Asymptotic Analysis of ISD algorithms for the q-ary case

Rodolfo Canto Torres

27 April 2017

INVENTEURS DU MONDE NUMÉRIQUE

Computational Syndrome Decoding

CSD(H, s, w)

Let $H \in \mathbb{F}_q^{(n-k)\times n}$, $s \in \mathbb{F}_q^{n-k}$, $w \geq 0$.
Find $e \in \mathbb{F}_q^n$, $\mathrm{wt}(e) \leq w$, $He = s$.

Computational Syndrome Decoding

CSD(H, s, w)

Let $H \in \mathbb{F}_q^{(n-k)\times n}$, $s \in \mathbb{F}_q^{n-k}$, $w \geq 0$.
Find $e \in \mathbb{F}_q^n$, $\mathrm{wt}(e) \leq w$, $He = s$.

- Generic problem is believed to be hard on average.
- Central for security of code-based cryptography.
- For the binary case, the first important algorithm (ISD) was proposed by Prange in 1962 and the last one by May and Ozerov in 2015.

Computational Syndrome Decoding

CSD(H, s, w)

Let $H \in \mathbb{F}_q^{(n-k)\times n}$, $s \in \mathbb{F}_q^{n-k}$, $w \geq 0$.
Find $e \in \mathbb{F}_q^n$, $\mathrm{wt}(e) \leq w$, $He = s$.

- Generic problem is believed to be hard on average.
- Central for security of code-based cryptography.
- For the binary case, the first important algorithm (ISD) was proposed by Prange in 1962 and the last one by May and Ozerov in 2015.
- For the q-ary case, the first generalisation was done by Peters in 2010 and the last one by Gueye, Klamti and Hirose in 2017. (ISD = Information Set Decoding)

Conventional Complexity of known algorithms

Asymptotic analysis

Let $R = k/n$ (code rate), $\tau = w/n$ (error rate) such that $0 \leq \tau \leq h_q^{-1}(1 - R)$. So, for an algorithm $\mathcal{A}$ we can express its *work factor* of solving $\mathrm{CSD}(n, Rn, \tau n)$ as

$$\mathrm{WF}_A(n, Rn, \tau n, q) = 2^{c' n(1+\mathbf{o}(1))}$$

where c is a constant which depends on R, τ, $\mathcal{A}$ and q.

In this Work

Main Aim

For many variants $\mathcal{A}$ of ISD, if R and τ are constants then

$$\lim_{q\to\infty} \mathrm{WF}_{\mathcal{A}}(n, Rn, \tau n, q) = \mathrm{WF}_{\mathrm{Prange}}(n, Rn, \tau n).$$

Decoding by Force Brute

- We produce one indexed list

$$\mathcal{L} = \{(He, e), e \in \mathbb{F}_q^n, \mathrm{wt}(e) = w\}$$

- We look for $(s, e) \in \mathcal{L}$
- The cost is

$$\binom{n}{w}(q-1)^w$$

Description

We divide the problem in half :

- $H = [H_1|H_2],\ H_1, H_2 \in \mathbb{F}_q^{(n-k)\times \frac{n}{2}}$
- $e = [e_1|e_2],\ e_1, e_2 \in \mathbb{F}_q^{\frac{n}{2}}$

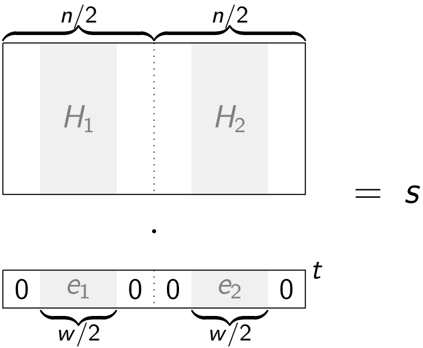
We produce two list:

$$\mathcal{L}_1 = \{(H_1 e_1, e_1), e_1 \in \mathbb{F}_q^{\frac{n}{2}}\},\ \mathcal{L}_2 = \{(s - H_2 e_2, e_2), e_2 \in \mathbb{F}_q^{\frac{n}{2}}\}$$

