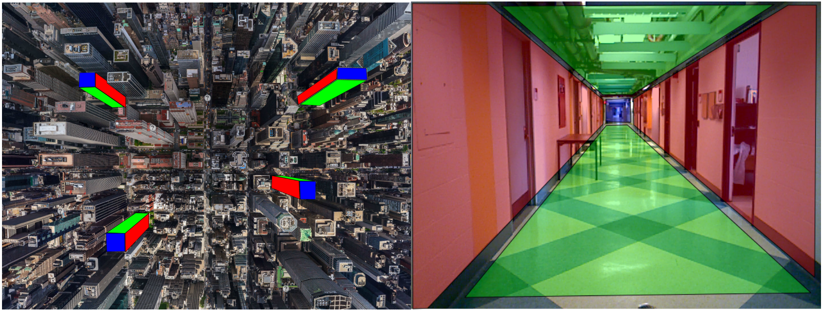
A Mixture of Manhattan Frames: Beyond the Manhattan World

Julian Straub
Massachusetts Institute of Technology
Computer Science and Artificial Intelligence Laboratory

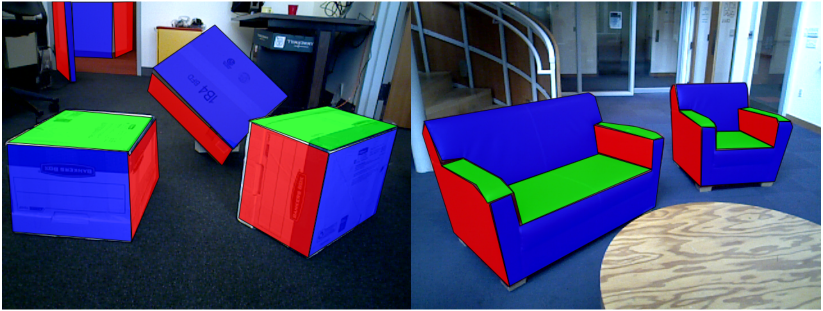
April 9, 2014

Manhattan World (MW) [Coughlan 1999]: All planes in the scene are parallel to one of the three major planes of one common coordinate system.



But a lot of **scenes break the Manhattan-World assumption**.

But a lot of **scenes break the Manhattan-World assumption**.



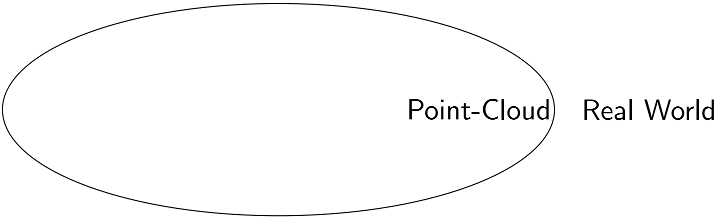
But a lot of **scenes break the Manhattan-World assumption**.

⇒ We describe a scene as a **Mixture of Manhattan Frames** (MMF).

Related Work – Scene Representation

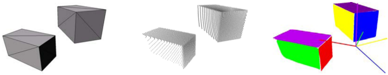
Real World

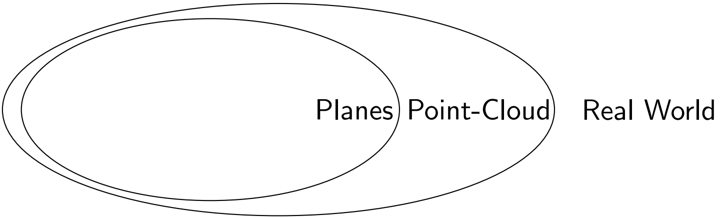
Related Work – Scene Representation



Related Work – Scene Representation

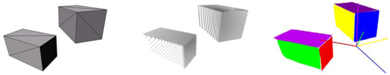
- **3D Plane Segmentation** [Triebel 2005, Stückler 2008, Holz 2011]



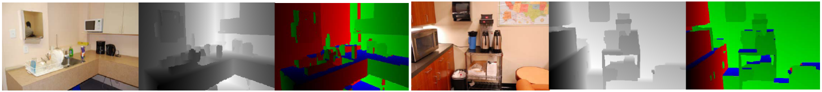


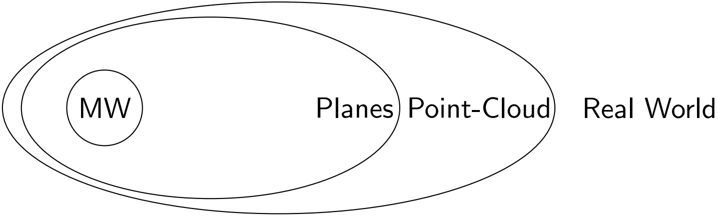
Related Work – Scene Representation

- **3D Plane Segmentation** [Triebel 2005, Stückler 2008, Holz 2011]



