Faster Optimal-rate Many-server Private Information Retrieval (PIR)

Syed Mahbub Hafiz and Ryan Henry

Indiana University—Bloomington and University of Calgary

Optimal-rate Private Information Retrieval (PIR)

- PIR enables client to fetch record(s) from remote and untrusted database with cryptographic privacy.
- Steps: query construction (Q) by the client, response generation (R) by the server, and record reconstruction (E) by the client.
- Cost-metrics: upload cost, download cost, computation cost at both client-side and server-side, and number of interaction round.
- Optimal download cost achieved by Shah et al. (ISIT 2014).
- Optimal upload cost achieved by Boyle et al. (CCS 2016).
- Optimal computation cost per server achieved by Chor et al. (FOCS 1995).
- Our proposed protocols can **CAPTURE AND OUTPERFORM** any of them with favourable settings!

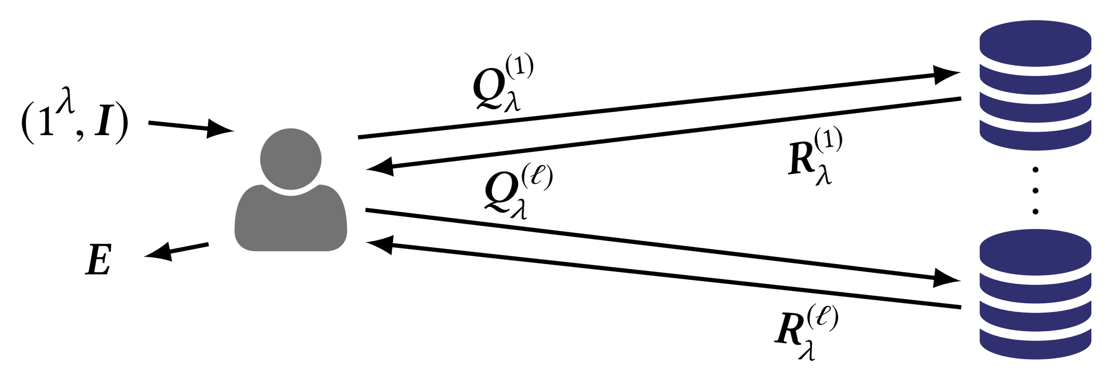

Figure: Information flow in single-round and ℓ-server PIR.

"One-extra-word" Protocols

- Augment the database $D \in \mathbb{F}^{r\times(s+1)}$ by the vector $\vec{v} \in \mathbb{F}^s$ as $D^* := D \| (D\vec{v}^T) \in \mathbb{F}^{r\times(s+1)}$.
- $M^{(r,s)} \subseteq \mathbb{F}^{r\times(s+1)}$ is the set of all height-r matrices, A, whose rows are vectors from the standard basis.
- Each of $\ell = (s+1)^r$ number of servers holds the *Frobenius inner product* $\langle D^*, A\rangle_F := \mathrm{tr}(D^* A^T)$ in its bucket.
- To retrieve D_j, client selects $A \in_R M^{(r,s)}$, and fetches $\langle D^*, B_j\rangle_F$ from bucket $\varphi(B_j)$ for each $B_j \in Eq(i; A)$.
- Thus client downloads $s+1$ **number of words, hence just an extra word beyond the whole record**, from the servers.
- Finally, client solves a system of s linear equations to reconstruct the desired record.

Perfectly 1-private "Bit-more-than-a-bit" protocols

- Governed by $\ell \geq 2$, s, the binary field $\mathbb{F} = \mathbf{GF}(2^w)$ where $w = \lceil \frac{b}{s} \rceil$, the all-0s vector $\vec{v} = \vec{0}$, and a particular mapping φ.
- New φ reduces the requirement of super-exponential number of servers (ℓ) to an arbitrary $\ell \geq 2$ with the condition of $\ell = s+1$.