We have: $\quad He = s \quad \Longrightarrow \quad H_1 e_s = s - H_2 e_2.$

$$\text{Produce } \mathcal{L} \quad \Longrightarrow \quad \text{Produce } \mathcal{L}_1, \mathcal{L}_2, \mathcal{L}_1 \bowtie \mathcal{L}_2$$

Scheme



$$\boxed{H_1 e_1 = s - H_2 e_2 ?}$$

Cost

- Total Cost: $2 * \binom{n/2}{w/2}(q-1)^{w/2} + \binom{n/2}{w/2}^2 (q-1)^w / q^{n-k}$
- Probability of succes:

$$\mathcal{P} = \frac{\left(\binom{n/2}{w/2}(q-1)^{w/2}\right)^2}{\binom{n}{w}(q-1)^w} = \frac{\binom{n/2}{w/2}^2}{\binom{n}{w}}$$

Description

- Choose randomly a matrix $P_1 \in \mathbb{F}_q^{n\times n}$
- Find matrices P_2 and U s.t. $H' = UHP_1P_2 = (\mathrm{Id}_{n-k}|Q)$

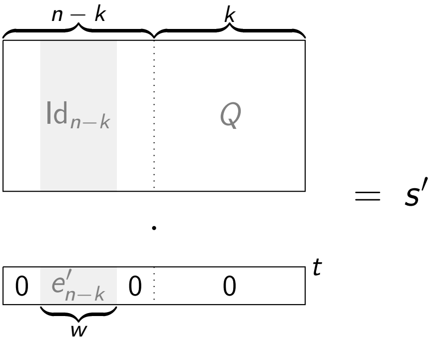
So, we obtain

$$He = s \qquad \Longrightarrow \qquad e'_{n-k} + Qe'_k = s'$$

If the first $n-k$ entries of e' were a *Information Set*, $e'_k = 0$.

$$\mathrm{CSD}(H, s, w) \qquad \Longrightarrow \qquad \text{Test } \mathrm{wt}(s') = w$$

Scheme



$$\boxed{\mathrm{wt}(s') = w?}$$

Work factor

- Cost of the permutation and diagonalization is polynomial
- Probability de succes:

$$\frac{\binom{n-k}{w}(q-1)^w}{\binom{n}{w}(q-1)^w}$$

- $\mathrm{WF}_{\mathrm{Pra}}(n,k,w) = \frac{\binom{n}{w}}{\binom{n-k}{w}}$

Description 1

- Choose randomly a matrix $P_1 \in \mathbb{F}_q^{n\times n}$
- Find matrices P_2 and U such that

$$H' = UHP_1P_2 = \begin{bmatrix} \mathrm{Id}_{n-k-\ell} & Q_{(n-k-\ell)} \\ 0 & Q_{[\ell]} \end{bmatrix}.$$

- We divide the error $e = [e_{(n-k-\ell)}, e_{[k+\ell]}]$.
- We divide the syndrome $s = [s_{(n-k-\ell)}, s_{[\ell]}]$.

Then

$$He = s \Longrightarrow \begin{cases} Q_{[\ell]}e_{[k+\ell]} = s_{[\ell]} \\ e_{(n-k-\ell)} = s_{(n-k-\ell)} - Q_{(n-k-\ell)}e_{[k+\ell]} \end{cases}$$

Description 2

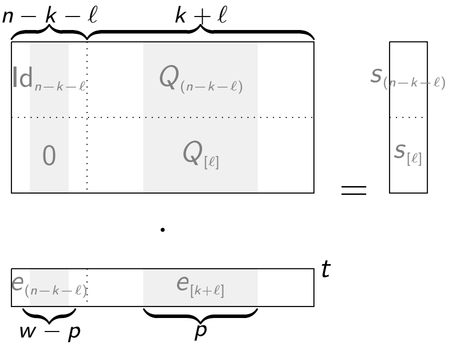
Then

$$He = s \Longrightarrow \begin{cases} Q_{[\ell]} e_{[k+\ell]} = s_{[\ell]} \\ e_{(n-k-\ell)} = s_{(n-k-\ell)} - Q_{(n-k-\ell)} e_{[k+\ell]} \end{cases}$$

