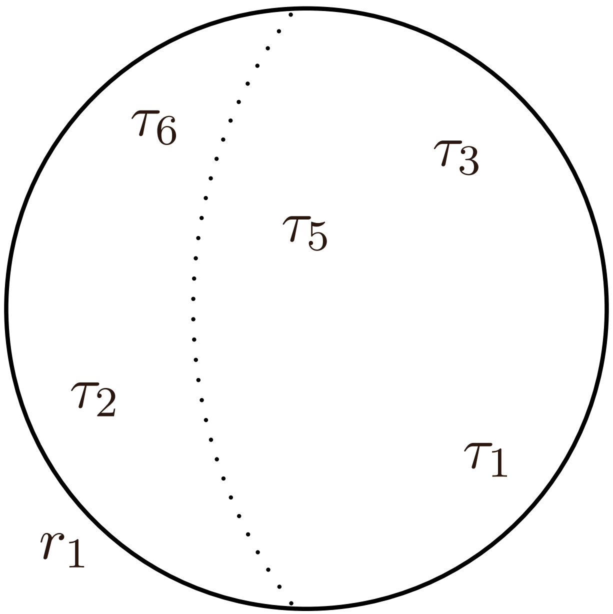
1. find $\hat{K}_{r_{1,2}} = \hat{N}_{r_{1,2}} \hat{p}_{r_{1,2}}{}^2$.
2. fuse $\hat{K}_{r_{1,2}}$ and $\hat{K}_{r_{3,4,\ldots,R}}$ into $\hat{K}$.
3. use $\hat{N}_{r_{1,2}}$ ($\hat{K}$), $\hat{p}_{r_{1,2}}$ ($\hat{p}'_{r_{3,4,\ldots,R}}$) and  to find $\hat{q}_{r_{1,2}}$ ($\hat{q}_{r_{3,4,\ldots,R}}$).
4. fuse $\hat{q}_{r_{1,2}}$ and $\hat{q}_{r_{3,4,\ldots,R}}$ into $\hat{q}$.
5. use $\hat{K}$, $\hat{q}$ and  to find fused $\hat{p}'$ and $\hat{p}_M = 1 - (1 - \hat{p}'^{R-2})^{\hat{K}}$.

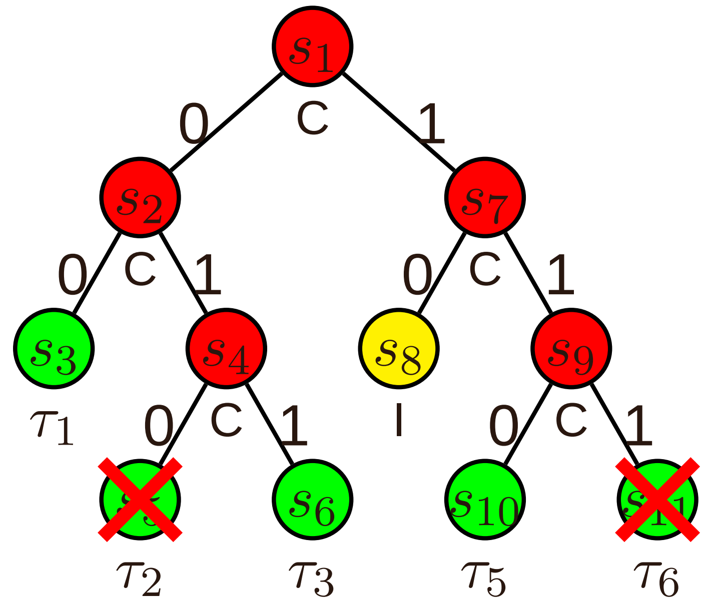
time efficient arbitration schemes, part 2



