Projets de codes correcteurs

Julien Coolen

2 avril 2022

Table des matières

1 Décodage par syndrôme

1.1 Algorithme de Prange

Données : $\boldsymbol{H} \in \mathbb{F}_q^{(n-k)\times n}$, $\boldsymbol{s} \in \mathbb{F}_q^{n-k}$, $w \in [\![0, n]\!]$.
Résultat : $\boldsymbol{e} \in \mathbb{F}_q^n$ tel que $\boldsymbol{He}^T = \boldsymbol{s}$ et $|\boldsymbol{e}| \leq w$.

répéter
  Choisir $I \subseteq [\![1, n]\!]$ tel que $|I| = k$ et $J = [\![1, n]\!] \setminus I$
  si $\boldsymbol{H}_J$ *est inversible* **alors**
    $\boldsymbol{s}' \leftarrow \boldsymbol{H}_J^{-1}\boldsymbol{s}^T$
  fin
jusqu'à $|\boldsymbol{s}'| \leq w$;
retourner $\boldsymbol{e}$ *tel que* $\boldsymbol{e}_I = 0$ *et* $\boldsymbol{e}_J = \boldsymbol{s}'^T$.

Algorithme 1 : Algorithme de Prange (1962) (type Las Vegas)

Preuve de l'algorithme de Prange. On a bien $|\boldsymbol{e}| = \underbrace{|\boldsymbol{e}_I|}_{=0} + |\boldsymbol{e}_J| \leq w$ car I et J sont disjoints. De plus la sortie $\boldsymbol{e}$ de l'algorithme vérifie bien

$$\begin{aligned} \boldsymbol{He}^T &= \boldsymbol{H}(\boldsymbol{e}_I + \boldsymbol{e}_J)^T \\ &= \boldsymbol{H}_I\boldsymbol{e}_I^T + \boldsymbol{H}_J\boldsymbol{e}_J^T \\ &= 0 + \boldsymbol{H}_J\boldsymbol{H}_J^{-1}\boldsymbol{s} \\ &= \boldsymbol{s}. \end{aligned}$$

□

Soit un code linéaire $[n, k]_q$ et $w \in [\![0, (1 - \frac{1}{q})(n - k)]\!]$. La complexité de Prange est

$$T_{\text{Prange}} = O\left(\frac{\min\left\{\binom{n}{w}(q-1)^w, q^{n-k}\right\}}{\binom{n-k}{w}(q-1)^w}\right).$$

Démonstration. La probabilité de succès d'une itération de l'algorithme est dans le pire cas ($|\boldsymbol{e}| = w$) :

$$\mathbb{P}_{\text{succès}} \geq \frac{\binom{n-k}{w}(q-1)^w}{\binom{n}{w}(q-1)^w} \max\{1, |\text{DecSynd}(\boldsymbol{H}, \boldsymbol{s}, w)|\}$$

avec

$$|\text{DecSynd}(\boldsymbol{H}, \boldsymbol{s}, w)| = q^k \frac{\binom{n}{w}(q-1)^w}{q^n}.$$

D'où

$$\begin{aligned}
\mathbb{P}_{\text{succès}} &\geq \frac{\binom{n-k}{w}(q-1)^w}{\binom{n}{w}(q-1)^w} \max\left\{1, \frac{\binom{n}{w}(q-1)^w}{q^{n-k}}\right\} \\
&\geq \binom{n-k}{w}(q-1)^w \max\left\{\frac{1}{\binom{n}{w}(q-1)^w}, \frac{1}{q^{n-k}}\right\} \\
&\geq \binom{n-k}{w}(q-1)^w \frac{1}{\min\{\binom{n}{w}(q-1)^w, q^{n-k}\}}.
\end{aligned}$$

□

Le coût C d'une itération est dominé par le coût polynomial de l'élimination gaussienne sur $\mathbb{F}_q$: $C = n(n-k)^2 \log^2 q$. En effet, on effectue $n(n-k)$ opérations sur les colonnes, et une multiplication sur $\mathbb{F}_q$ requiert $O(\log^2 q)$ opérations en bits.