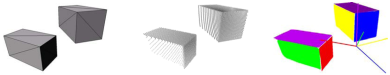
- **Manhattan World in 3D** (MW) [Furukawa 2009, Neverova 2013]



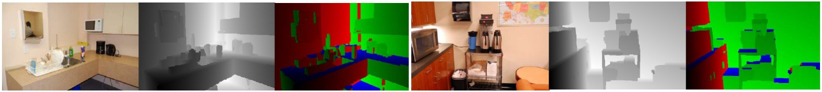


Related Work – Scene Representation

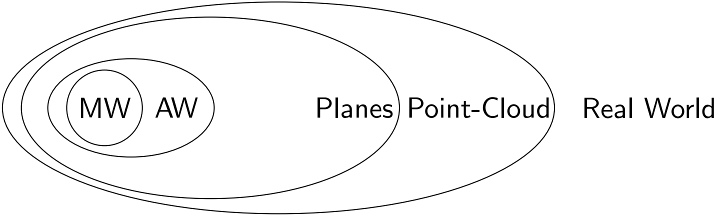
- **3D Plane Segmentation** [Triebel 2005, Stückler 2008, Holz 2011]



- **Manhattan World in 3D** (MW) [Furukawa 2009, Neverova 2013]

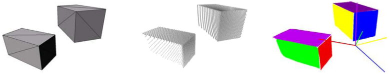


- **Atlanta World** (AW) [Schindler 2004]

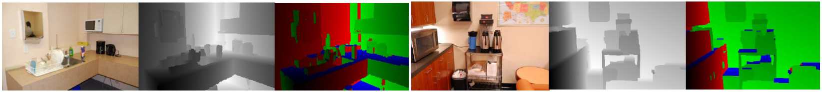


Related Work – Scene Representation

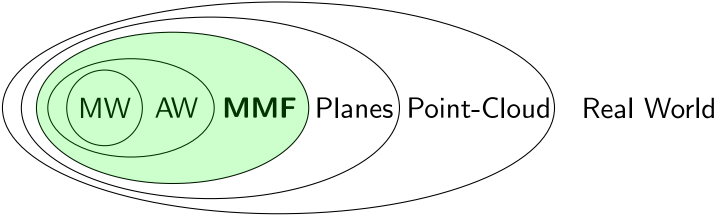
- **3D Plane Segmentation** [Triebel 2005, Stückler 2008, Holz 2011]



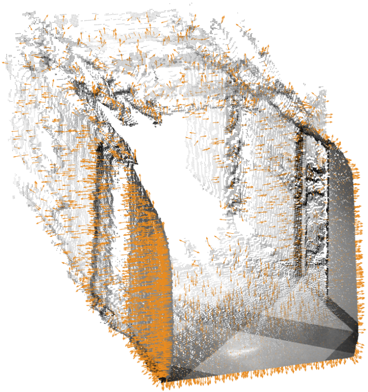
- **Manhattan World in 3D** (MW) [Furukawa 2009, Neverova 2013]



- **Atlanta World** (AW) [Schindler 2004]