If we divide the weight $w = p + (w - p)$:

$$\mathrm{CSD}(H, s, w) \Longrightarrow \begin{cases} \mathrm{CSD}(Q_{[\ell]}, s_{[\ell]}, p) \\ \text{Test } \mathrm{wt}(s_{(n-k-\ell)} - Q_{(n-k-\ell)} e_{[k+\ell]}) = w - p \end{cases}$$

Scheme



Solve CSD$(Q_{[\ell]}, s_{[\ell]}, p, q)$
$\mathrm{wt}(s_{(n-k-\ell)} - Q_{(n-k-\ell)} e_{[k+\ell]}) = w?$

Example: Stern-Dumer's Algorithm

We solve $\mathrm{CSD}(Q_{[\ell]}, s_{[\ell]}, p)$ by the Birthday Decoding

- The cost of solving $\mathrm{CSD}(Q_{[\ell]}, s_{[\ell]}, p)$:

$$\sqrt{\binom{k+\ell}{p}(q-1)^{p/2}} + \binom{k+\ell}{p}(q-1)^{p}/q^{\ell}$$

- Probability of succes is

$$\mathcal{P} = \frac{\binom{n-k-\ell}{w-p}\binom{(k+\ell)/2}{p/2}^2}{\binom{n}{w}} \approx \frac{\binom{n-k-\ell}{w-p}\binom{k+\ell}{p}}{\binom{n}{w}}$$

- $\mathrm{WF}_{\mathrm{SD-ISD}} =$
$\min_{p,\ell}\left\{\mathcal{P}^{-1}\left(\sqrt{\binom{k+\ell}{p}(q-1)^{p/2}} + \binom{k+\ell}{p}(q-1)^{p}/q^{\ell}\right)\right\}$

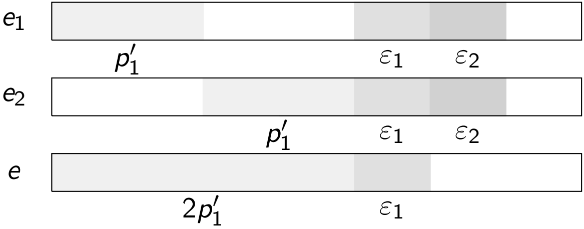
Description 1

- New sub-procedure for $\mathrm{CSD}(Q_{[\ell]}, s_{[\ell]}, p)$: *ColumnMatch*
- *ColumnMatch*:

$$Q_{[\ell]} e_1 = s - Q_{[\ell]} e_2, \qquad e = e_1 + e_2, \quad e_1, e_2 \in \mathbb{F}_q^{\ell}$$

where the supports of e_1, e_2 can have intersection.

Description 2



$$\mathrm{wt}(e_1) = \mathrm{wt}(e_2) = p_1 = p_1' + \varepsilon_1 + \varepsilon_2$$

$$\mathrm{wt}(e) = p = 2p_1' + \varepsilon_1$$

Probability $p_1 \oplus p_1 = p$

For p, p_1 fixed:

$$\mu_{p'_1,\varepsilon_1,\varepsilon_2} = \frac{\binom{p_1}{\varepsilon_2}\binom{p'_1+\varepsilon_1}{\varepsilon_1}\binom{k+\ell-p_1}{p'_1}}{\binom{k+\ell}{p_1}}\left(\frac{1}{q-1}\right)^{\varepsilon_2}\left(\frac{q-2}{q-1}\right)^{\varepsilon_1}$$