Computationally 1-private "Bit-more-than-a-bit" protocols

- Our most efficient construction with $\ell = 2^L$ reduces per-server upload cost from rL bits to just $(\lambda+2)\lceil \lg \frac{r}{\lambda} \rceil L$ bits.
- Client samples independently and distributes L-tuple of $(2,2)$-DPF key pairs to the servers.
- Server performs a full-domain evaluation on the received keys and concatenates the resulting bit vectors component-wise to obtain a length-r vector.

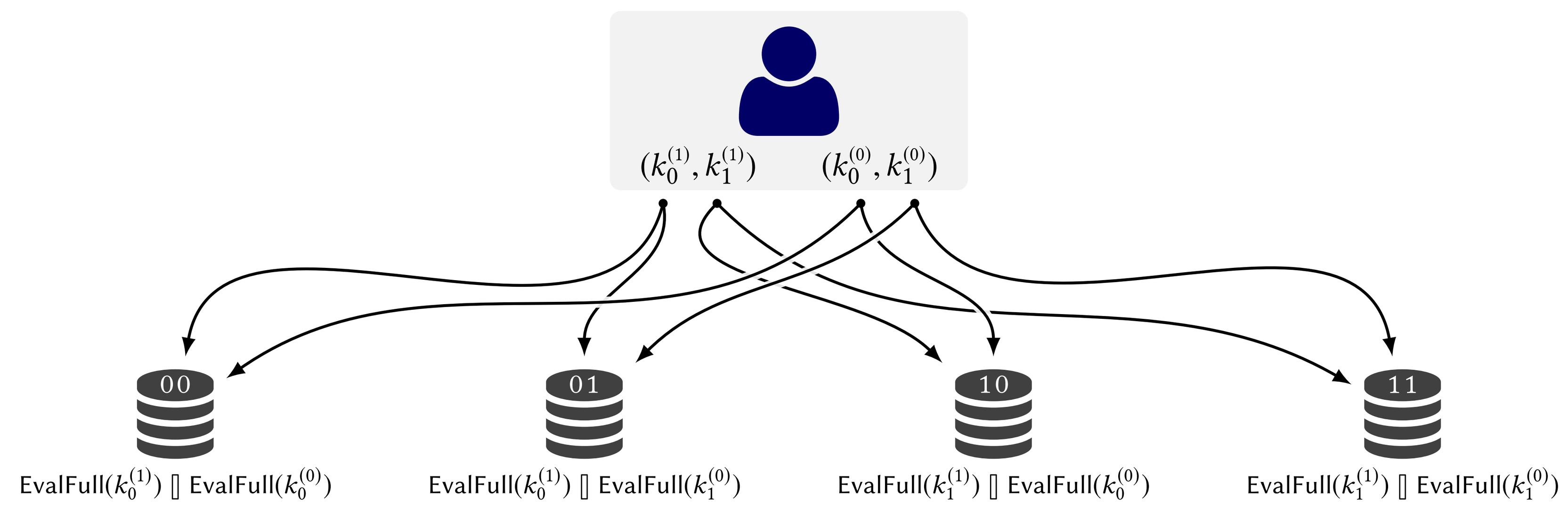

Figure: $(2,2)$-DPF key distribution and query expansion procedure for $\ell = 2^2$ servers.

