Wednesday, April 11, 2018 2:04 PM

Metric Embeddings :-

n-points, distance function 'd'

$d(x,x) = 0$

$d(y,x) = d(x,y) \geq 0 \quad \forall x, y$

$d(x,y) + d(y,z) \geq d(x,z)$

Examples of metrics

① n Points in a plane

Extend to higher dimensions $(d \geq 2)$

$(x_1, y_1) \quad (x_2, y_2) \quad \cdots \quad (x_n, y_n)$
$\quad p_1 \qquad\quad p_2 \quad - \quad - \qquad p_n$

$d(p_i, p_j) = |x_i - x_j| + |y_i - y_j|$
$\hookrightarrow l_1$ - metric

$d(p_i, p_j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$
$\hookrightarrow l_2$ - metric

$d(p_i, p_j) = \left((|x_i - x_j|)^p + (|y_i - y_j|)^p \right)^{1/p}$
$\hookrightarrow l_p$ - norm

$d(p_i, p_j) = \max\left(|x_i - x_j|, |y_i - y_j| \right)$
$\hookrightarrow l_\infty$ - norm

Natural Example of non-metric

$d(p_i, p_j) = (x_i - x_j)^2 + (y_i - y_j)^2$
(Won't satisfy Δ^{le} inequality) !!

Graph Metric

$G = (V, E)$ & have $w_e \geq 0$ are weights on edges,

$$\boxed{d(u,v) = \text{shortest path w.r.t weights } \{w_e\}}$$

→ Where do metrics appear in algorithm design ?

(A1) Metrics arise as part of the problem!

(A2) Metrics appear as part of the solution technique

What are Metric Embeddings ?

How well do one family of metrics embed into another family of metrics ?

Example

d be any n-point metric (X, d)

embed isometrically into l_∞ metrics embed isometrically into ℓ_∞ metrics

$$\{p_1, p_2, \dots, p_n\} = X \xrightarrow{\quad} \{q_1, q_2, \dots, q_n\}$$

points in $\mathbb{R}^n$

$$\boxed{d(p_i, p_j) = \|q_i - q_j\|_\infty}$$

$$\begin{array}{c|ccccc} & a & b & c & d & e \\ \hline a & 0 & 3 & 8 & 6 & 1 \\ b & 3 & 0 & 9 & 7 & 2 \\ c & 8 & 9 & 0 & 2 & 7 \\ d & 6 & 7 & 2 & 0 & 5 \\ e & 1 & 2 & 7 & 5 & 0 \end{array}$$

→ 5 point metric space.

$$\begin{array}{c|ccccc} & a & b & c & d & e \\ \hline v_a & (0, & 3, & 8, & 6, & 1) \\ v_b & (3, & 0, & 9, & 7, & 2) \\ v_c & (8, & 9, & 0, & 2, & 7) \\ v_d & (6, & 7, & 2, & 0, & 5) \\ v_e & (1, & 2, & 7, & 5, & 0) \end{array}$$

Example

$$v_a - v_d = (-6, -4, 6, 6, -4)$$

$$\|v_a - v_d\| = 6$$

$\forall\ p, q$

(A) $\|v_p - v_q\|_\infty \geq d(p,q)$ ← look at p^{th} column of v_p & v_q.

(B) $\|v_p - v_q\|_\infty \leq d(p,q)$

look at any coordinate 'r'

$$\underbrace{|d(p,r) - d(q,r)|}_{\text{follows because}} \leq d(p,q)$$

Δ^{le} ineq.:
$$\begin{cases} d(p,r) - d(q,r) \leq d(p,q) \ \& \\ d(q,r) - d(p,r) \leq d(p,q) \end{cases}$$

Provided enough dimensions, ℓ_∞ metrics are all powerful to capture any finite metric !!

?

How powerful are ℓ_1-metrics? for any metric (X, d), can we embed it into vectors s.t. $d(x,y) = \|f(x) - f(y)\|_1$??

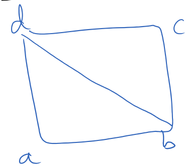
SAME QN for ℓ_2 ??

Can ℓ_1-metrics embed into ℓ_2-metrics?

$\ell_2 \longrightarrow \ell_1$?

