

# SPLOCIS – Extending SPLOMs to a Scatterplot Cube with Interactable Shadows for Immersive Analysis in Virtual Reality (Supplementary Material)

Melanie Derksen\*  
TU Dortmund University,  
Germany

Viktoria Diekel†  
TU Dortmund University,  
Germany

Torsten Kuhlen‡  
RWTH Aachen University,  
Germany

Mario Botsch §  
TU Dortmund University,  
Germany

Tim Weissker¶  
RWTH Aachen University,  
Germany

## SPLOCIS INTERACTION FORMALIZATION

In this section, we present the mechanisms through which users can interactively make specific 3D plots visible within the SPLOCIS application. Exemplary interactions are also shown in the supplementary video.

Initially, the whole SPLOC is invisible (see Figure 1 (a)). By selecting one or more attributes per axis by either selecting the labels directly or by selecting 2D plots in the SPLOMs, the user specifies the conditions the 3D plots should fulfill. Only those 3D plots turn visible that meet the strictest conditions.

This process can be seen as calculating volumetric intersections – selecting a label for one spatial axis defines a plane paved by 3D plots (see Figure 1 (b)); selecting a 2D plot defines an elongated cuboid (see Figure 1 (c)). As soon as there is an intersection between two or more volumes, their intersections are displayed – meaning only the plots matching all of their conditions (see Figure 1 (d)). If there is a set of plots matching one condition that does not intersect with any other plot set, its whole set of plots is displayed (so all 3D plots matching that single condition). One exception is if labels on all three different axes are selected: In this case, their pairwise-intersections are omitted and only the intersection of all three 3D plot planes is shown.

More formally, let  $A_d$  be the set of attributes per room axis  $d \in \{x, y, z\}$  and  $p(a_x, a_y, a_z)$  a 3D plot with attributes  $a_x \in A_x, a_y \in A_y, a_z \in A_z$ . The set of all 3D plots within a SPLOC is then defined as follows:

$$P := \{p(a_x, a_y, a_z) \mid (a_x, a_y, a_z) \in A_x \times A_y \times A_z\}.$$

At the beginning, all plots  $P$  are hidden. By specifying conditions through user input, plots that match the strictest conditions turn visible.

Each condition set by the user results in a subset of the original plot set  $P$  and can be of two kinds: (i) it can either specify one attribute per room axis  $d$ , (ii) or even a pair of attributes for two distinct room axes  $d'$  and  $d''$ , with  $d', d'' \in \{x, y, z\}, d' \neq d''$ .

Regarding (i), the complete set  $C_d$  of plot subsets  $C_d^i \subseteq P$  provided by attribute specification for room axis  $d$  is then defined as follows:

$$C_d := \{C_d^i \mid a_d^i \in A_d\}.$$

\*e-mail: melanie.derksen@tu-dortmund.de

†e-mail: viktoria.diekkel@tu-dortmund.de

‡e-mail: kuhlen@vr.rwth-aachen.de

§e-mail: mario.botsch@tu-dortmund.de

¶e-mail: me@tim-weissker.de

Therefore,  $C_x^i$  specifies the subset of plots matching the single condition of having the attribute  $x_i \in A_x$  on their x-axis. More formally:

$$C_x^i := \{p(a_x^i, a_y, a_z) \mid (a_x^i, a_y, a_z) \in \{a_x^i\} \times A_y \times A_z\}.$$

For  $y$  and  $z$ , with  $y_i \in A_y$  and  $z_i \in A_z$ , it is defined analogously:

$$C_y^i := \{p(a_x, a_y^i, a_z) \mid (a_x, a_y^i, a_z) \in A_x \times \{a_y^i\} \times A_z\},$$

$$C_z^i := \{p(a_x, a_y, a_z^i) \mid (a_x, a_y, a_z^i) \in A_x \times A_y \times \{a_z^i\}\}.$$

Regarding (ii), the combined set  $C_{d'd''}$  of plot subsets  $C_{d'd''}^i \subseteq P$  provided by attribute specification on room axes  $d'$  and  $d''$  is then defined as follows:

$$C_{d'd''} := \{C_{d'd''}^i \mid d'_i \in A_{d'}, d''_i \in A_{d''}, d' \neq d''\}.$$