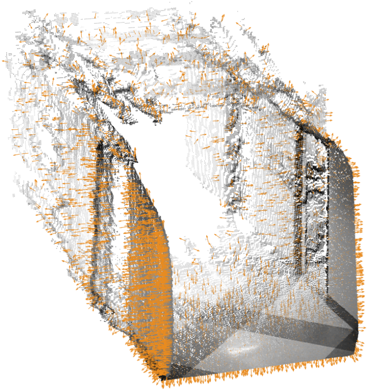


Input Data

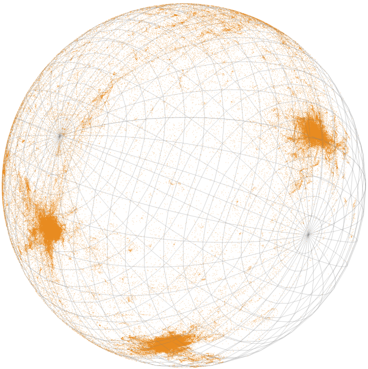


Point-Cloud with Normals

Input Data

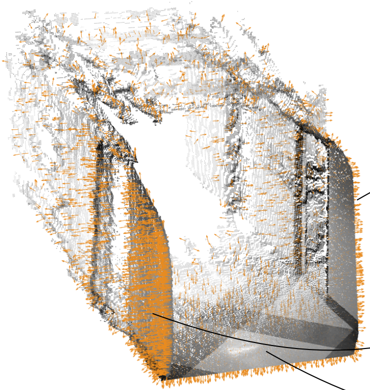


Point-Cloud with Normals

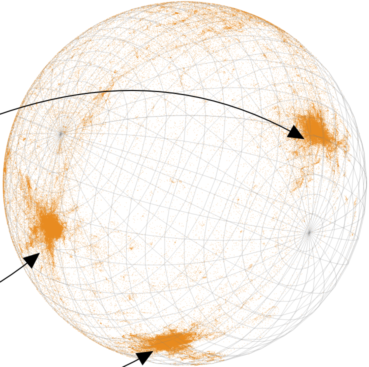


Normals on Unit Sphere

Input Data



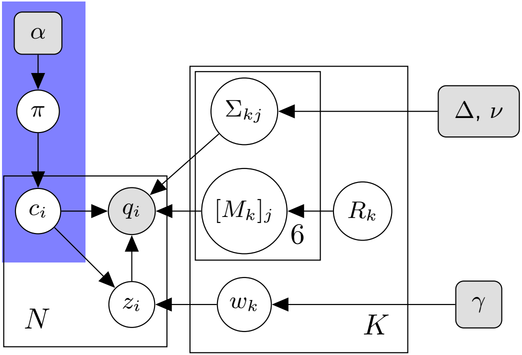
Point-Cloud with Normals

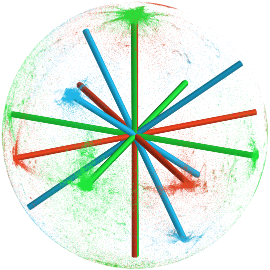


Normals on Unit Sphere

Mixture of Manhattan Frames Model

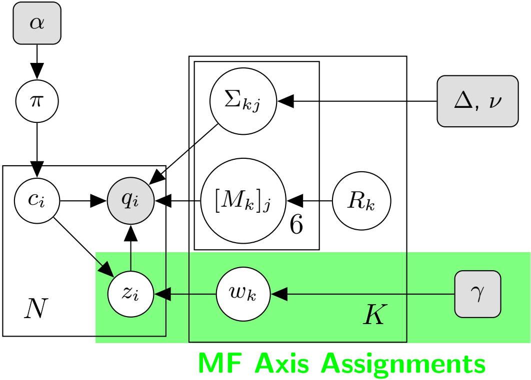
MF Associations

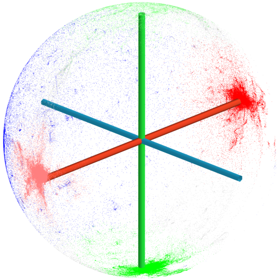




R_k: rotations of MFs; q_i: normals; z_i: associations to MF axes; Σ_{kj} : covariance
$[M_k]_j$: jth axis of kth MF; c_i: association to MFs;

Mixture of Manhattan Frames Model

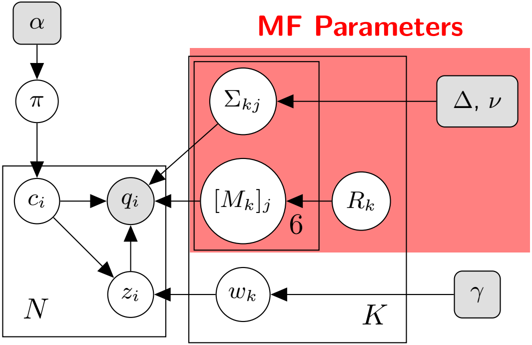


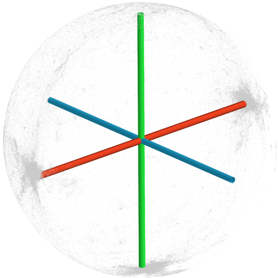


R_k: rotations of MFs; q_i: normals; z_i: associations to MF axes; Σ_{kj} : covariance

$[M_k]_j$: jth axis of kth MF; c_i: association to MFs;

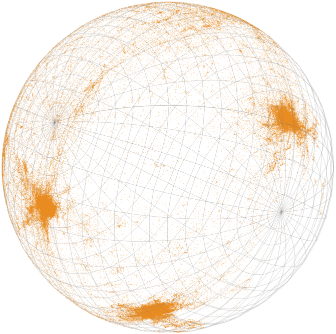
Mixture of Manhattan Frames Model

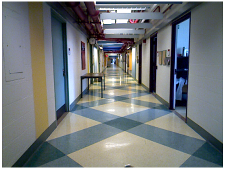




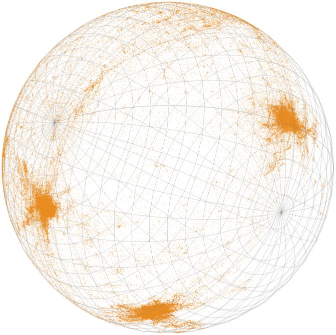
R_k: rotations of MFs; q_i: normals; z_i: associations to MF axes; Σ_{kj} : covariance
$[M_k]_j$: jth axis of kth MF; c_i: association to MFs;