So, for $\varepsilon_2 = 0, 1, \ldots, p_1 - p/2$

$$\mu_{\varepsilon_2} = \frac{\binom{p_1}{\varepsilon_2}\binom{p_1-\varepsilon_2}{2p_1-p-2\varepsilon_2}\binom{k+\ell-p_1}{p-p_1+\varepsilon_2}}{\binom{k+\ell}{p_1}}\left(\frac{q-1}{(q-2)^2}\right)^{\varepsilon_2}\left(\frac{q-2}{q-1}\right)^{2p_1-p}$$

The total probability $p_1 \oplus p_1 = p$: $\mu = \sum_{\varepsilon_2} \mu_{\varepsilon_2}$.

Number of representations $p_1 \oplus p_1 = p$

For p, p_1 fixed:

$$\rho_{p_1',\varepsilon_1,\varepsilon_2} = \binom{2p_1' + \varepsilon_1}{\varepsilon_1}(q-2)^{\varepsilon_1}\binom{2p_1'}{p_1'}\binom{k+\ell-2p_1'-\varepsilon_1}{\varepsilon_2}(q-1)^{\varepsilon_2}$$

We can prove

$$\rho_{\varepsilon_2} = \mu_{\varepsilon_2}\frac{\binom{k+\ell}{p_1}^2}{\binom{k+\ell}{p}}(q-1)^{2p_1-p}$$

The total number of representations $p_1 \oplus p_1 = p$

$$\rho = \sum_{\varepsilon_2} \rho_{\varepsilon_2} = \mu\frac{\binom{k+\ell}{p_1}^2}{\binom{k+\ell}{p}}(q-1)^{2p_1-p}$$

- We PRODUCE with Birthday decoding:

$$\mathcal{L}_1 = \{(Q_{[\ell]}e_1, e_1), e_1 \text{ solution } \mathrm{CSD}(Q_{[r]}, 0, p_1)\}$$

$$\mathcal{L}_2 = \{(s - Q_{[\ell]}e_2, e_2), e_2 \text{ solution } \mathrm{CSD}(Q_{[r]}, s_{[r]}, p_1)\}$$

- We calcule $\mathcal{L}_1 \bowtie \mathcal{L}_2$ to solve $\mathrm{CSD}(Q_{[\ell]}, s_{[\ell]}, p)$

$$\mathrm{CSD}(H, s, w) \Longrightarrow \begin{cases} \mathrm{CSD}(Q_{[r]}, 0, p_1), \mathrm{CSD}(Q_{[r]}, s_{[r]}, p_1), \\ \mathrm{CSD}(Q_{[\ell]}, s_{[\ell]}, p) \\ \text{ Test } \mathrm{wt}(s_{(n-k-\ell)} - Q_{(n-k-\ell)}e_{[k+\ell]}) = w - p \end{cases}$$

- We divide the size $\mathcal{L}_1$, $\mathcal{L}_2$ by a factor ρ, so we have to do $r = \log_q(\rho)$.
- La probability of succes $\mathcal{P}$ is the same as SD-ISD.

So,

$$\mathrm{WF}_{\mathrm{MMT}-ISD} = \min_{p,\ell,p_1} \left\{ \mathcal{P}^{-1} \left(\sqrt{\binom{k+\ell}{p_1}(q-1)^{p_1}} + \binom{k+\ell}{p_1} \frac{(q-1)^{p_1}}{\rho} + \binom{k+\ell}{p} \frac{(q-1)^{p}}{q^{\ell}\mu} \right) \right\}$$

We give a lower bound for the SD-ISD and MMT-ISD algorithms.

Lemma

For sufficiently large values of n and k, we have

$$\mathrm{WF}_{\mathcal{A}}(n,k,w,q) \geq \min_{p,\ell} \frac{\binom{n}{w}}{\binom{n-k-\ell}{w-p}} \left(\frac{(q-1)^{(1-a)p}}{\binom{k+\ell}{ap}} + \frac{(q-1)^p}{q^\ell} \right),$$

where a is equals to $1/2$ and $3/4$, when $\mathcal{A}$ is SD-ISD and MMT-ISD.