Therefore,  $C_{xy}^i$  specifies the subset of plots having the attribute  $x_i \in A_x$  on their x-axis and the attribute  $y_i \in A_y$  on their y-axis. More formally:

$$C_{xy}^i := \{p(a_x^i, a_y^i, a_z) \mid (a_x^i, a_y^i, a_z) \in \{a_x^i\} \times \{a_y^i\} \times A_z\}.$$

For the the remaining two attribute combinations the definitions are analogously

$$C_{xz}^i := \{p(a_x^i, a_y, a_z^i) \mid (a_x^i, a_y, a_z^i) \in \{a_x^i\} \times A_y \times \{a_z^i\}\},$$

$$C_{yz}^i := \{p(a_x, a_y^i, a_z^i) \mid (a_x, a_y^i, a_z^i) \in A_x \times \{a_y^i\} \times \{a_z^i\}\}.$$

Let  $\ominus$  be a new “set-of-sets-theoretic-operator” defined as follows:

$$B \ominus D = \bigcup_{B_i \in B} (B_i \ominus D), \text{ with}$$

$$B_i \ominus D = \begin{cases} B_i & \text{if } B_i \cap (\bigcup_{D_j \in D} D_j) = \emptyset, \\ B_i \cap (\bigcup_{D_j \in D} D_j) & \text{otherwise.} \end{cases}$$

Generally speaking, the operator checks for every set  $B_i \in B$  if there is any intersection with one of the sets in  $D$ . If there is no intersection at all, the operator simply gives the original set  $B_i$  as is. If there is an intersection, it gives the union of all intersections between  $B_i$  and each subset  $D_j \in D$ .

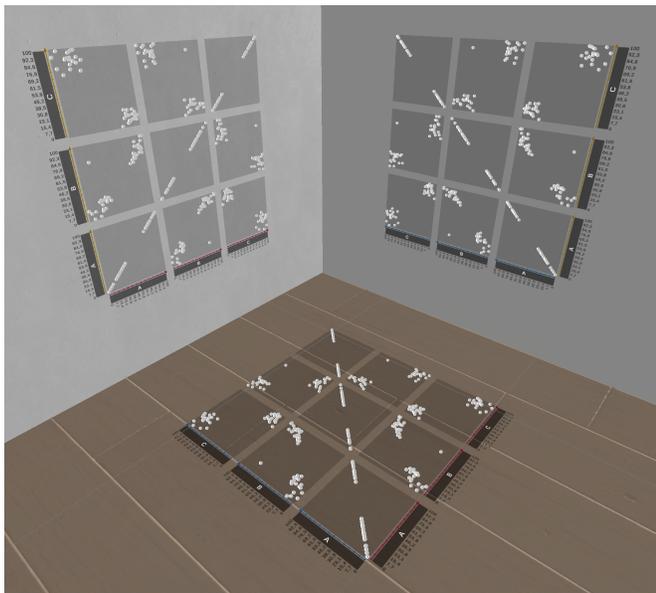
The plot subset  $P'$  of 3D plots to be visualized is provided by taking all conditions and their resulting plot subsets into account, and applying the new operator  $\ominus$  on each set of plot subsets that may affect them – meaning that the other sets may further specify

the plots that already matched the conditions (in other words: that may cause an intersection with the plot subset). More formally:

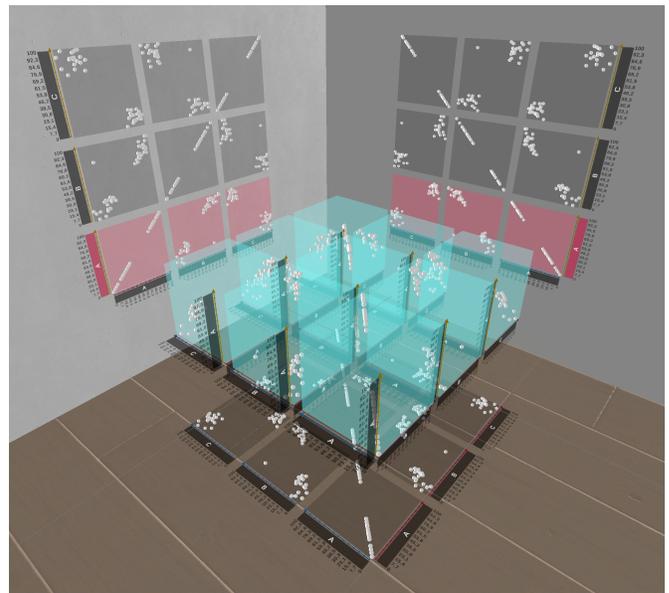
$$\begin{aligned}
 P' = & (C_{xy} \ominus (C_{xz} \cup C_{yz} \cup C_z)) \\
 & \cup (C_{xz} \ominus (C_{xy} \cup C_{yz} \cup C_y)) \\
 & \cup (C_{yz} \ominus (C_{xy} \cup C_{xz} \cup C_x)) \\
 & \cup (C_x \ominus (C_{yz} \cup C_{xy} \cup C_{xz} \cup (C_y \ominus C_z) \cup (C_z \ominus C_y))) \\
 & \cup (C_y \ominus (C_{xz} \cup C_{xy} \cup C_{yz} \cup (C_x \ominus C_z) \cup (C_z \ominus C_x))) \\
 & \cup (C_z \ominus (C_{xy} \cup C_{xz} \cup C_{yz} \cup (C_x \ominus C_y) \cup (C_y \ominus C_x))).
 \end{aligned}$$