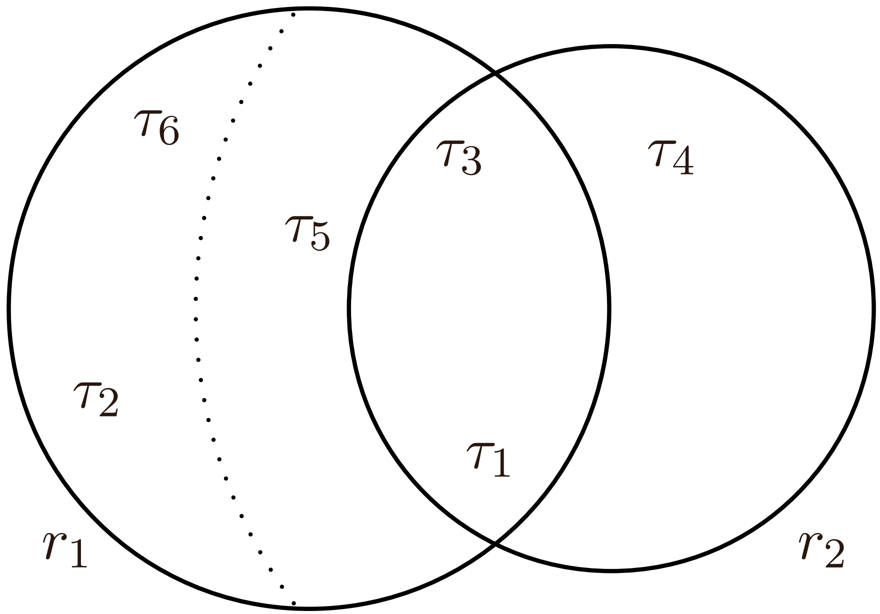


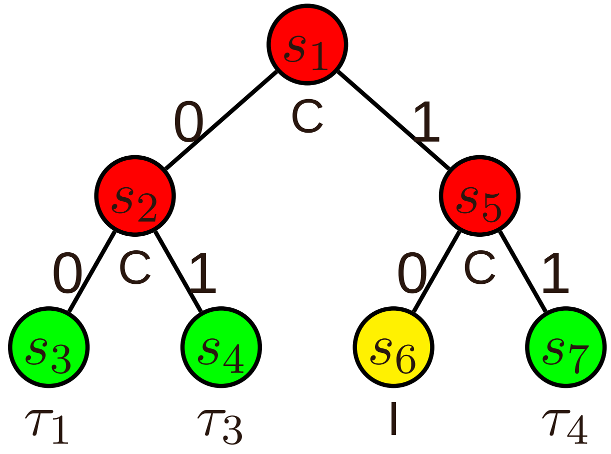
- capture–recapture techniques are based on the relationsship between the initially captured and the recaptured samples.
- more samples means more accurate estimation, but the accuracy comes at a price.



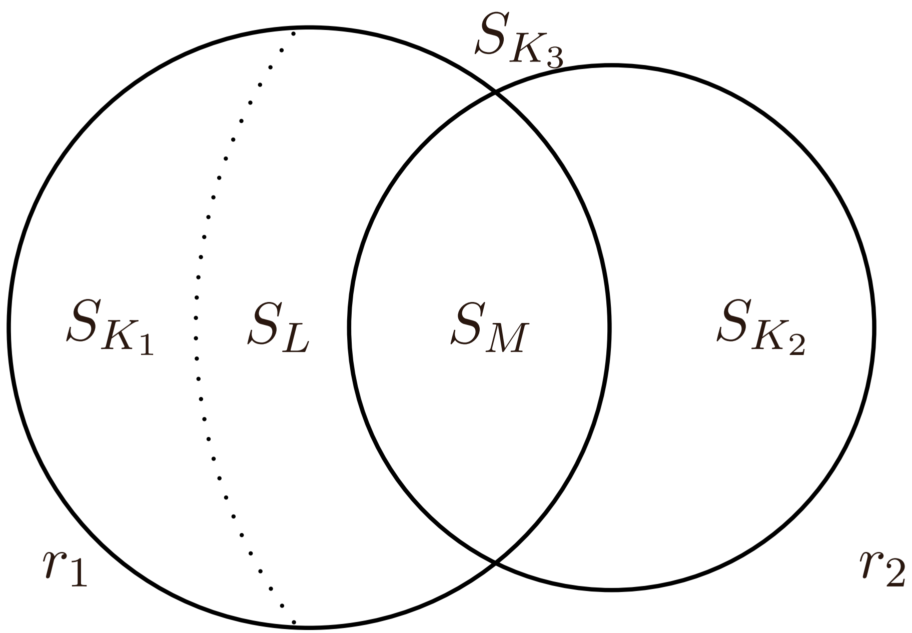


- if the set found in r_1 is reduced before commencing with r_2, the accuracy is reduced, but less slots are needed to resolve the tag set.
- can a trade–off be found, where a desired accuracy of estimation is attained with the minimum number of participating tags?



