For a data set $S$ of $n$ points the clique complex $K_\Sigma$ is hard to compute when $n$ is large. $K_\Sigma$ has

- $n$ vertices
- edges determined by the $\frac{1}{2}n(n-1)$ distances $d(x,y)$, $x,y \in S$.
- 2-simplexes determined by the $\frac{1}{6}n(n-1)(n-2)$ triples $x,y,z \in S$.
- and so on.

Yesterday's textons example had $|S| = 8\,000\,000$.

# Mapper Clustering

Again based on the nerve $N\mathcal{U}$ of an open cover of an unknown population space $X$.

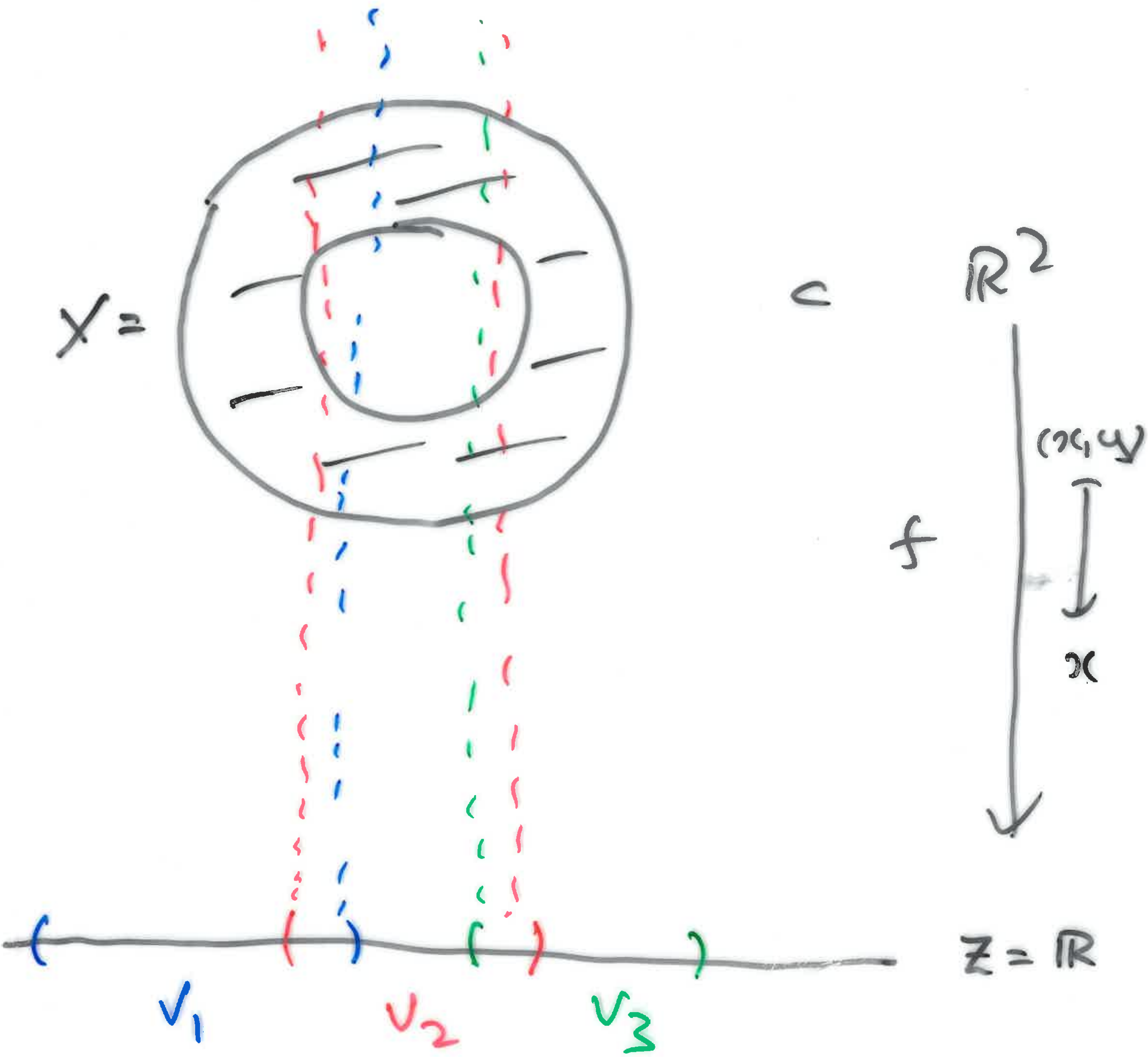
To construct such a cover $\mathcal{U}$ we could choose:

1) a continuous map $f : X \to Z$ where $Z$ is some known (parameter) space, and where we only know $f(x)$ for $x$ in some finite sample $S \subset X$.