## USER STUDY TASKS

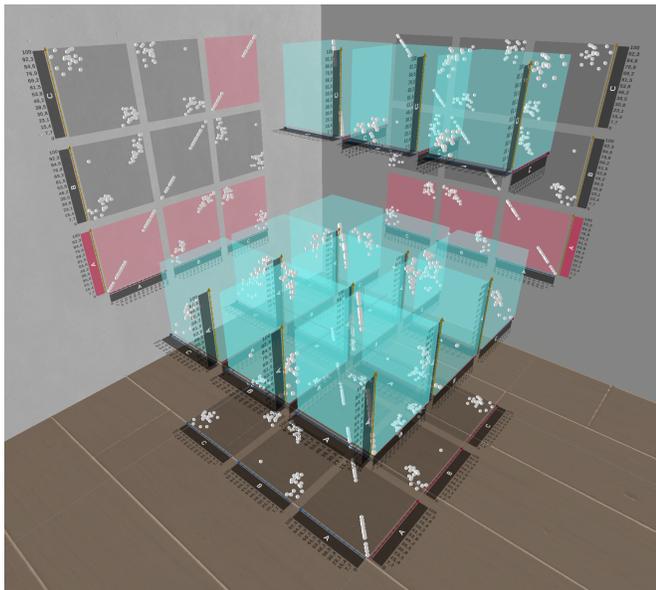
In Table 1 the formulations per task type are listed. In Tables 2 and 3, we list details about the tasks the participants had to solve during our user study. They contain the task number, the name of the respective dataset (which are in our supplementary material), the type of task and the valid plot attribute combinations. We only list the attributes the 3D plot must have in general; it does not matter on which axis the attributes are located: E.g., an entry of (A, C, E) in the tables covers the plots  $p(A, C, E)$ ,  $p(A, E, C)$ ,  $p(C, A, E)$ ,  $p(C, E, A)$ ,  $p(E, A, C)$ , and  $p(E, C, A)$ .



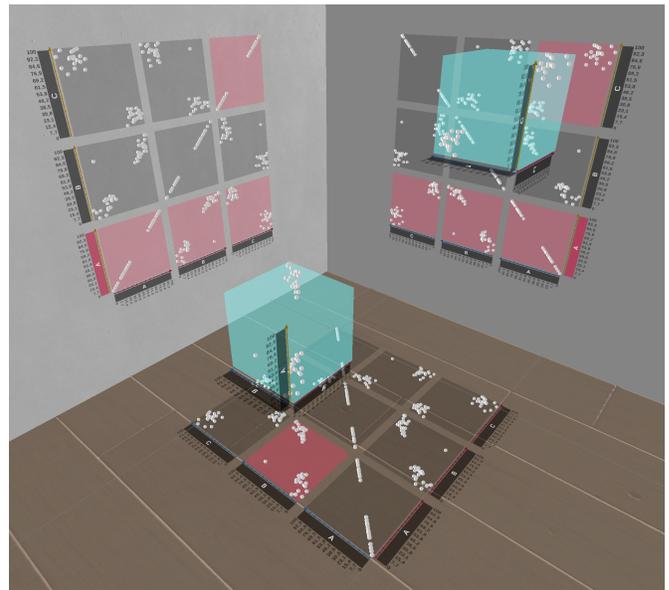
(a)



(b)



(c)



(d)

Figure 1: Three SPLOMs build the SPLOCIS' information and interaction basis. Initially, no 3D plot is visible (a). 3D Plots can be turned visible (highlighted in light blue) by selecting the labels located at the borders of the SPLOMs (b) and/or by selecting individual 2D plots (selected labels and 2D plots are highlighted in red) in the SPLOMs directly (c). This way, the user specifies step-by-step which attributes are of interest, filtering which 3D plots should remain visible for further investigation (d).