Comparison in Query Construction

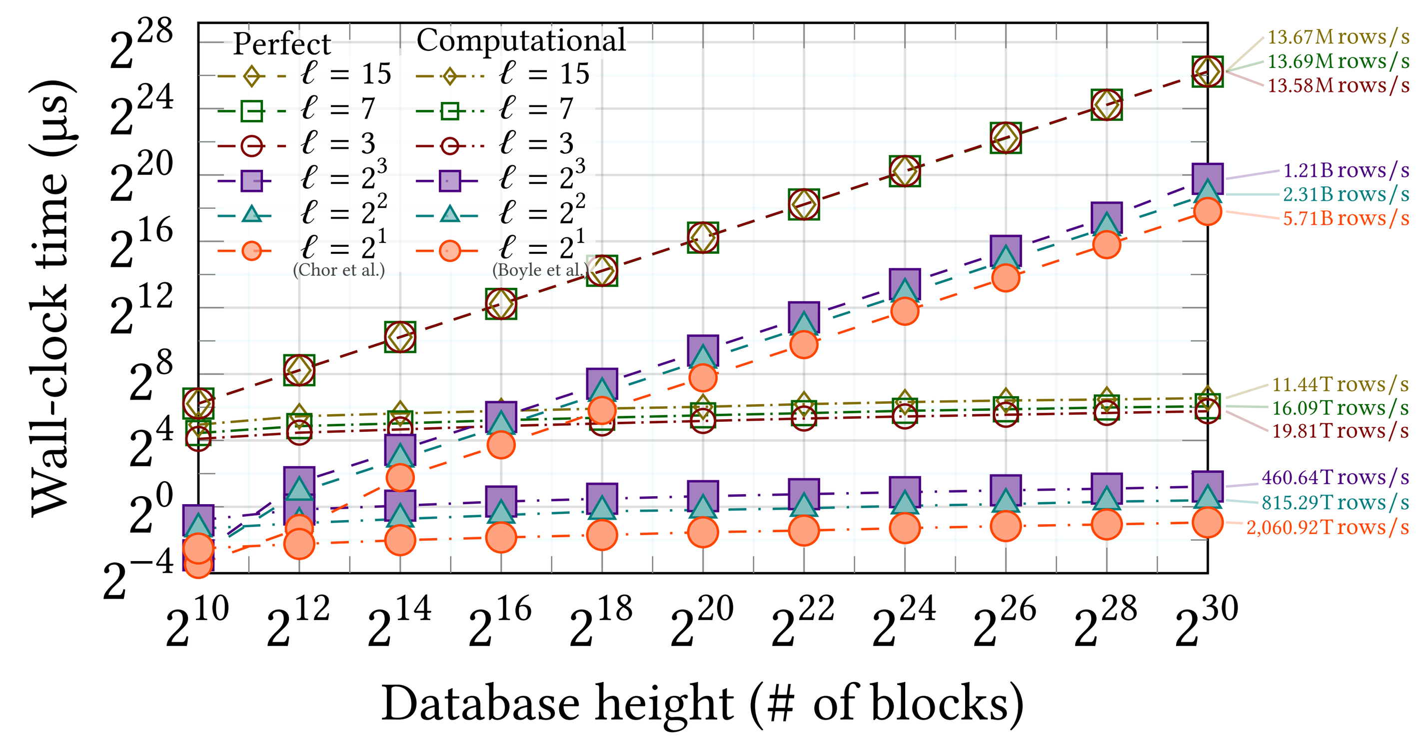

Figure: Wall-clock time for (client-side) query construction in bit-more-than-a-bit protocols.