Bien que le coût d'une itération est dominé par les facteurs exponentiels, on doit le garder dans le grand O car c'est un facteur non constant (polynomial) :

$$\begin{aligned}
T_{Prange} &= O\left(C \cdot \frac{1}{\mathbb{P}_{\text{succès}}}\right) \\
&= O\left(\frac{n(n-k)^2(\log^2 q)\min\left\{\binom{n}{w}(q-1)^w, q^{n-k}\right\}}{\binom{n-k}{w}(q-1)^w}\right).
\end{aligned}$$

On pose $w = \lfloor \omega n \rfloor$. La complexité asymptotique de Prange est $\tilde{O}(q^{n\alpha})$ avec

$$\alpha \coloneqq \min\left\{h_q(\omega), 1-R\right\} - (1-R)h_q\left(\frac{\omega}{1-R}\right).$$

Remarque. T_{Prange} est maximal lorsque $\binom{n}{w}(q-1)^w = q^{n-k}$ donc lorsque $w = d_{GV} \coloneqq h_q^{-1}(1-R)$. En effet $\delta_{GV}(R) \coloneqq d_{GV}/n = h_q^{-1}(1-R) + o(1)$ et $\binom{n}{d_{GV}}(q-1)^{d_{GV}} = O(q^{nh_q(d_{GV}/n)}) = O(q^{nh_q(h_q^{-1}(1-k/n))}) = O(q^{n-k})$.

La complexité de cet algo de décodage par ensemble d'information est donc exponentielle en le nombre d'erreurs w du mot à décoder si l'on cherche une solution particulière. Et pour w fixé, il faut n grand pour obtenir une complexité (= borne supérieure sur la sécurité) de l'ordre de 2^{100} (donc la taille des clés = longueur des lignes de la matrice de parité = longueur des colonnes de la matrice génératrice du code explose).

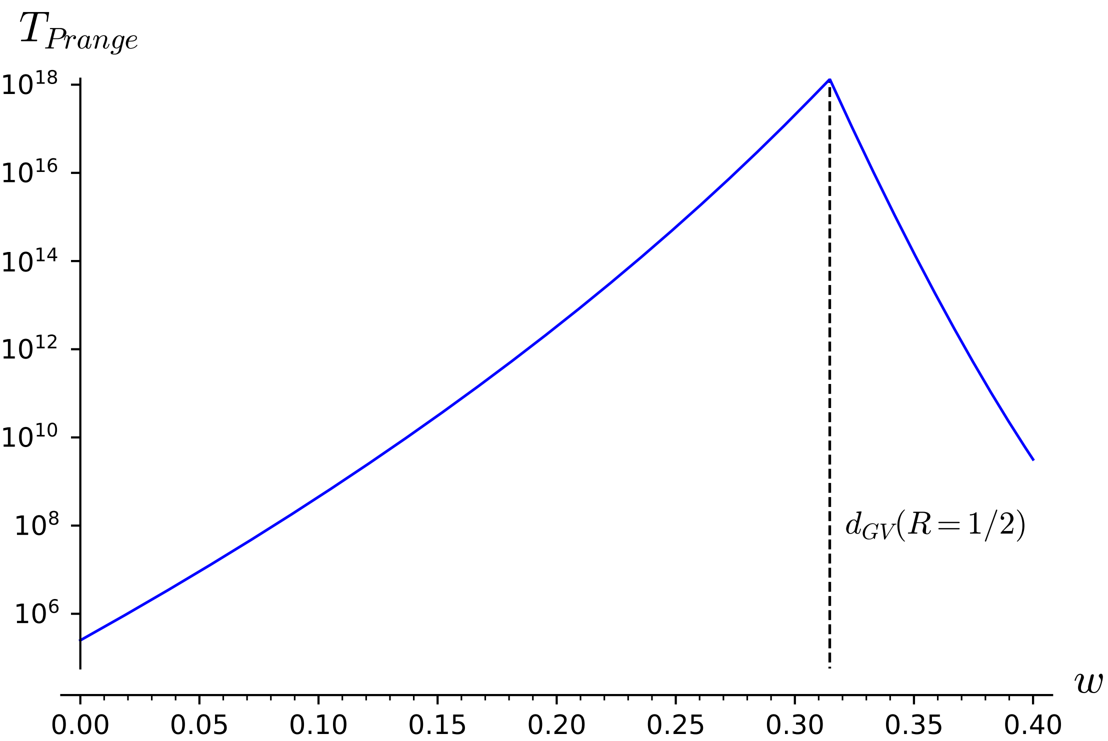

FIGURE 1 – Complexité asymptotique de Prange pour un rendement $R = \frac{1}{2}$. T_{Prange} est bien maximal pour la distance de Gilbert-Varshamov, la distance la plus difficile à décoder.