Thm :- $\exists$ metrics (X, d) which don't embed isometrically into ℓ_1

$\exists$ metrics (X, d) which don't embed isometrically into l_2.



$$\begin{array}{c|cccc} & a & b & c & d \\ \hline a & 0 & 1 & 2 & 1 \\ b & 1 & 0 & 1 & 2 \\ c & 2 & 1 & 0 & 1 \\ d & 1 & 2 & 1 & 0 \end{array}$$

Next qn: Can we 'approximately' embed all metrics into l_1 (or) l_2 ?

Thm [BOURGAIN]

Any metric can be embedded into an l_1 metric with distortion $O(\log n)$

$\Downarrow$

$\exists \; f : X \to \mathbb{R}^d$ s.t

$$d(x,y) \le \|f(x) - f(y)\|_1 \le O(\log n)\, d(x,y) \quad \forall x, y \in X.$$

Similarly, any metric also embeds into l_2 with distortion $O(\log n)$

In both cases, d will be reasonally small! $(\approx O(\log^2 n))$

Cool property

→ Originally needed $\frac{n^2}{2}$ entries to remember the metric!

→ If we are OK with approximately remember distances, can do it with $n \cdot d$ entries $\approx n \cdot \log^2 n$

LB : Can't improve on $O(\log n)$ distortion!
(ie) $\exists$ metrics s.t. best possible embedding into l_1 has distortion $\Omega(\log n)$.

Dimensionality Reduction

Given n points in $\mathbb{R}^d$, ~~can I red~~
& l_2 - norm metric between them.

$v_1, v_2 \ldots v_n$

$$d(v_i, v_j) = \|v_i - v_j\|_2$$

& suppose d is very very large

Theorem: $\forall \; \epsilon > 0$, $\exists$ a mapping $f : \mathbb{R}^d \to \mathbb{R}^L$

Theorem: $\forall\ \epsilon > 0$, $\exists$ a mapping $f : \mathbb{R}^d \to \mathbb{R}^L$ s.t. $\forall\ i, j$

$$(1-\varepsilon)\,\|v_i - v_j\|_{2} \le \|f(v_i) - f(v_j)\|_2 \le (1+\varepsilon)\,\|v_i - v_j\|_2$$

where $L = \dfrac{O(\log n)}{\varepsilon^2}$.

entries per point comes down from 'd' $\longrightarrow L = \dfrac{O(\log n)}{\varepsilon^2}$

$\boxed{\text{Johnson Lindenstrauss Lemma}}$

Bourgain's Theorem

Given a metric (X, d) on n points, we can always embed it into ℓ_1 with distortion $\le O(\log n)$ in $\mathbb{R}^d$ where $d = \Theta(\log^2 n)$

(ie) $\exists\ f : X \to \mathbb{R}^d$ s.t.

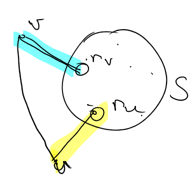
$$d(u,v) \le \|f(u) - f(v)\|_1 \le O(\log n)\, d(u,v) \quad \forall\ u, v \in X$$

We'll carefully choose many subsets $S \subseteq X$, & create one co-ordinate in the embedding

$$f_S(v) = d(v, S)$$

$$f(v) = \left(f_{S_1}(v),\ f_{S_2}(v),\ \dots\ f_{S_D}(v) \right)$$

$$\boxed{d(v, S) = \min_{u \in S} d(u,v)}$$



Claim ①

$$\|f(u) - f(v)\|_1 \le D\, d(u,v)$$

Proof: $\forall\ S \subseteq X$

$$|d(u,S) - d(v,S)| \le d(u,v)$$

$$\Leftrightarrow \quad d(u,S) - d(v,S) \le d(u,v) \ \ \& \quad d(v,S) - d(u,S) \le d(u,v)$$

We'll set $D = \Theta(\log^2 n)$

Theorem ②

We can carefully choose $S_1, S_2, \dots S_D$ s.t.

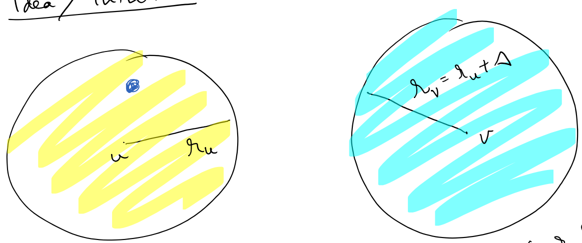
$$\|f(u) - f(v)\|_1 \ge \Omega(\log n) \cdot d(u,v)$$

How to choose $S_1, S_2, \dots, S_D$?