2) an open cover

$$\mathcal{V} = \{V_\alpha\}_{\alpha \in A}$$

of $Z$




$A = \{1, 2, 3\}$

For $\alpha \in A$ the pre-image

$$U_\alpha = f^{-1}(V_\alpha)$$

is open since $f$ is continuous.

In general $U_\alpha$ may have many connected components.

In the example 

$$U_1 = f^{-1}(V_2)$$

In general

$$U_\alpha = U_{\alpha 1} \sqcup U_{\alpha 2} \sqcup \cdots \sqcup U_{\alpha n_\alpha},$$

with $U_{\alpha i}$ a connected component.

So we get an open cover

$$\mathcal{U} = \{U_{\alpha,i}\}_{\substack{\alpha\in A \\ 1\le i\le n_\alpha}}$$

of $X$.

The nerve $N\mathcal{U}$ is our "approximation" to $X$.

If $f$ and $\mathcal{V}$ are well chosen, we would hope to get $|N\mathcal{U}|$ homotopy equivalent to $X$.

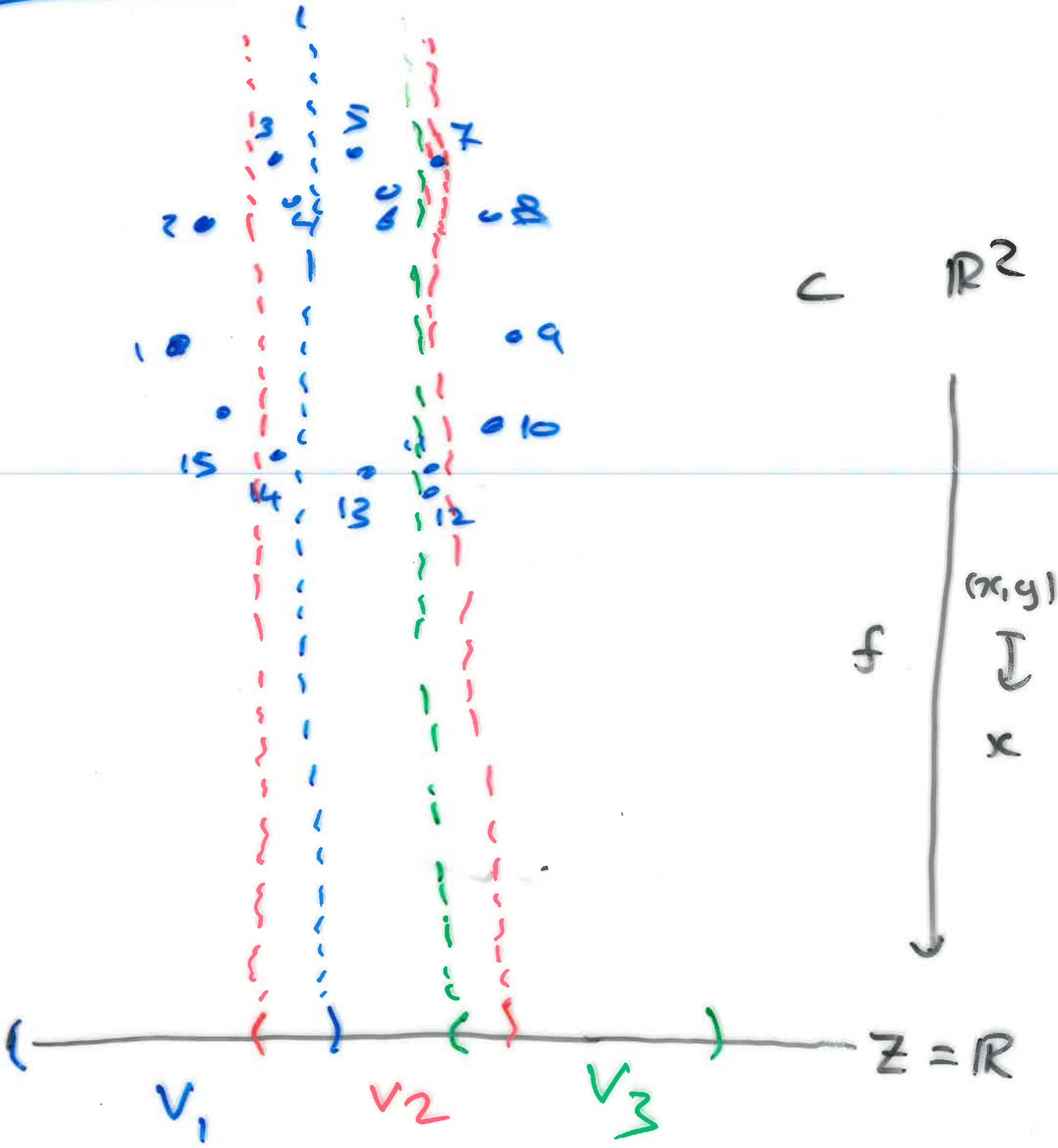
<u>However</u> we can't construct the $U_{\alpha,i}$ since we don't know $X$. We approximate $U_{\alpha,i}$ by