1.2 Algorithme de Lee-Brickell

L'idée est de relaxer la condition de Prange ($|\boldsymbol{e}_I| = 0$) pour amortir le coût de l'élimination gaussienne : on prend $p \geq 0$ petit et on l'on retourne un vecteur $\boldsymbol{e}$ tel que $|\boldsymbol{e}_I| = p$ et $|\boldsymbol{e}_J| \leq w - p$. Si $p = 0$ il s'agit exactement de l'algo de Prange.

Données : $\boldsymbol{H} \in \mathbb{F}_q^{(n-k)\times n}$, $\boldsymbol{s} \in \mathbb{F}_q^{n-k}$, $w \in [\![0, n]\!]$.
Résultat : $\boldsymbol{e} \in \mathbb{F}_q^n$ tel que $\boldsymbol{H}\boldsymbol{e}^T = \boldsymbol{s}$ et $|\boldsymbol{e}| \leq w$.

$X = \{\boldsymbol{x} \in \mathbb{F}_q^k : |\boldsymbol{x}| = p\}$
$\mathcal{L}$ est une table de hachage, indexée par des éléments de $\mathbb{F}_q^{n-k}$
répéter
    Choisir un ensemble d'information $I \subseteq [\![1, n]\!]$, $|I| = k$, et son complémentaire J
    pour tous les $\boldsymbol{x} \in X$ **faire**
        si $\boldsymbol{H}_J$ *est inversible* **alors**
            $\mathcal{L}[\boldsymbol{H}_J^{-1}\boldsymbol{s} - \boldsymbol{H}_J^{-1}\boldsymbol{H}_I\boldsymbol{x}^T] \leftarrow \boldsymbol{x}$
        fin
    fin
jusqu'à $\exists(\boldsymbol{y}, \boldsymbol{x}) \in \mathcal{L}, |\boldsymbol{y}| \leq w - p$;
retourner $\boldsymbol{e}$ *tel que* $\boldsymbol{e}_I = \boldsymbol{x}$ *et* $\boldsymbol{e}_J = \boldsymbol{y}^T$ *($\boldsymbol{y}$ clé obtenue à partir de $\boldsymbol{x}$).*

Algorithme 2 : Algorithme de Lee-Brickell

Preuve de l'algorithme de Lee-Brickell. On a bien $|\boldsymbol{e}| = |\boldsymbol{e}_I| + |\boldsymbol{e}_J| \leq p + (w-p) \leq w$ car I et J sont disjoints. De plus la sortie $\boldsymbol{e}$ de l'algorithme vérifie bien

$$\begin{aligned}\boldsymbol{H}\boldsymbol{e}^T &= \boldsymbol{H}(\boldsymbol{e}_I + \boldsymbol{e}_J)^T \\ &= \boldsymbol{H}_I\boldsymbol{x}^T + \boldsymbol{H}_J\boldsymbol{y} \\ &= \boldsymbol{H}_I\boldsymbol{x}^T + \boldsymbol{H}_J(\boldsymbol{H}_J^{-1}\boldsymbol{s} - \boldsymbol{H}_J^{-1}\boldsymbol{H}_I\boldsymbol{x}^T) \\ &= \boldsymbol{s}.\end{aligned}$$

□

Le coût d'une itération est dominé par l'énumération $C = n(n-k)^2(\log^2 q) + \binom{k}{p}(q-1)^p = O(\binom{k}{p}(q-1)^p)$ et non plus par le pivot de Gauss.