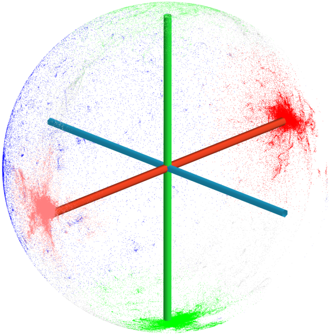
Single Manhattan Frame Scene

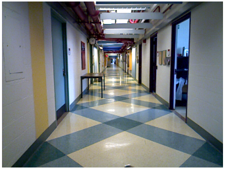




Single Manhattan Frame Scene



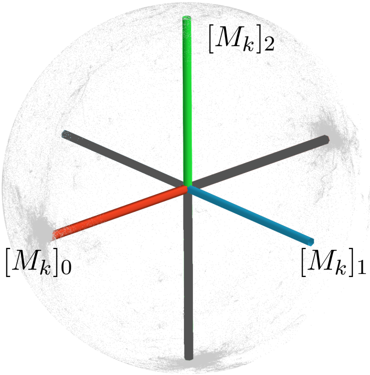




Manhattan Frame Definition

We represent a **Manhattan Frame** M_k with rotation R_k by its six axes:

$$M_k = M(R_k) = [R_k, -R_k]$$

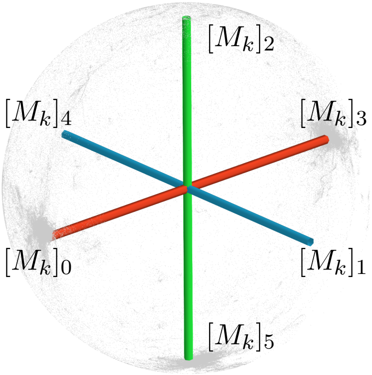


R_k: rotations of MFs;
$[M_k]_j$: jth axis of kth MF;

Manhattan Frame Definition

We represent a **Manhattan Frame** M_k with rotation R_k by its six axes:

$$M_k = M(R_k) = [R_k, -R_k]$$

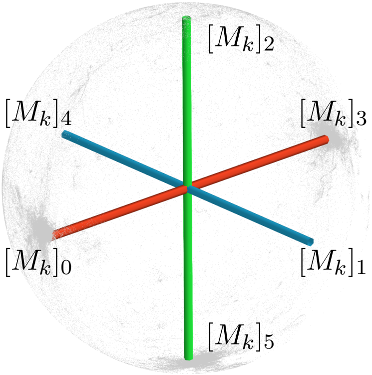


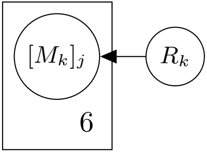
R_k: rotations of MFs;
$[M_k]_j$: jth axis of kth MF;

Manhattan Frame Definition

We represent a **Manhattan Frame** M_k with rotation R_k by its six axes:

$$M_k = M(R_k) = [R_k, -R_k]$$

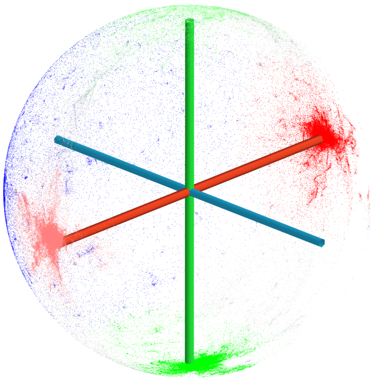


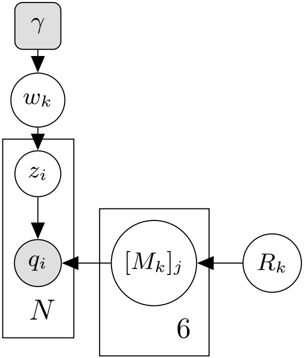


$R_k \sim$ Unif$(\mathrm{SO}(3))$

R_k: rotations of MFs;
$[M_k]_j$: jth axis of kth MF;

Manhattan Frame: Mixture over Axes

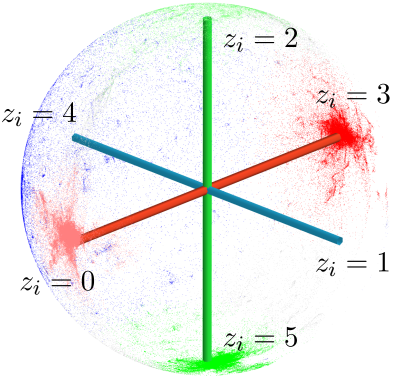


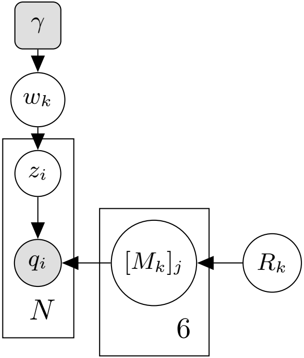


$$
\begin{aligned}
w_k &\sim \text{Dir}(\gamma) \\
z_i &\sim \text{Cat}(w_k)
\end{aligned}
$$