We asociate this bound to the function

$$B_{a,q}(\ell, p) = \frac{\binom{n}{w}}{\binom{n-k-\ell}{w-p}} \left(\frac{(q-1)^{(1-a)p}}{\binom{k+\ell}{ap}} + \frac{(q-1)^p}{q^\ell} \right).$$

Lemma

For $w < \frac{n-k}{2}$, there is ℓ^*, p^* such that

$$\min_{\ell,p} B_{a,q}(\ell, p) = B_{a,q}(\ell^*, p^*) \quad \text{and} \quad q^{\ell^*} = \binom{k+\ell^*}{ap^*}(q-1)^{ap^*}$$

We obtain the following inequalitiy

$$|\ell - ap^*| \leq \log\left(\binom{k+\ell^*}{ap^*}\right) / \log(q).$$

So, $\ell \to ap^*$ when $q \to \infty$. Moreover,

$$\mathrm{WF}_{\mathrm{Pra-ISD}} \geq \mathrm{WF}_{\mathcal{A}} \geq q^{(1-ap^*)} \frac{\binom{n}{w}}{\binom{n-k-\ell^*}{w-p^*}} \frac{(q-1)^{ap^*}}{q^{\ell^*}}$$

Therefore,

$$\lim_{q\to\infty} \ell^* = \lim_{q\to\infty} p^* = 0.$$

.

Main Result

Theorem

For all $\mathcal{A}$ among the q-ary version of SD-ISD, MMT-ISD and BJMM-NN-ISD[a]*, code rate R and error rate $\tau \leq h^{-1}(1-R)$, we have*

$$\lim_{q\to\infty} \mathrm{WF}_{\mathcal{A}}(n, Rn, \tau n, q) = \mathrm{WF}_{\mathrm{Pra-ISD}}(n, Rn, \tau n).$$

[a] In march 2017, binary BJMM algorithm was extented to q-ary case by using the $q-ary$ nearest neighbor technique

Numerical Results

	Pra-ISD	SD-ISD			MMT-ISD		
q	c	c	ℓ/n	p/n	c	ℓ/n	p/n
3	0.1826	0.1773	0.0291	0.0119	0.1631	0.0471	0.0287
8	0.3084	0.3038	0.0507	0.0327	0.2903	0.0162	0.0099
64	0.4875	0.4861	0.0033	0.0027	0.4832	0.0118	0.0072
128	0.5286	0.5278	0.0018	0.0015	0.5262	0.0060	0.0035
256	0.5633	0.5628	0.0010	0.0008	0.5620	0.0030	0.0017