La proba de succès est dans le pire cas ($|\boldsymbol{e}| = w$) :

$$\begin{aligned}\mathbb{P}_{\text{succès}} &\geq \frac{\binom{n-k}{w-p}(q-1)^{w-p}\binom{k}{p}(q-1)^p}{\binom{n}{w}(q-1)^w}\max\{1, |\mathrm{DecSynd}(\boldsymbol{H},\boldsymbol{s},w)|\} \\ &\geq \frac{\binom{n-k}{w-p}\binom{k}{p}}{\binom{n}{w}}\max\{1, |\mathrm{DecSynd}(\boldsymbol{H},\boldsymbol{s},w)|\} \\ &\geq \frac{\binom{n-k}{w-p}\binom{k}{p}}{\binom{n}{w}}\max\{1, q^k\frac{\binom{n}{w}(q-1)^w}{q^n}\} \\ &\geq \binom{n-k}{w-p}\binom{k}{p}\max\{\frac{1}{\binom{n}{w}}, \frac{(q-1)^w}{q^{n-k}}\} \\ &\geq \binom{n-k}{w-p}\binom{k}{p}\frac{1}{\min\{\binom{n}{w}, \frac{q^{n-k}}{(q-1)^w}\}} \\ &\geq \binom{n-k}{w-p}\binom{k}{p}\frac{(q-1)^w}{\min\{\binom{n}{w}(q-1)^w, q^{n-k}\}}.\end{aligned}$$

On ne gagne jamais plus qu'un facteur polynomial sur Prange (le coût du pivot) :

$$\begin{aligned}T_{LB} &= C \cdot \frac{1}{\mathbb{P}_{\text{succès}}} \\ &= \left(n(n-k)^2(\log^2 q) + \binom{k}{p}(q-1)^p\right)\frac{\min\{\binom{n}{w}(q-1)^w, q^{n-k}\}}{\binom{n-k}{w-p}\binom{k}{p}(q-1)^w} \\ &= \left(\frac{n(n-k)^2(\log^2 q)}{\binom{k}{p}} + (q-1)^p\right)\frac{\min\{\binom{n}{w}(q-1)^w, q^{n-k}\}}{\binom{n-k}{w-p}(q-1)^w} \\ &\geq \frac{\min\{\binom{n}{w}(q-1)^w, q^{n-k}\}}{\binom{n-k}{w-p}(q-1)^w} \\ &\geq \frac{\min\{\binom{n}{w}(q-1)^w, q^{n-k}\}}{\binom{n-k}{w}(q-1)^w} \\ &\geq \frac{1}{n(n-k)^2(\log^2 q)}T_{Prange}.\end{aligned}$$

La complexité de Lee-Brickell est donc (pivot dominé par l'énumération), pour $w = \lfloor \omega n \rfloor$ et $\rho = \lfloor \frac{p}{n} \rfloor$:

$$O\left(\frac{\min\{\binom{n}{w}(q-1)^w, q^{n-k}\}}{\binom{n-k}{w-p}(q-1)^{w-p}}\right).$$

La complexité asymptotique de Lee-Brickell est $\tilde{O}(q^{n\alpha})$ pour

$$\alpha := \min\{h_q(\omega), 1-R\} - (1-R)h_q\left(\frac{\omega-\rho}{1-R}\right).$$