Idea / Intuition

How to ...

Idea / Intuition



$\forall u \in X, r > 0$, $B(u, r) = \{v : d(u,v) \le r\}$

Consider $B(u, r_u)$ & $B(v, r_v)$ s.t $r_v = r_u + \Delta$

& suppose $S \subseteq X$ is st $S \cap B(u, r_u) \neq \phi$ $\Rightarrow f_S(u) \le r_u$

& $S \cap B(v, r_v) = \phi \Rightarrow f_S(v) \ge r_v = r_u + \Delta$

In this case, $|f_S(u) - f_S(v)| \ge \Delta$

Algorithm

For $i = 1, 2, \ldots \cancel{\theta}\,100(\log n)$

For $\ell = 0, 1, \ldots \log n$

Add u to S_ℓ^i w.p $\frac{1}{2^\ell}$ independently for all u

$\forall v : f_{S_\ell^i}(v) = d(v, S_\ell^i)$

$$f(v) = \left(\{f_{S_\ell^i}(v)\}_{\substack{1 \le i \le \theta(\log n) \\ 0 \le \ell \le \log n}}\right)$$

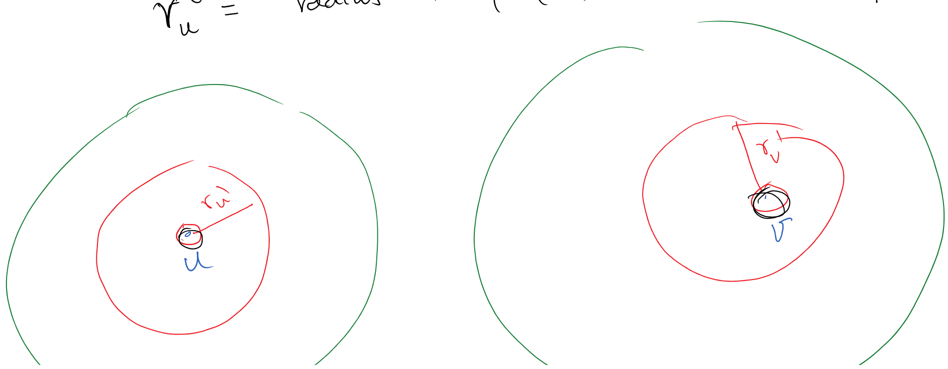
WTS

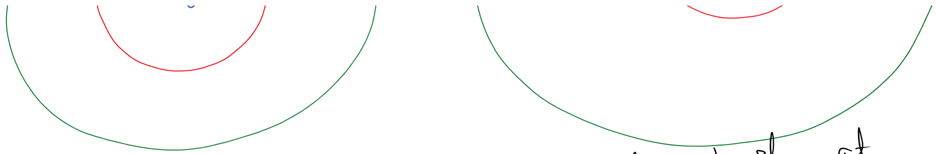
for each $u \neq v$

$\|f(u) - f(v)\|_1 \ge \Omega(\log n) \cdot d(u, v)$.

For every (u, v),

r_u^ℓ = radius s.t $|B(u, r_u^\ell)| \approx 2^\ell$ points in it





Let t be the level at which the 2 balls overlap

$$\Rightarrow \boxed{r_u^t + r_v^t \geq d(u,v)}$$

$$\boxed{\text{WLOG, let } r_u^t \geq \frac{d(u,v)}{2}}$$

CLAIM 1

$\forall\, 0 \leq l < t$, we'll show the following

$$\| f_{S_l}(u) - f_{S_l}(v) \|_1 \geq 0.1 \left(r_u^l - r_u^{l-1} \right)$$

with probability ≥ 0.1

Call S_l good if $| f_{S_l}(u) - f_{S_l}(v) | \geq 0.1 \left(r_u^l - r_u^{l-1} \right)$

it's bad otherwise.

$$E[\#\text{ good sets } S_l \text{ over } i = 1, 2, \ldots 100(\log n)]$$
$$\geq \cancel{0.5}\ 0.1 \times 100 \cdot \log n$$
$$= 10 \log n$$

Chernoff Bounds

$\Rightarrow$ with probability $\geq 1 - \frac{1}{n^6}$,

the # good sets $S_l \geq 5 \log n$.