Table 1: Values in code rate $R = 0.45$ and error rate $\tau = h_2^{-1}(1 - R)$

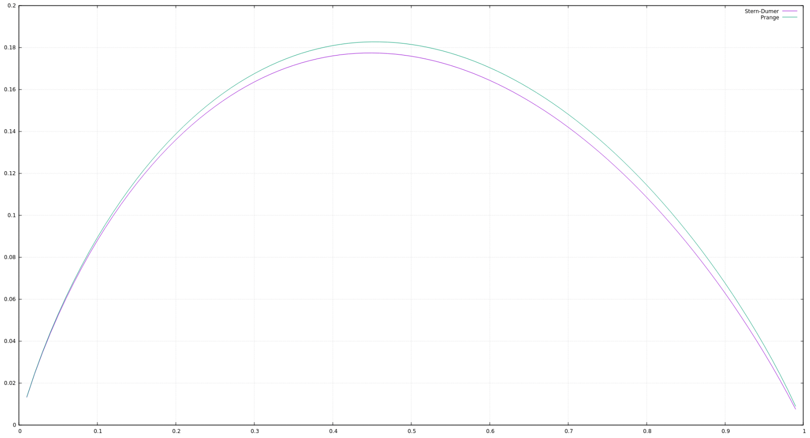


Figure 1: $R, \tau = h_q^{-1}(1-R)$ vs $c \quad (q=3)$

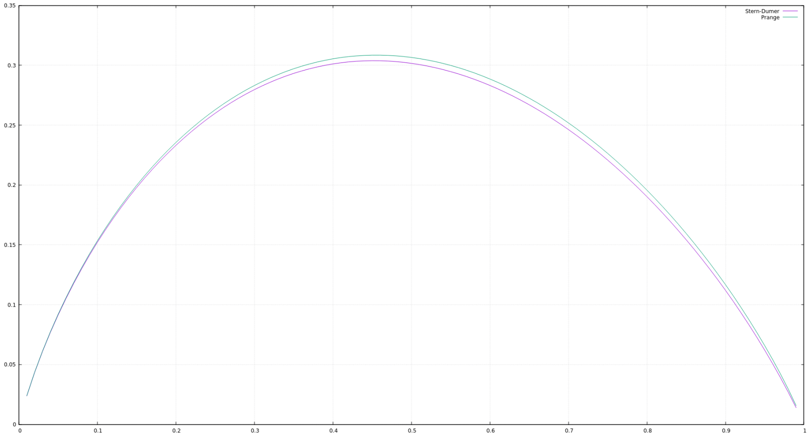


Figure 2: $R, \tau = h_q^{-1}(1 - R)$ vs c $\quad (q = 8)$

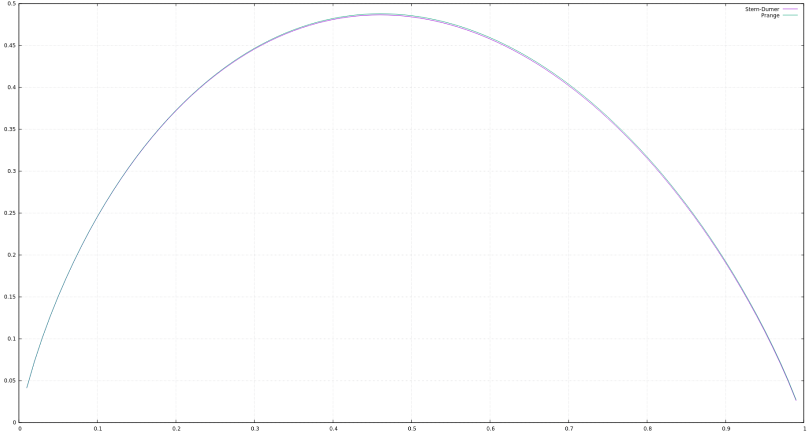


Figure 3: $R, \tau = h_q^{-1}(1 - R)$ vs $c \quad (q = 64)$