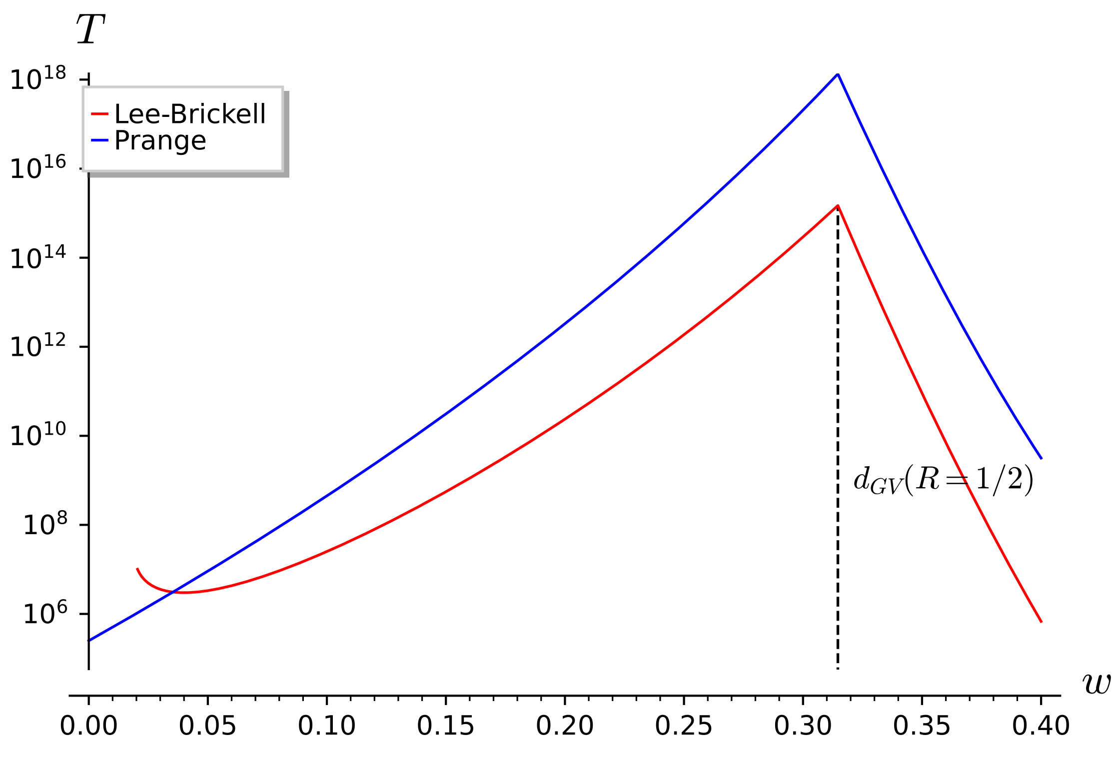

FIGURE 2 – Complexité asymptotique de Lee-Brickell pour un rendement $R = \frac{1}{2}$ et $p = 2$. La distance de Gilbert-Varshamov est la plus difficile à décoder. On observe bien le gain d'un facteur polynomial $O(n^3)$ avec l'échelle logarithmique : pour de petites valeurs de w l'amélioration se dégrade.

Quelle est la valeur optimale de p ?

On suppose que $R = \frac{1}{2}$ et n très grand constant. Étudions les variations de la fonction

$$f(w) = \frac{1}{2}h_q\left(\frac{2(w-p)}{n}\right).$$

Sauf pour w très petit ou très grand, et $q = 2$, $p = 2$ est optimal.

1.3 Algorithme de Stern

L'idée est de combiner Lee-Brickell et le décodage par paradoxe des anniversaires. On effectue une élimination gaussienne partielle.

Décodage par le paradoxe des anniversaires : Problème : trouver au plus w colonnes de $\boldsymbol{H}$ dont la somme vaut $\boldsymbol{s}$ sur $\mathbb{F}_q$.

Solution : scinder $\boldsymbol{H}$ en deux parties égales et énumérer les deux ensembles

$$\mathcal{L}_1 = \left\{\boldsymbol{s} - \boldsymbol{H}_1\boldsymbol{e}_1^T \,:\, wt(\boldsymbol{e}_1) = \frac{w}{2}\right\}$$

et

$$\mathcal{L}_2 = \left\{\boldsymbol{H}_2\boldsymbol{e}_2^T \,:\, wt(\boldsymbol{e}_2) = \frac{w}{2}\right\}.$$

Si $\mathcal{L}_1 \cap \mathcal{L}_2 \neq \varnothing$, on a des solutions $\boldsymbol{s} - \boldsymbol{H}\boldsymbol{e}_1^T - \boldsymbol{H}_2\boldsymbol{e}_2^T = 0$.

Données : $\boldsymbol{H} \in \mathbb{F}_q^{(n-k)\times n}$, $\boldsymbol{s} \in \mathbb{F}_q^{n-k}$, $w \in [\![0, n]\!]$.
Résultat : $\boldsymbol{e}_{\in}\mathbb{F}_q^n$ tel que $\boldsymbol{H}\boldsymbol{e}^T = \boldsymbol{s}$ et $|\boldsymbol{e}| \leq w$.