$$f^{-1}(V_\alpha) \cap S = S_{\alpha,1} \sqcup S_{\alpha,2} \sqcup \cdots \sqcup S_{\alpha,n_\alpha}$$

$$\left( f^{-1}(V_\alpha) = U_{\alpha,1} \sqcup \cdots \sqcup U_{\alpha n_\alpha} \right)$$

The connected component $U_{\alpha,i}$ is replaced by a cluster $S_{\alpha,i}$ obtained by applying <u>any</u> clustering algorithm to $f^{-1}(V_\alpha) \cap S$.

<u>Example</u>



$f^{-1}(V_1) \cap S = \{1, 2, 3, 4, 14, 15\} = S_{11}$

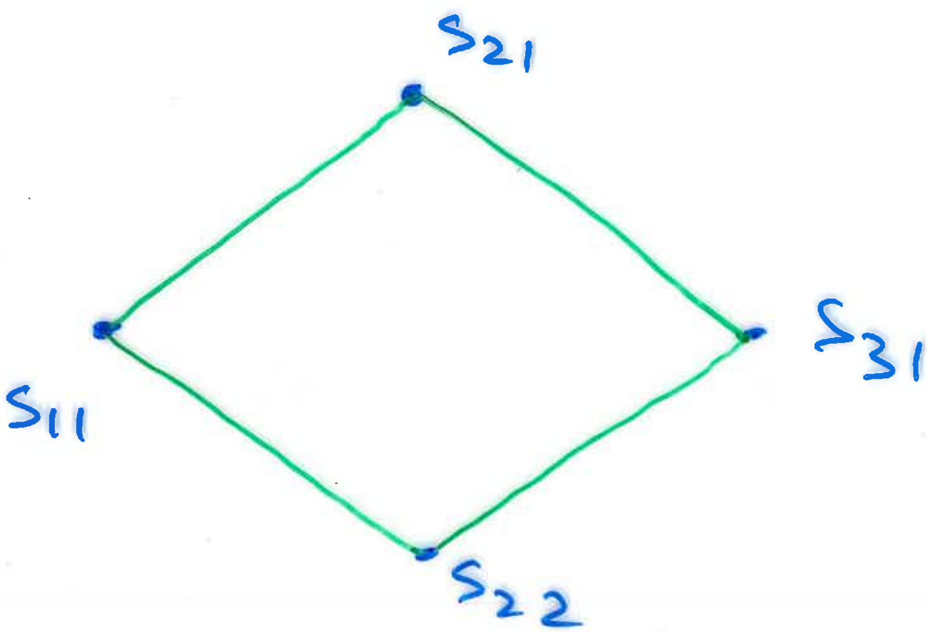
$f^{-1}(V_2) \cap S = \underbrace{\{3, 4, 5, 6, 7\}}_{S_{2,1}} \cup \underbrace{\{11, 12, 13, 14\}}_{S_{2,2}}$

$f^{-1}(U_3) \cap S = \underbrace{\{7, 8, 9, 10, 11, 12\}}_{S_{3,1}}$

So $S_{11} \cup S_{21} \cup S_{22} \cup S_{31} = S$

or $\{S_{\alpha i}\}$ is a cover of $S$.

The nerve $N\{S_{\alpha i}\}$ is




So this nerve is homeomorphic to a circle $S^1$.