Comparison in PIR Response Generation

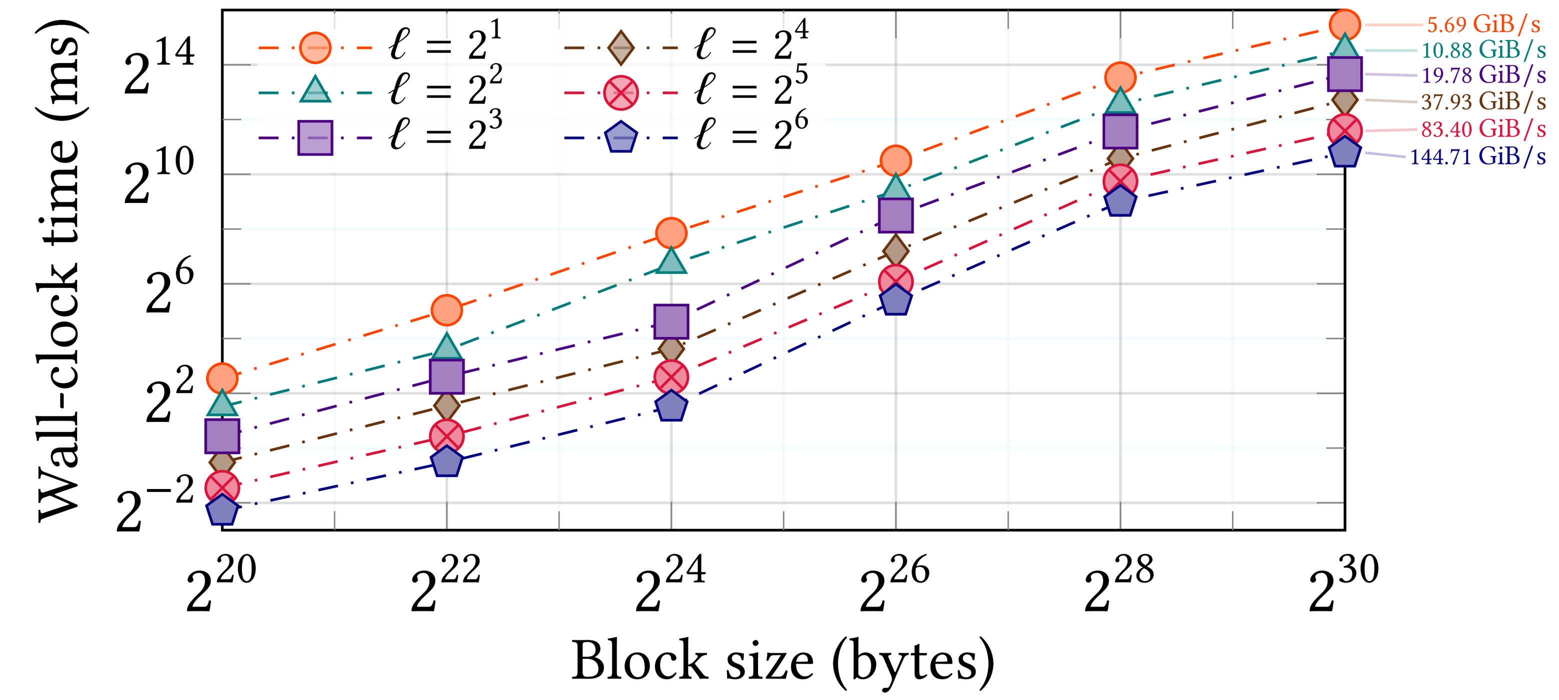

Figure: Wall-clock time for (server-side) response generation for bit-more-than-a-bit protocols. DB size scales up to 256 GiB.

Head-to-head comparison with Percy++ (2014)

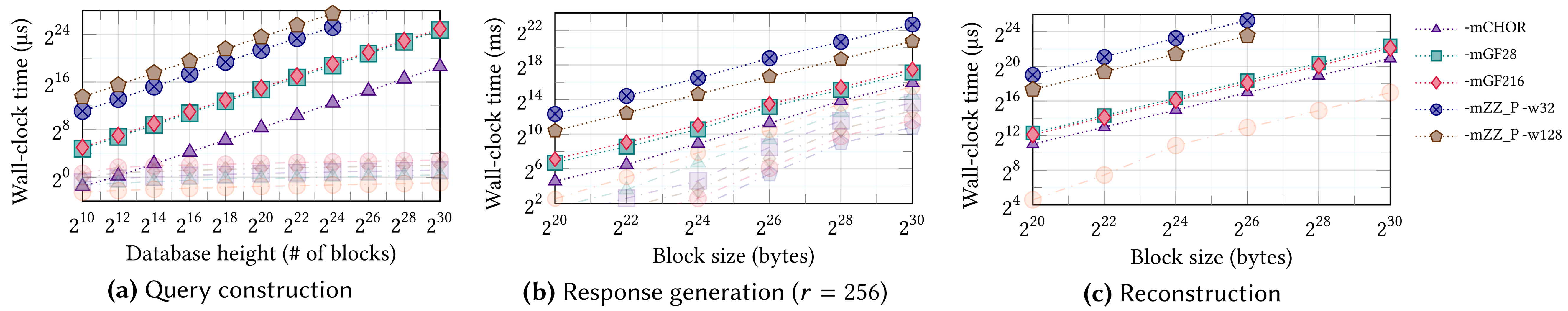

Fig. 7. Head-to-head comparison with various 1-private, 2-server instances from Percy++ v1.0. The faint plots near the bottoms show corresponding costs for bit-more-than-a-bit protocol.