$X = \{\boldsymbol{x} \in \mathbb{F}_q^{k/2} \,:\, |\boldsymbol{x}| = \frac{p}{2}\}$
$\mathcal{L}_1, \mathcal{L}_2$ des tables de hachage, indexées par des éléments de $\mathbb{F}_q^{n-k}$
répéter
  Choisir un ensemble d'information $I \subseteq [\![1, n]\!]$, $|I| = k$, et son complémentaire J
  si $\boldsymbol{H}_J$ *est inversible* **alors**
    $\bar{\boldsymbol{s}} \leftarrow \boldsymbol{H}_J^{-1}\boldsymbol{s}$
    $\bar{\boldsymbol{P}} \leftarrow \boldsymbol{H}_J^{-1}\boldsymbol{H}_I$
    On extrait les deux sous-matrices $\bar{\boldsymbol{P}} = [\bar{\boldsymbol{P}}_1\bar{\boldsymbol{P}}_2]$
    pour tous les $\boldsymbol{x} \in X$ **faire**
      $\mathcal{L}_1[\bar{\boldsymbol{s}} - \bar{\boldsymbol{P}}_1\boldsymbol{x}^T] \leftarrow \boldsymbol{x}$
      $\mathcal{L}_2[\bar{\boldsymbol{P}}_2\boldsymbol{x}^T] \leftarrow \boldsymbol{x}$
    fin
    Choisir un ensemble L de l positions
  fin
jusqu'à $\exists(\boldsymbol{y}, \boldsymbol{x}_1), (\boldsymbol{y}, \boldsymbol{x}_2) \in \mathcal{L}_1 \times \mathcal{L}_2, |\boldsymbol{y}_{[\![1,n]\!]\setminus L}| \leq w - p$;
retourner $\boldsymbol{e}$ *tel que* $\boldsymbol{e}_I = \boldsymbol{x}$ *et* $\boldsymbol{e}_J = \boldsymbol{y}^T$ *(pour une clé* $\boldsymbol{y}$ *obtenu à partir de* $\boldsymbol{x}$, $\mathcal{L}$ *est une table de hachage).*

Algorithme 3 : Algorithme de Stern

L'algorithme est correct, et la preuve est analogue à celle de Lee-Brickell (même calcul avec $\boldsymbol{e}_I = \boldsymbol{x}_1 + \boldsymbol{x}_2$).

On se place dans le pire des cas ($|\boldsymbol{e}| = w$). Une solution est trouvée à la fin d'une itération si $\boldsymbol{e}_I$ est de poids p (événement E_1), que de plus le poids de $\boldsymbol{e}_I$ soit distribué de manière égale en $\boldsymbol{x}_1$ et $\boldsymbol{x}_2$ (événement E_2), et enfin que l positions de $\boldsymbol{e}_J$ de poids $w - p$ soient nulles (événement E_3).

On obtient que la probabilité de trouver une solution spécifique est, les événements étant indépendants :

$$
\begin{aligned}
\mathbb{P}_{\text{succès}} &\geq \mathbb{P}(E_1 \wedge E_2 \wedge E_3) \\
&\geq \mathbb{P}(E_1)\mathbb{P}(E_2)\mathbb{P}(E_3) \\
&\geq \frac{\binom{k}{p}\binom{n-k}{w-p}}{\binom{n}{w}} \frac{\binom{k/2}{p/2}^2}{\binom{k}{p}} \frac{\binom{n-k-(w-p)}{l}}{\binom{n-k}{l}} \quad \text{car les puissances de } (q-1) \text{ se simplifient} \\
&\geq \frac{\binom{k/2}{p/2}^2\binom{n-k-l}{w-p}}{\binom{n}{w}} \quad \text{car } \frac{\binom{n-k}{w-p}\binom{n-k-(w-p)}{l}}{\binom{n-k}{l}} = \binom{n-k-l}{w-p} \text{ en explicitant les factorielles} \\
&\geq \frac{\binom{k/2}{p/2}^2\binom{n-k-l}{w-p}}{\binom{n}{w}}.
\end{aligned}
$$

Donc la probabilité de trouver n'importe quelle solution est

$$
\begin{aligned}
\mathbb{P}_{\text{succès}} &\geq \frac{\binom{k/2}{p/2}^2\binom{n-k-l}{w-p}}{\binom{n}{w}} \max\{1, |\mathrm{DecSynd}(\boldsymbol{H}, \boldsymbol{s}, w)|\} \\
&\geq \frac{\binom{k/2}{p/2}^2\binom{n-k-l}{w-p}}{\binom{n}{w}\min\{1, \frac{q^{n-k}}{\binom{n}{w}(q-1)^w}\}}.
\end{aligned}
$$

Le coût d'une itération est dominé par la recherche de collision (parcours des clés d'une des listes $\mathcal{L}_1$, linéaire en la taille de $\mathcal{L}_1$) et les multiplications de matrices : $C = \binom{k/2}{p/2}(q-1)^{p/2}(k/2)^3$. Il faut aussi ajouter la probabilité d'obtenir une collision