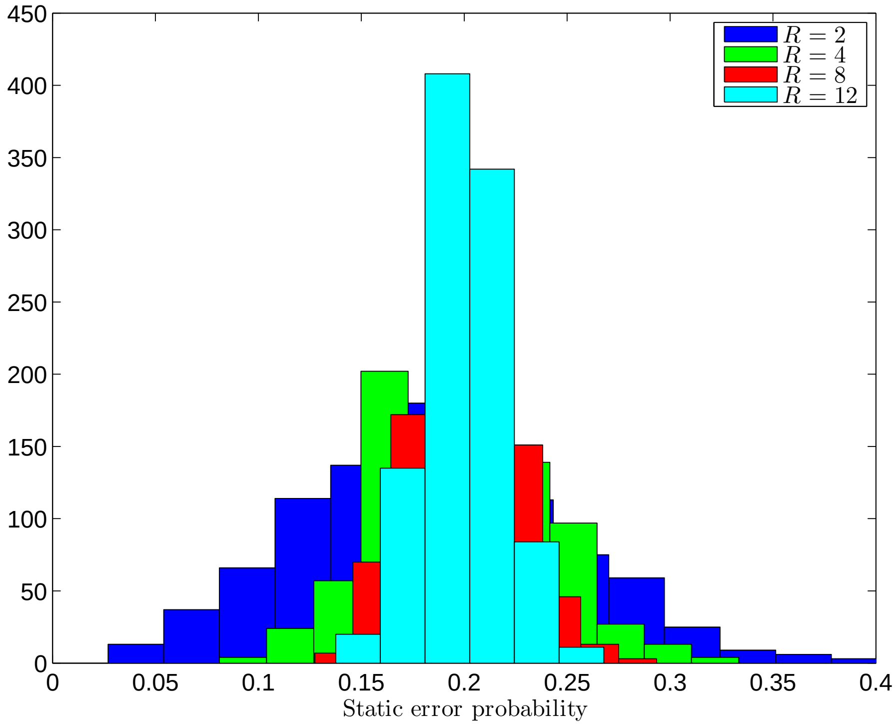


- when reducing the set that participates in r_2 by silencing the tags τ_2 and τ_6, fewer slots are needed in the arbitration to resolve the tag set.
- however, it is important to only silence tags if the estimation maintains the desired accuracy.

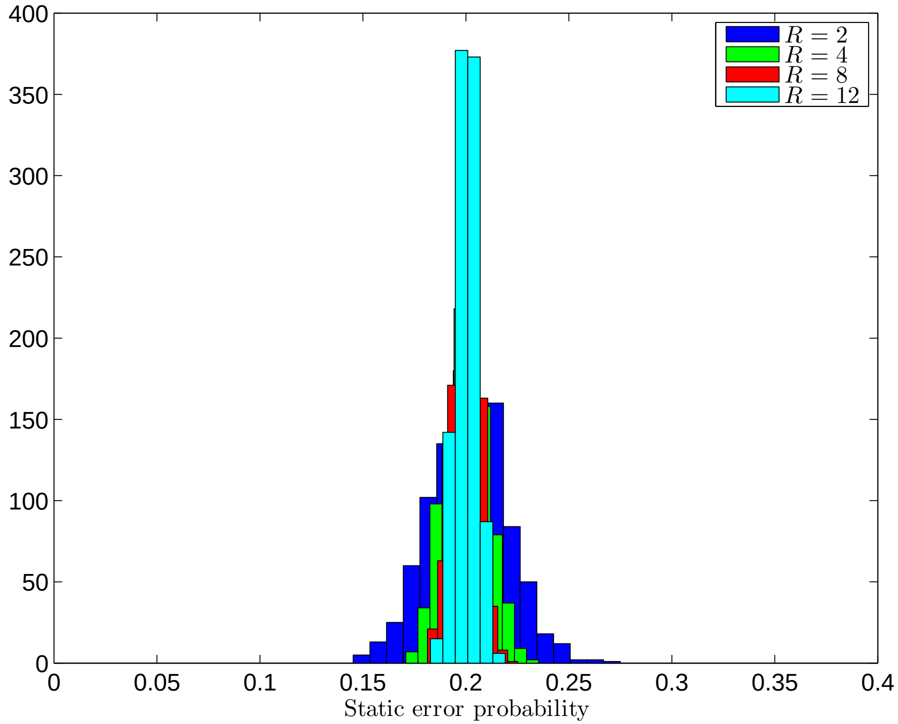


$$g(l,m) = \hat{p} = \begin{cases} 1 & \text{if } l+m=0, \\ \frac{l}{l+m} & \text{otherwise.} \end{cases}$$

- the bias of this estimator is less dependent on N, than the estimator presented in previous work.
- the variance of this estimator is less affected by the choice of v, than other attempted estimators.
- more analysis should go into investigating possible estimators, that utilizes all the observed sets of tags.