R_k: rotations of MFs; q_i: normals; z_i: associations to MF axes; $[M_k]_j$: jth axis of kth MF;

Manhattan Frame: Mixture over Axes





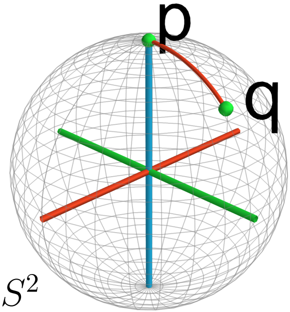
$$\begin{aligned} w_k &\sim \text{Dir}(\gamma) \\ z_i &\sim \text{Cat}(w_k) \end{aligned}$$

R_k: rotations of MFs; q_i: normals; z_i: associations to MF axes; $[M_k]_j$: jth axis of kth MF;

Space of the Normals: S^2

- The **Geodesic Distance**

$$d_G(p,q) = \arccos(p^T q)$$



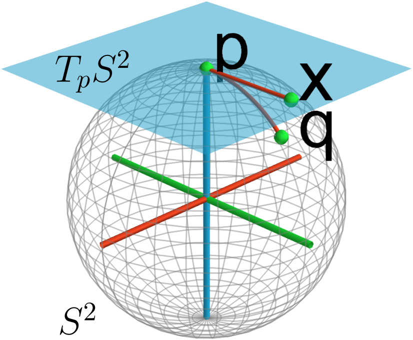
Space of the Normals: S^2

- The **Geodesic Distance**

$$d_G(p, q) = \arccos(p^T q)$$

- The **Riemannian exponential**

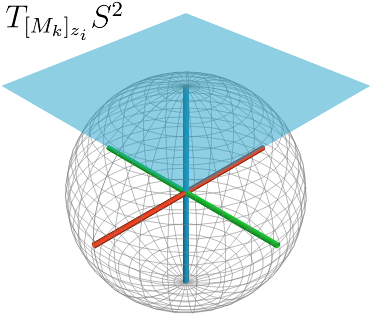
$$\mathrm{Exp}_p : T_pS^2 \to S^2$$

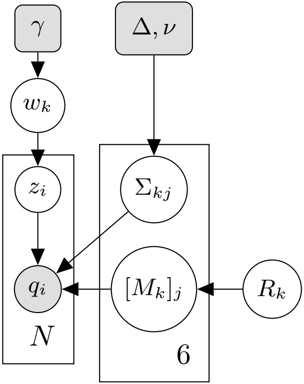


Distribution of Normals on MF Axes

The distribution over normals is defined in the tangent plane $T_{[M_k]_{z_i}} S^2$:

$$q_i \sim \mathrm{Exp}_{[M_k]_{z_i}} \left(\mathcal{N}(0, \Sigma_{k z_i}) \right) \ .$$



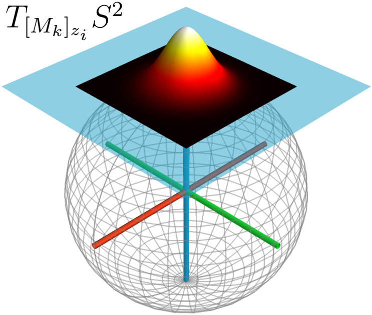


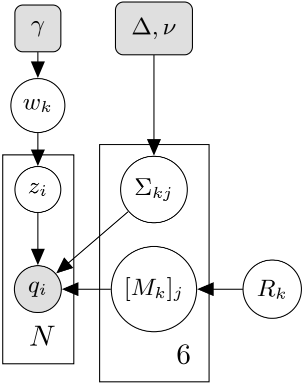
R_k: rotations of MFs; q_i: normals; z_i: associations to MF axes; Σ_{kj} : covariance
$[M_k]_j$: jth axis of kth MF; c_i: association to MFs;

Distribution of Normals on MF Axes

The distribution over normals is defined in the tangent plane $T_{[M_k]_{z_i}} S^2$:

$$q_i \sim \mathrm{Exp}_{[M_k]_{z_i}} \left(\mathcal{N}(0, \Sigma_{k z_i}) \right) \ .$$