On obtient donc comme complexité temporelle

$$
\begin{aligned}
T_{Stern} &= C \cdot \frac{1}{P_{Stern}} \\
&= \binom{k/2}{p/2}(q-1)^{p/2}(k/2)^3 \frac{\binom{n}{w}(q-1)^w \min\{1, \frac{q^{n-k}}{\binom{n}{w}(q-1)^w}\}}{\binom{k/2}{p/2}^2(q-1)^p\binom{n-k-l}{w-p}(q-1)^{w-p}} \\
&= O\Big(k^3 \frac{\min\{\binom{n}{w}(q-1)^w, q^{n-k}\}}{\binom{k/2}{p/2}(q-1)^{p/2}\binom{n-k-l}{w-p}(q-1)^{w-p}}\Big).
\end{aligned}
$$

Sa complexité asymptotique est, pour $w = \lfloor \omega n \rfloor$, $p = \lfloor \rho n \rfloor$, $l = \lfloor \lambda n \rfloor$, $\tilde{O}(q^{n\alpha})$ avec

$$
\alpha := \min\{h_q(\omega), 1-R\} - \frac{R}{2}h_q(\frac{\rho}{R}) - (1-R-\lambda)h_q(\frac{\omega-\rho}{1-R-\lambda}).
$$

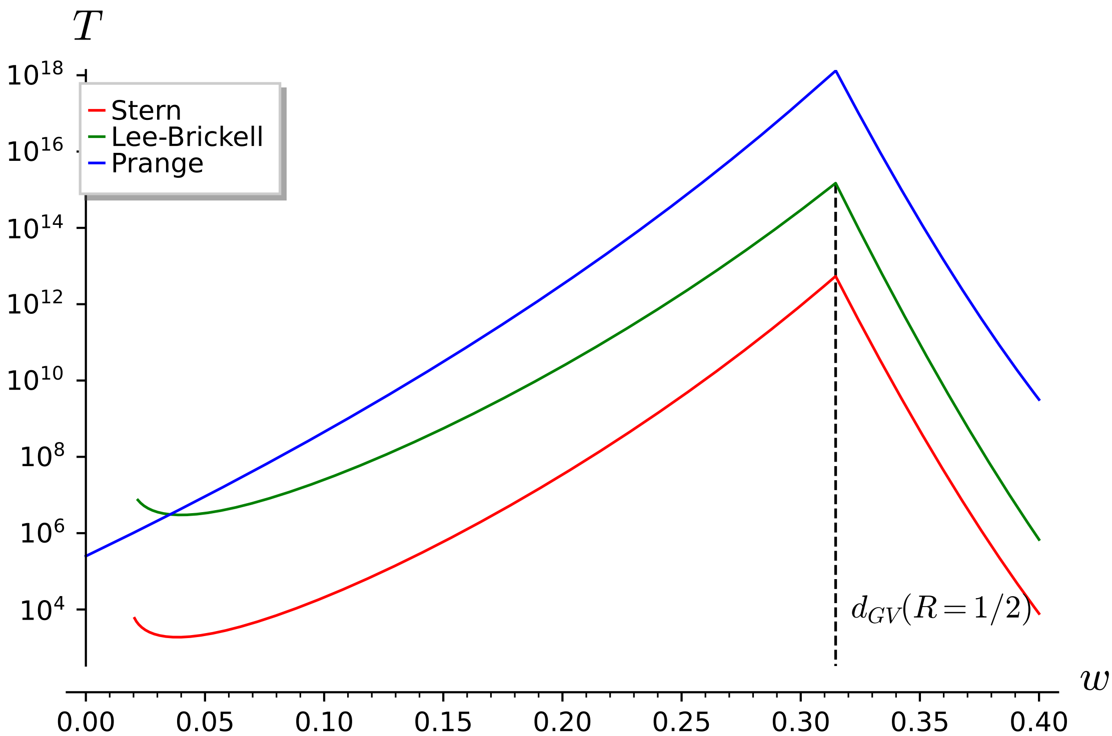

FIGURE 3 – Complexité asymptotique de Stern pour un rendement $R = \frac{1}{2}$.

2 Décodage en liste