For each good set, coordinate value

$$| f_{S_l}(u) - f_{S_l}(v) | \geq 0.1 \left(r_u^l - r_u^{l-1} \right)$$

$\Rightarrow$ total co-ordinate value over all good sets S_l

$$\geq 0.5 \log n \left(r_u^l - r_u^{l-1} \right)$$

Sum over all $0 \leq l < t$

$$\text{total } l_1 \text{ dist} \geq 0.5 \log n \left(r_u^t - r_u^0 \right)$$
$$\geq 0.5 \log n \, \frac{d(u,v)}{2}$$

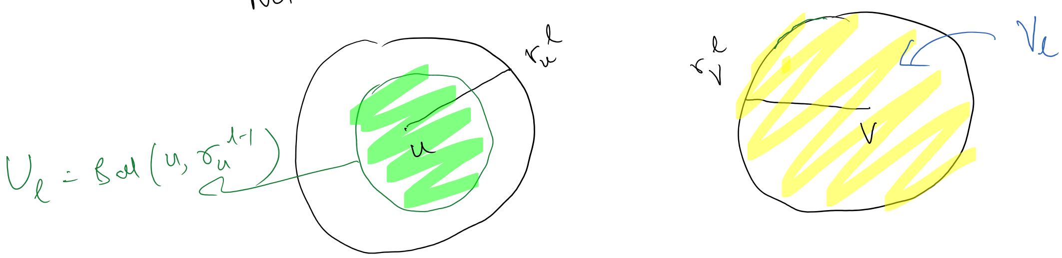
Suffices to show Claim 1

Call a set S_ℓ good if

$$|f_{S_\ell}(u) - f_{S_\ell}(v)| \geq 0.1\,(r_u^\ell - r_u^{\ell-1})$$

Need to show that picking S_ℓ randomly by including each vertex with prob $\frac{1}{2^\ell}$ gives a good set w.p. ≥ 0.1

Note that $\forall\ \ell \in [0, t)$



Case ①

$r_u^\ell < r_v^\ell$: Consider $U_\ell = Ball(u, r_u^{\ell-1})$

& $V_\ell = Ball(v, r_v^\ell)$

Suppose S_ℓ intersects U_ℓ but completely misses V_ℓ

$$\Rightarrow \quad d(u, S_\ell) \leq r_u^{\ell-1} \quad \&$$

$$d(v, S_\ell) > r_v^\ell \geq r_u^\ell$$

$$\Rightarrow \|f_{S_\ell}(u) - f_{S_\ell}(v)\|_1 \geq r_u^\ell - r_u^{\ell-1}$$

$\Pr(S_\ell \text{ intersects } U_\ell \ \&\ S_\ell \text{ misses } V_\ell)$

$$= \Pr(S_\ell \text{ misses } V_\ell)\,(1 - \Pr(S_\ell \text{ misses } U_\ell))$$

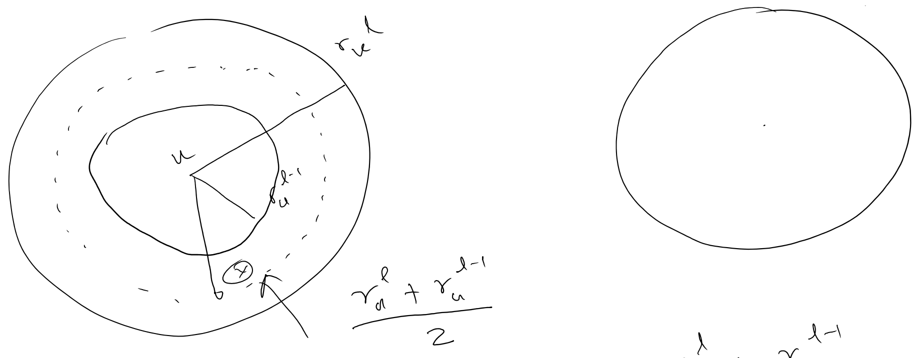
$$= \left(1 - \frac{1}{2^\ell}\right)^{2^\ell}\left[1 - \left(1 - \frac{1}{2^\ell}\right)^{2^{\ell-1}}\right]$$

$$\approx \frac{1}{e}\left[1 - \frac{1}{e}\right]$$

$$\geq 0.1$$

Case ② : $r_u^\ell \geq r_v^\ell$

r_v^ℓ



Case 2a $\quad r_v^l \geq \frac{r_u^l + r_u^{l-1}}{2}$

Set $U_l = \text{Ball}(u, r_u^{l-1})$ &
$V_l = \text{Ball}(v, r_v^l)$

We'll get that if
$S_l \cap V_l = \phi$ & $S_l \cap U_l \neq \phi$
$|d(u, S_l) - d(v, S_l)| \geq \frac{r_u^l - r_u^{l-1}}{2}$