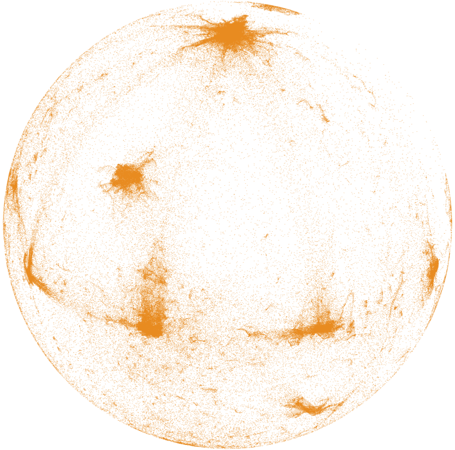




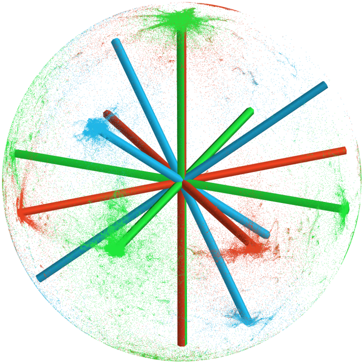
R_k: rotations of MFs; q_i: normals; z_i: associations to MF axes; Σ_{kj} : covariance

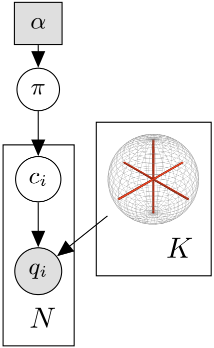
$[M_k]_j$: jth axis of kth MF; c_i: association to MFs;

Scenes with multiple Manhattan Frames



Mixture of Manhattan Frames

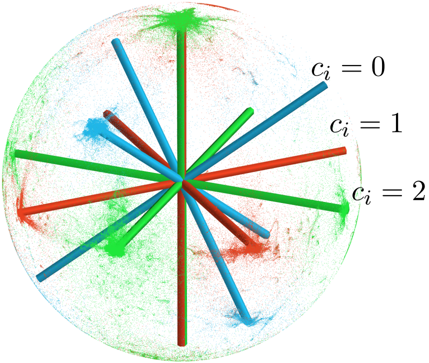


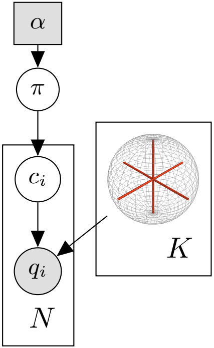


$$\pi \sim \mathrm{Dir}(\alpha)$$
$$c_i \sim \mathrm{Cat}(\pi)$$

R_k: rotations of MFs; q_i: normals; z_i: associations to MF axes; Σ_{kj} : covariance
$[M_k]_j$: jth axis of kth MF; c_i: association to MFs;

Mixture of Manhattan Frames

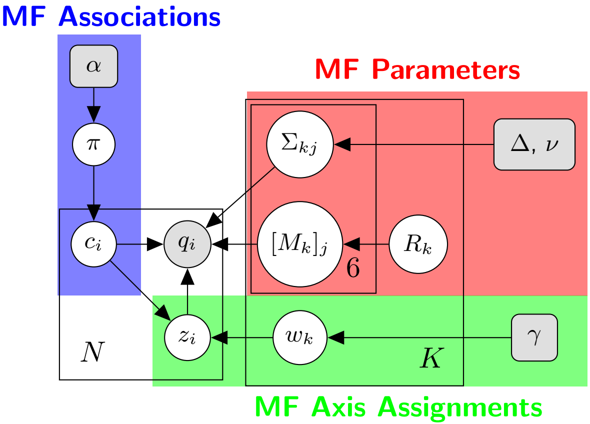




$$\begin{aligned} \pi &\sim \mathrm{Dir}(\alpha) \\ c_i &\sim \mathrm{Cat}(\pi) \end{aligned}$$

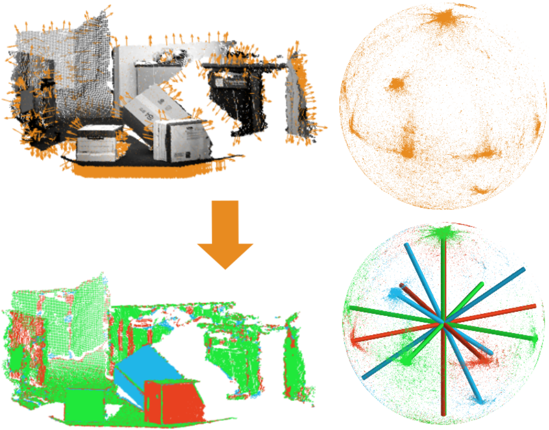
R_k: rotations of MFs; q_i: normals; z_i: associations to MF axes; Σ_{kj} : covariance
$[M_k]_j$: jth axis of kth MF; c_i: association to MFs;

Mixture of Manhattan Frames Model



R_k: rotations of MFs; q_i: normals; z_i: associations to MF axes; Σ_{kj} : covariance
$[M_k]_j$: jth axis of kth MF; c_i: association to MFs;

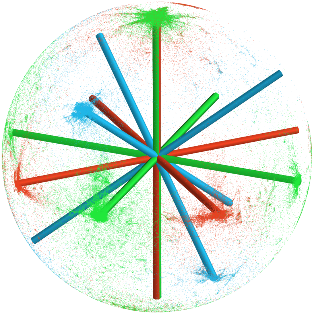
Let's do inference!