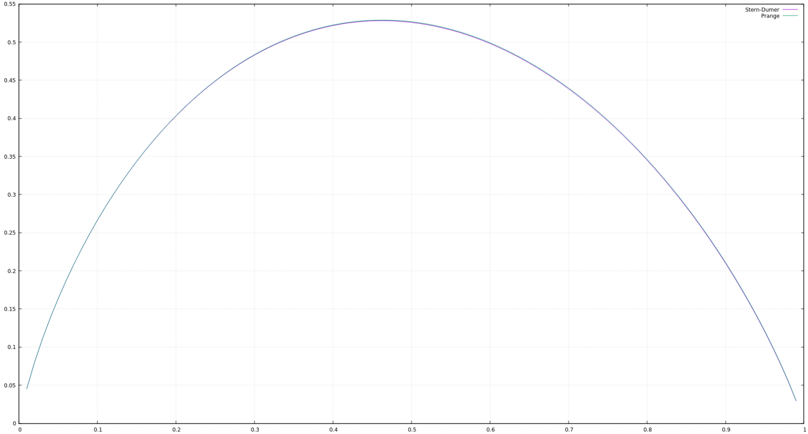


Figure 4: $R, \tau = h_q^{-1}(1-R)$ vs c $\quad (q = 128)$

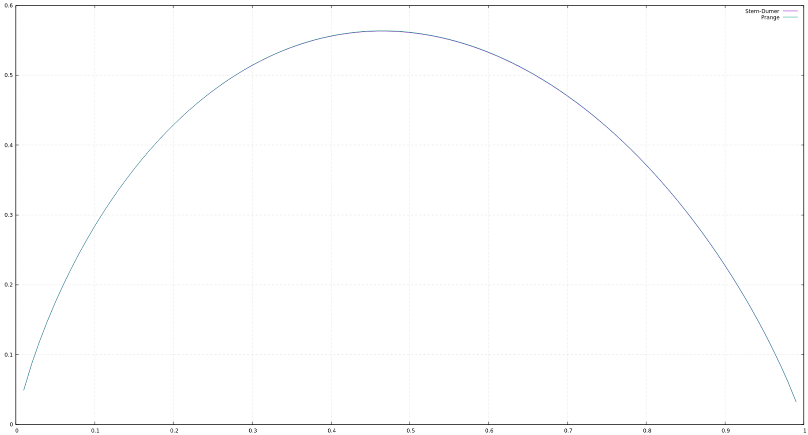


Figure 5: $R, \tau = h_q^{-1}(1 - R)$ vs c $\quad (q = 256)$

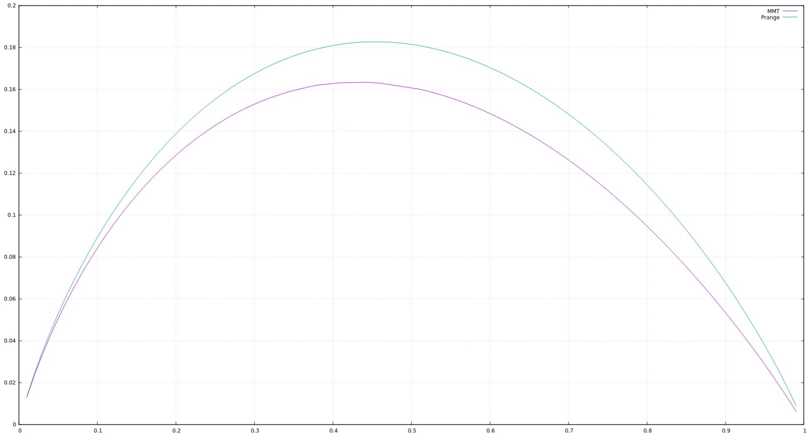


Figure 6: $R, \tau = h_q^{-1}(1 - R)$ vs $c \quad (q = 3)$

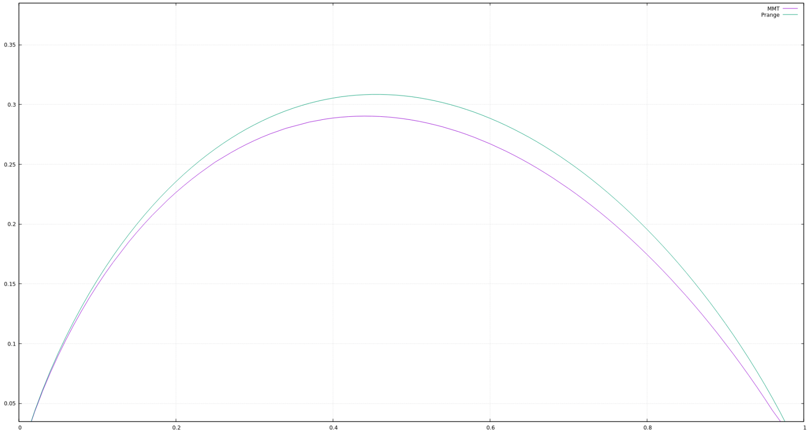