Case 2b
$r_v^l \leq \frac{r_u^l + r_u^{l-1}}{2}$

In this case, set $U_l = \text{Ball}(u, r_u^l)$
$V_l = \text{Ball}(v, r_v^l)$

Same proof to show that
if $S_l \cap U_l = \phi$ & $S_l \cap V_l \neq \phi$
$|d(u, S_l) - d(v, S_l)| \geq r_u^l - r_v^l$
$\geq \frac{r_u^l - r_u^{l-1}}{2}$

Proof of Claim 1 QED

Similar proof technique can be used for embedding

$\boxed{\text{Frechet Embedding}}$

~~There~~ is
Every metric can embed into l_1 with distortion $\leq O(\log n)$
... and ...

Bourgain's Theorem:

~~There is~~ Every metric can embed into ℓ_1 $\le O(\log n)$

(ie) $\exists \; f : X \to \mathbb{R}^d$ st

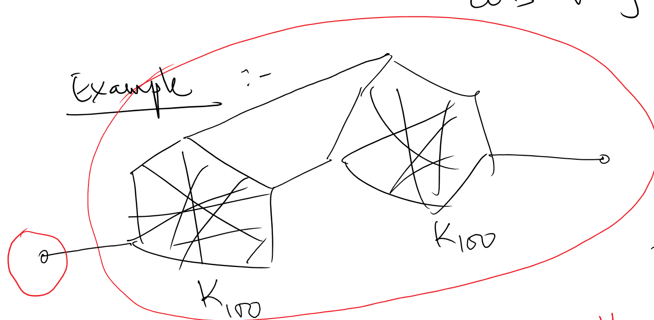
$\forall (u,v) \quad d(u,v) \le \|f(u) - f(v)\|_1 \le O(\log n) \cdot d(u,v)$

Sparsest Cut :-

Given a graph $G = (V, E)$,

high level goal: find a "balanced" partitioning which cuts very few edges.

Example :-



$n = 102$

min-cut will return a very unbalanced cut

Sparsest Cut :-

$$\min_{S} \frac{|E(S, \bar{S})|}{|S|\,|\bar{S}|}$$

Very impt problem in real world !!

$\downarrow$

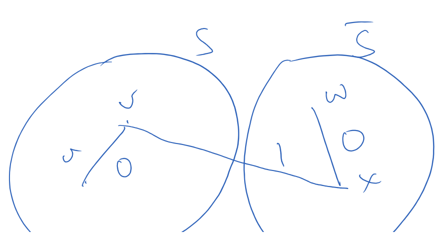
NP-hard !!

Every cut naturally defines a metric !! (cut-metric)

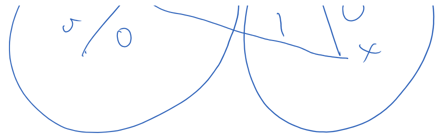
$\forall S : S \subseteq V$,

let $\delta_S(u,v) = 1$ if $|S \cap \{u,v\}| = 1$

$= 0$ otherwise



$d(u,v) + d(v,w) \geq d(u,w)$



$\geq d(u,w)$

Special type of metric !!

Observe: $|E(S,\bar{S})| = \sum_{(u,v)\in E} \delta_S(u,v)$

$$|S||\bar{S}| = \sum_{\substack{u\in V\\ v\in V}} \delta_S(u,v)$$

Sparsest cut equivalent to

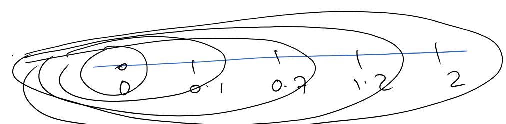
$$\min_{\delta\,:\,\delta \text{ is a cut-metric}} \frac{\sum_{(u,v)\in E} \delta(u,v)}{\sum_{u,v\in V} \delta(u,v)}$$

Same probs! So still NP-hard.