Table 1: Here, the specific formulations per task type are listed.

Task	Wording
<b>Attribute</b>	“Put a 3D plot into the box that has following attributes:” – <i>followed by the three attributes.</i>
<b>Correlations</b>	“Put a 3D plot into the box that has three different attributes, all of which show pairwise linear relationships.”
<b>Positive Correlations</b>	“Put a 3D plot into the box that has three different attributes, all of which show pairwise linear positive relationships.”
<b>Clusters</b>	“Put a 3D plot into the box that has three different attributes, in which two clusters are visible (each cluster should contain at least nine data points).”
<b>Outlier</b>	“Put a 3D plot into the box that has three different attributes and includes a clear outlier.”
<b>No Relation</b>	“Put a 3D plot into the box that shows neither linear relationships nor multiple clusters.”

Table 2: This table shows the task number, the name of the respective dataset (in our supplementary material), the type of task and the valid plots are listed. During the tutorial the datasets consisted of four attributes A, B, C, and D. This resulted in a total of 64 3D plots.

Task Nr.	(Dataset Name).csv	Task	Valid Plot Attribute Combinations
1	1_Tut	<b>Attributes: B, C, D.</b>	(B, C, D)
2	1_Tut	<b>Correlations</b>	(A, B, C)
3	-1_Tut	<b>Positive Correlations</b>	(B, C, D)
4	0_Tut	<b>Clusters</b>	(A, B, C), (A, B, D), (A, C, D)
5	-1_Tut	<b>Outlier</b>	(A, B, D), (A, C, D), (B, C, D)
6	0_Tut	<b>No Relation</b>	(B, C, D)

Table 3: This table shows the task number, the name of the respective dataset (in our supplementary material), the type of task and the valid plots are listed. During the actual study the datasets consisted of five attributes A, B, C, D, and E. This resulted in a total of 125 3D plots. Per task type the number of valid plot combinations remained the same in order to have the same level in difficulty.

Task Nr.	(Dataset Name).csv	Task	Valid Plot Attribute Combinations
1	-1	<b>Attributes: A, A, A.</b>	(A, A, A)
2	-1	<b>Attributes: B, B, C.</b>	(B, B, C)
3	-1	<b>Attributes: A, C, E.</b>	(A, C, E)
4	-1	<b>Attributes: B, D, E.</b>	(B, D, E)
5	6	<b>Correlations</b>	(A, B, C), (A, B, E), (A, C, E), (B, C, E)
6	2	<b>Positive Correlations</b>	(B, D, E), (B, C, E), (B, D, E), (C, D, E)
7	5	<b>Correlations</b>	(A, C, D), (A, C, E), (A, D, E), (C, D, E)
8	4	<b>Positive Correlations</b>	(A, B, C), (A, B, D), (A, C, D), (B, C, D)
9	7	<b>Clusters</b>	(A, B, C), (A, B, D), (A, B, E), (A, C, D), (A, C, E), (A, D, E)
10	6	<b>Clusters</b>	(A, B, D), (A, C, D), (A, D, E), (B, C, D), (B, D, E), (C, D, E)
11	5	<b>Clusters</b>	(A, B, C), (A, B, D), (A, B, E), (B, C, D), (B, C, E), (B, D, E)
12	4	<b>Clusters</b>	(A, B, E), (A, C, E), (A, D, E), (B, C, E), (B, D, E), (C, D, E)
13	8	<b>Outlier</b>	(A, B, C), (B, C, D), (B, C, E)
14	3	<b>Outlier</b>	(A, B, D), (A, C, D), (A, D, E)
15	2	<b>Outlier</b>	(A, B, E), (B, C, E), (B, D, E)
16	1	<b>Outlier</b>	(A, B, E), (A, C, E), (A, D, E)
17	0	<b>No Relation</b>	(A, C, E), (B, C, E), (C, D, E)
18	8	<b>No Relation</b>	(A, B, D), (A, C, D), (A, D, E)
19	3	<b>No Relation</b>	(A, B, E), (B, C, E), (B, D, E)
20	1	<b>No Relation</b>	(A, B, C), (B, C, D), (B, D, E)