⇒ Gibbs sampling with Metropolis-Hastings split-merge proposals

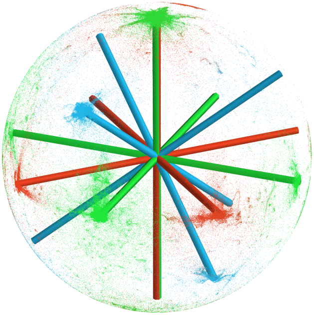
Gibbs Sampling with M-H Split-Merge Proposals

1: sample MF assignments $\mathbf{c}$
2: **for** MF $k \in \{1, \dots, K\}$ **do**
3: sample MF axis assignments $\mathbf{z}_{\{\mathbf{c}=k\}}$
4: sample new MF rotation R_k
5: sample axis covariances $\Sigma_{k\{1\dots6\}}$
6: **end for**
7: propose splits for all MFs
8: propose merges for all MF combinations

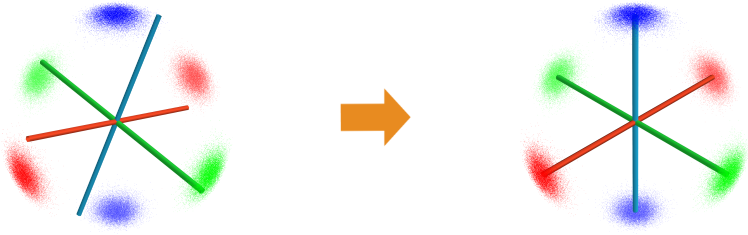


Gibbs Sampling with M-H Split-Merge Proposals

```
sample MF assignments c
for MF k ∈ {1, ..., K} do
    sample MF axis assignments z_{c=k}
    sample new MF rotation R_k
    sample axis covariances Σ_{k{1...6}}
end for
propose splits for all MFs
propose merges for all MF combinations
```



Posterior Distribution for MF Rotations

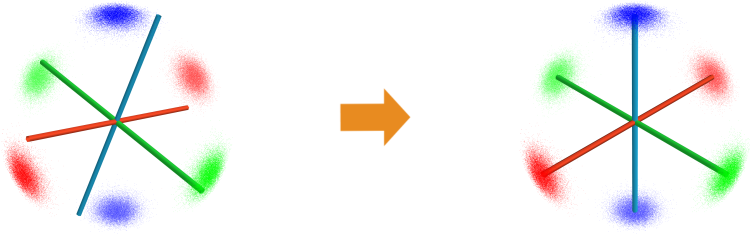


Posterior Distribution for MF Rotations

Approximate the posterior over a MF's rotation R as

$$p(R|\mathbf{z}, \mathbf{c}, \mathbf{q}, R^\star) \approx \mathcal{N}(R; R^\star(R_0, \mathbf{z}, \mathbf{c}, \mathbf{q}), \Sigma) \ , \tag{1}$$

where $R^\star$ is the (locally-) optimal rotation.

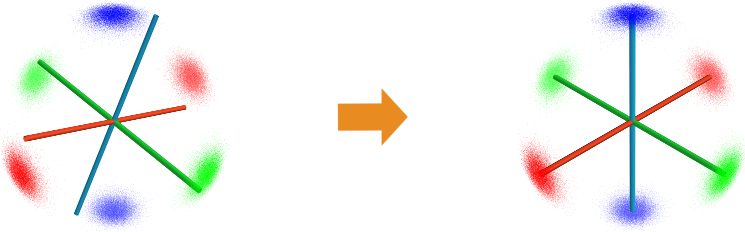


Posterior Distribution for MF Rotations

Approximate the posterior over a MF's rotation R as

$$p(R|\mathbf{z},\mathbf{c},\mathbf{q},R^{\star}) \approx \mathcal{N}(R; R^{\star}(R_0,\mathbf{z},\mathbf{c},\mathbf{q}),\Sigma) \;, \tag{1}$$

where $R^{\star}$ is the (locally-) optimal rotation.