$\boxed{\text{HW}}$ { Cut metrics are effectively the same as l_1 - metrics }

$$= \min_{\delta\,:\,\delta \text{ is } l_1\text{-metric}} \frac{\sum_{(u,v)\in E} \delta(u,v)}{\sum_{u,v\in V} \delta(u,v)}$$

Relax l_1 requirement



$$\min_{\delta\,:\,\delta \text{ is a metric}} \frac{\sum_{(u,v)\in E} \delta(u,v)}{\sum_{u,v\in V} \delta(u,v)}$$

OPT(metric) $\leq$ OPT(l_1 metric) = Sparsest Cut Opt.

$$= \min \sum_{(u,v)\in E} \delta(u,v)$$

$$\sum_{u,v\in V} \delta(u,v) = 1$$

$\forall\, u,v,w \qquad \delta(u,v) + \delta(v,w) \geq \delta(u,w)$

$\delta(u,v) = \delta(v,u)$

$\forall\,(u,v) \qquad \delta(u,v) \geq 0$

LP !!

$\forall (u,v) \quad \ldots$

Bourgains theorem

embed this solution metric $\delta \to l_1$

$l_1 \to$ cut metrics $\to$ good soln.

Overall Result :-

If the embedding has distortion α, then

we can get α- approximation to Sparsest Cut.

$\Rightarrow O(\log n)$ - approximation to Sparsest cut.

$$\min_{\delta:\text{ metric}} \frac{\sum_{(u,v)\in E} \delta(u,v)}{\sum_{u,v} \delta(u,v)}$$

P3

$$\min_{\delta} \sum_{(u,v)\in E} \delta(u,v)$$

$$\sum_{u,v} \delta(u,v) = 1$$

$$\delta(u,v) + \delta(v,w) \geq \delta(u,w)$$

$$\delta(u,v) \geq 0$$

P4

$$OPT(P4) = OPT(P3) \leq OPT(P1)$$

Solve P4 optimally and get δ^*

Use Bourgain's Theorem & embed $\delta^* \to l_1$ with distortion $\log n$.

$$\forall (u,v) : \delta^*(u,v) \leq \|f(u) - f(v)\|_1 \leq O(\log n)\, \delta^*(u,v)$$

$$\Rightarrow \sum_{(u,v)\in E} \|f(u) - f(v)\|_1 \leq O(\log n) \sum_{(u,v)\in E} \delta^*(u,v)$$

$$\& \sum_{u,v\in V} \|f(u) - f(v)\|_1 \geq \sum_{u,v\in V} \delta^*(u,v) = 1$$

$\Rightarrow$ Sparsity of my l_1 - solution

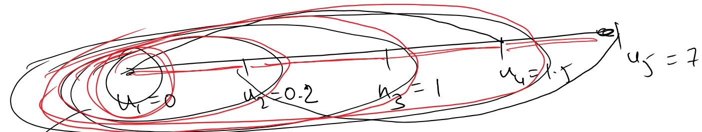
$\Rightarrow$ Sparsity of my l_1 - solution

$$= \boxed{\frac{\sum_{(u,v)\in E} \|f(u) - f(v)\|_1}{\sum_{u,v \in V} \|f(u) - f(v)\|_1}} \leq O(\log n)\, OPT(l_4)$$

$\Rightarrow$ I've found a sol^n to P2 with cost $\leq O(\log n)\, OPT(l_4)$

$\Rightarrow$ Can find sol^n to P1 with cost $\leq O(\log n)\, OPT(l_4)$

Equivalence between l_1 & cut metrics:



$$\sum_{(u,v)\in E} \delta(u,v) = \sum_{(u,v)\in E} \|f(u) - f(v)\|_1$$

Write it as

$$\delta = 0.2\, \delta_{S_1} + 0.8\, \delta_{S_2} + 0.5\, \delta_{S_3} + 5.5\, \delta_{S_4}$$

$\forall (u,v)$, $\delta(u,v) = 0.2\, \delta_{S_1}(u,v) + 0.8\, \delta_{S_2}(u,v) + 0.5\, \delta_{S_3}(u,v) + 5.5\, \delta_{S_4}(u,v)$