Head-to-head comparison with RAID-PIR (CCSW 2014)

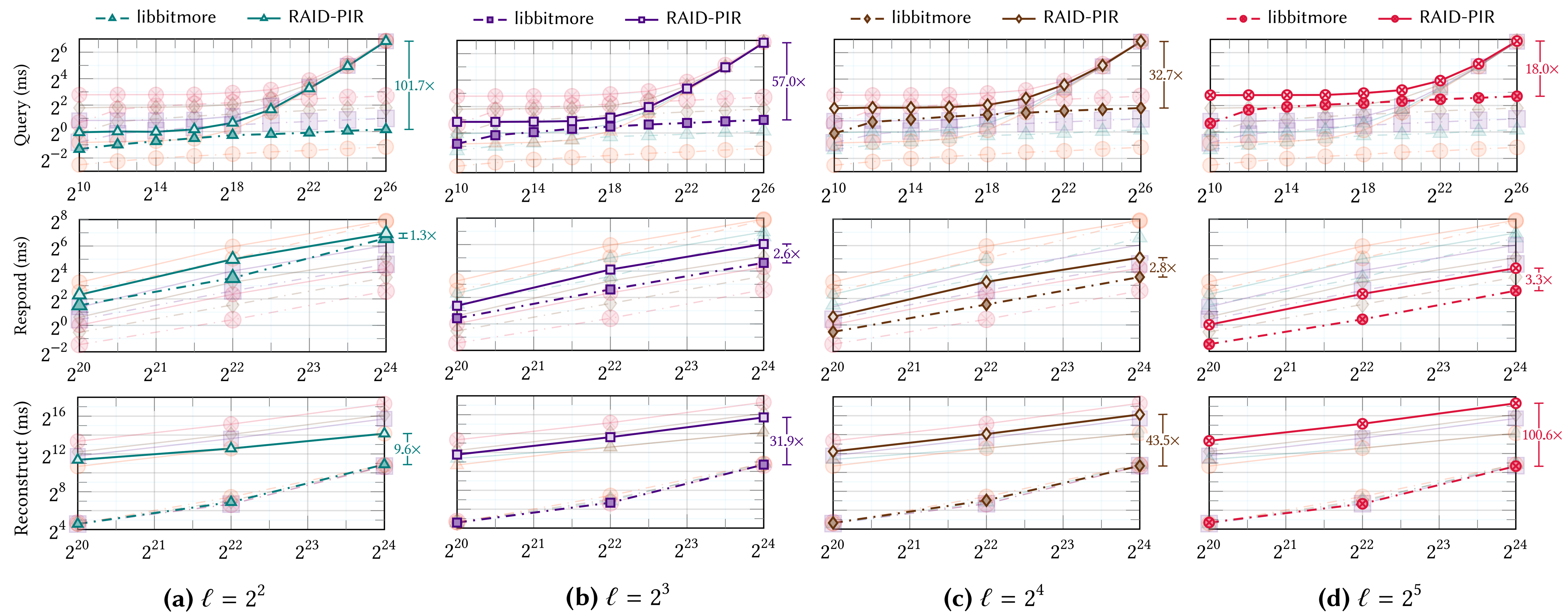

Fig. 8. Head-to-head comparison with computationally 1-private RAID-PIR v0.9.5 instances for ℓ ranging from 4 to 32. The scale of some experiments was limited because RAID-PIR v0.9.5 cannot handle databases that exceed physical memory.

This work is supported by the NSF Grant No. 1718475 and an NSERC Discovery Grant.

E-mail: shafiz@iu.edu and ryan.henry@ucalgary.ca

Homepage: sites.google.com/view/syedhafiz/