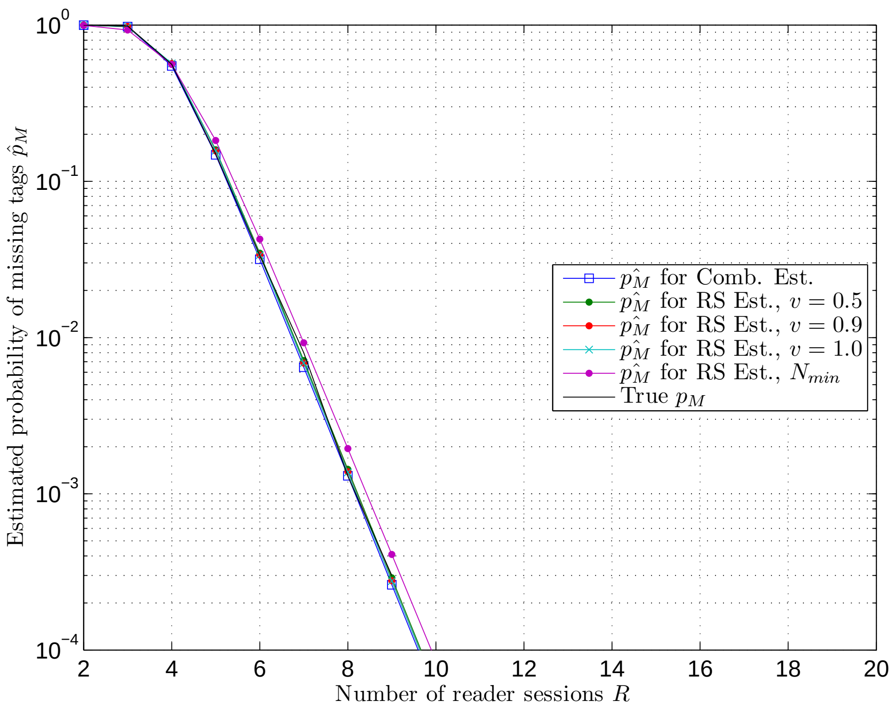


$(N_{min} = 37)$

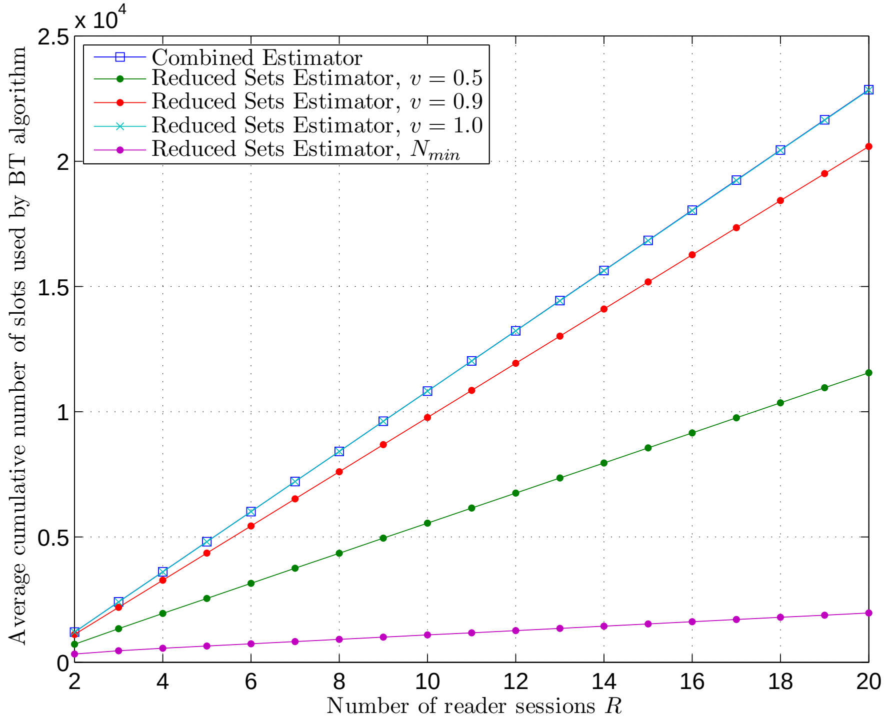


$(N_{min} = 500)$

- the number of samples decreases the bias and the variance.
- we define desired accuracy as $|(1-p)(p+(1-p)(1-v))^{N_{min}}| \leq 0.01$, which enables us to find the minimum number of tags that must be reused, denoted N_{min}, to be 37 tags.
- future work should investigate the desired accuracy of the estimate of the probability of missing tags p_M instead.