Figure 7: $R, \tau = h_q^{-1}(1 - R)$ vs $c \quad (q = 8)$

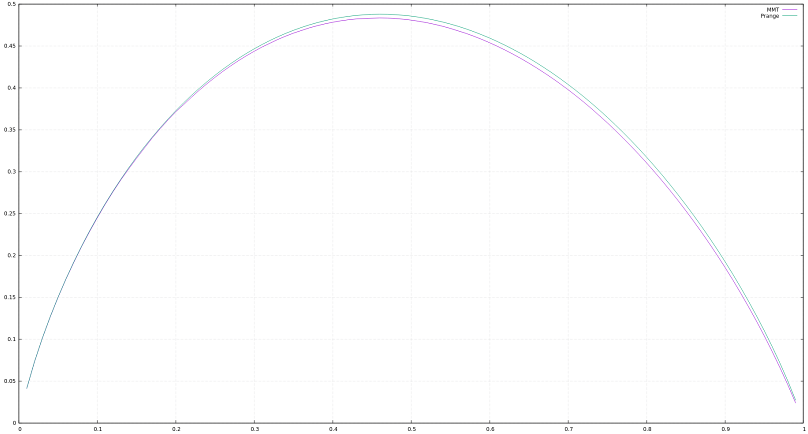


Figure 8: $R, \tau = h_q^{-1}(1 - R)$ vs $c \quad (q = 64)$

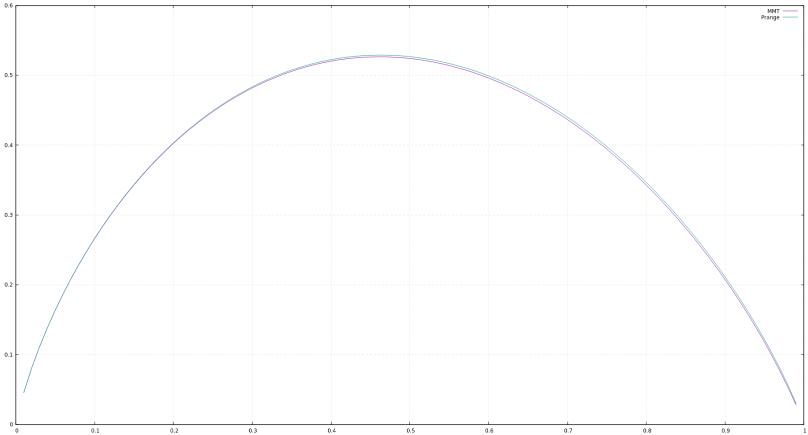


Figure 9: $R, \tau = h_q^{-1}(1-R)$ vs c $\quad (q = 128)$

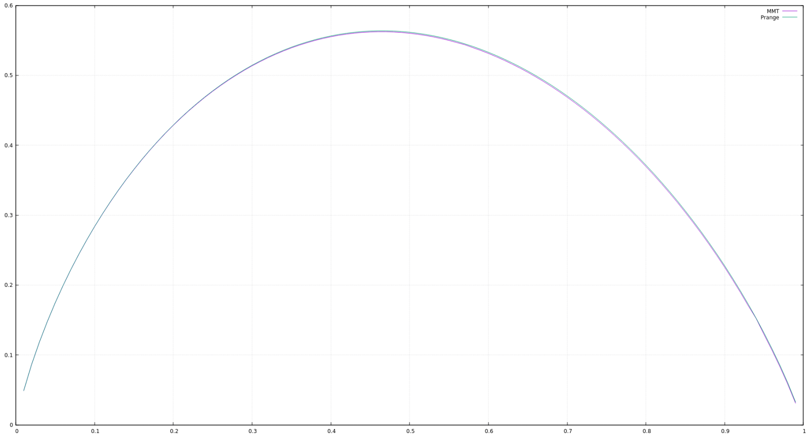


Figure 10: $R, \tau = h_q^{-1}(1 - R)$ vs $c \quad (q = 256)$