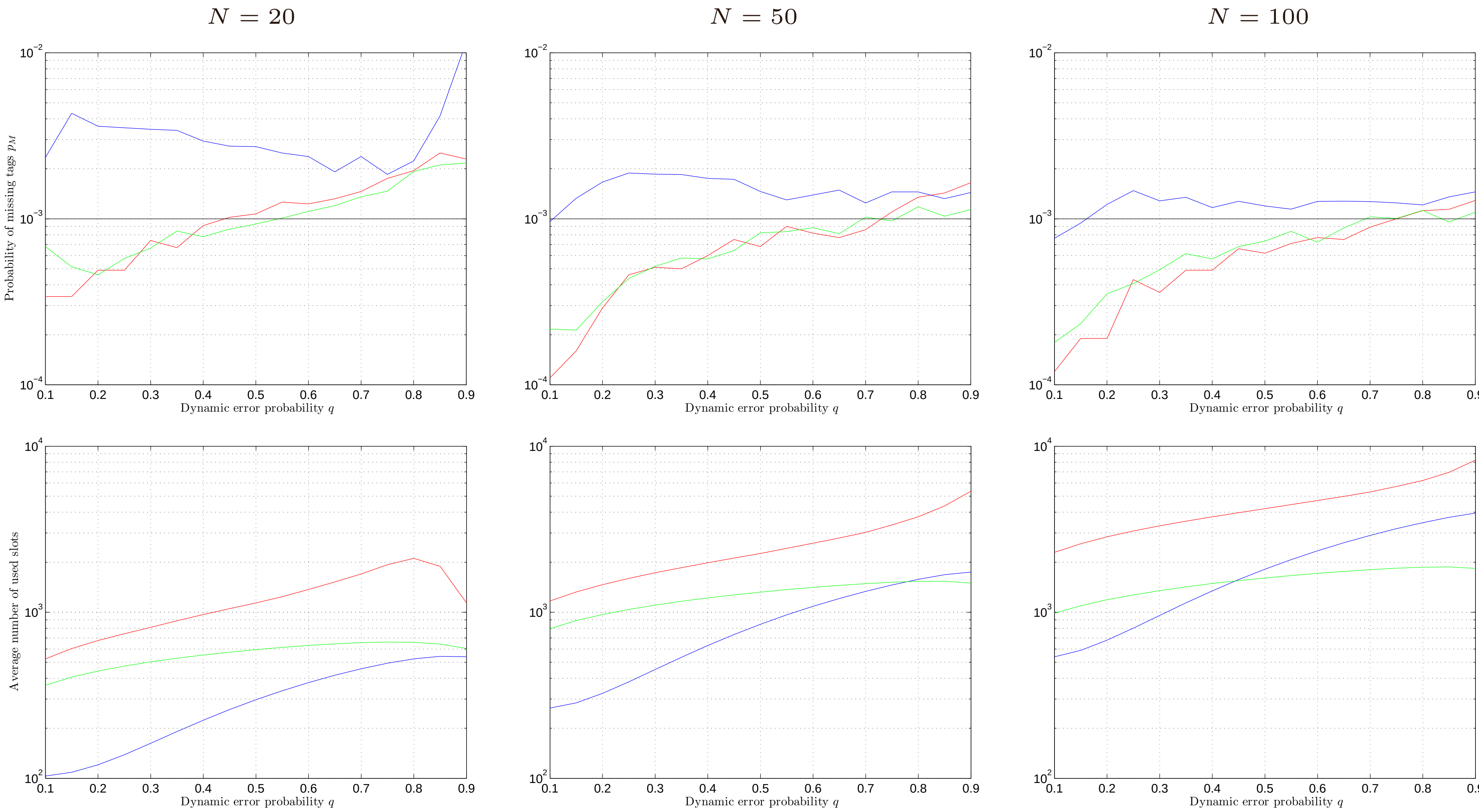
estimated probability of error.



cumulative average number of used slots.

- the estimates of p_M using the Reduced Sets estimator is as reliable as that of the Combined estimator, and follows the true p_M, except N_{min}, which slightly overestimates p_M.
- the amount of saved slots is proportional to the choice of v.
- the small overestimation may be due to a bad choice of N_{min}.

talk roundup



- Combined estimator
- scheme for reliable estimation using the relation between q and p
- Reduced Sets estimator

- the proposed Combined estimator is as reliable as present state of the art techniques in independent reader sessions, and more reliable in dependent reader sessions.
- a relation can be found between the static and dynamic error probability and for $N \geq 13$, this relation seems to become independent of N.
- the relation can be used in the sequential decision process, where it provides a slightly less reliable estimate, but uses significantly fewer slots than the original sequential decision process.
- the idea of silencing tags is further investigated, and a novel technique is found which is as reliable as the Combined estimator, but much more time efficient in terms of used slots.