Using the exponential map $\mathrm{Exp}_{R^{\star}}(R) : T_{R^{\star}}\mathrm{SO}(3) \rightarrow \mathrm{SO}(3)$ we sample:

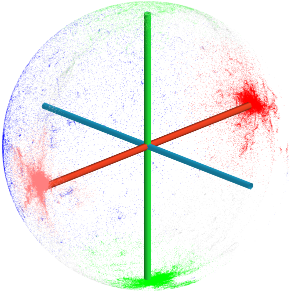
$$R \sim \mathrm{Exp}_{R^{\star}}\left[\mathcal{N}(0,\Sigma_{3\times 3})\right] \;.$$

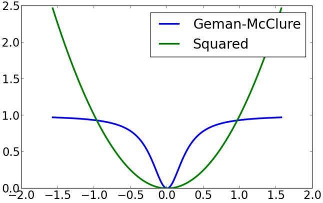
MF Rotation Optimization

$$R_k^\star = \arg\min_{R_k} \sum_{i|c_i=k} \rho(d_G(q_i, [M(R_k)]_{z_i})) \ ,$$

where we use the **Geman-McClure robust function** [Geman 1987]

$$\rho : x \mapsto x^2/(x^2+\sigma^2) \ .$$





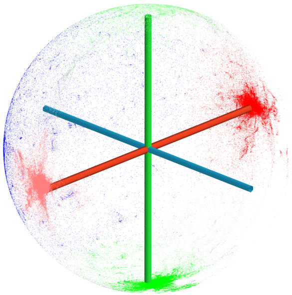
MF Rotation Optimization

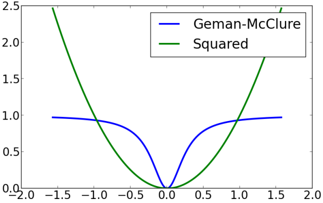
$$R_k^\star = \arg\min_{R_k} \sum_{i|c_i=k} \rho(d_G(q_i, [M(R_k)]_{z_i}))\ ,$$

where we use the **Geman-McClure robust function** [Geman 1987]

$$\rho : x \mapsto x^2/(x^2+\sigma^2)\ .$$

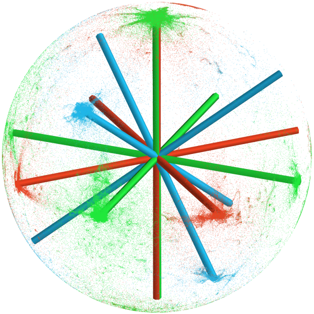
⇒ **conjugate gradient algorithm to optimize** R on $\mathrm{SO}(3)$. [Edelman 1998]



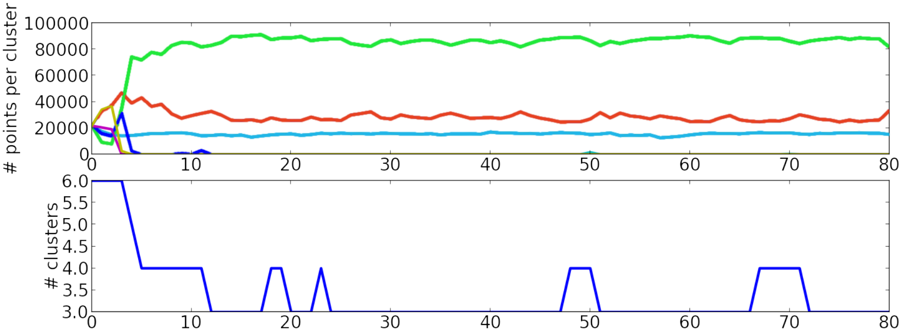


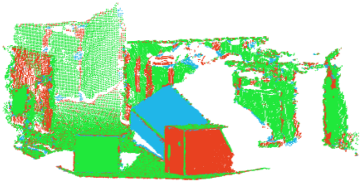
Gibbs Sampling with M-H Split-Merge Proposals

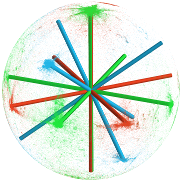
sample MF assignments **c**
for MF $k \in \{1, \ldots, K\}$ **do**
    sample MF axis assignments $\mathbf{z}_{\{\mathbf{c}=k\}}$
    sample new MF rotation R_k
    sample axis covariances $\Sigma_{k\{1 \ldots 6\}}$
end for
propose splits for all MFs
propose merges for all MF combinations



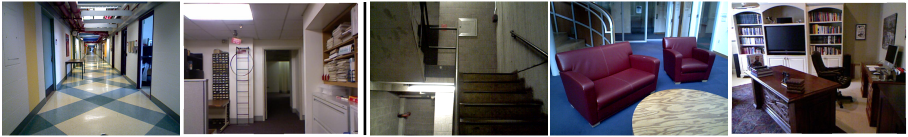
Inference Example



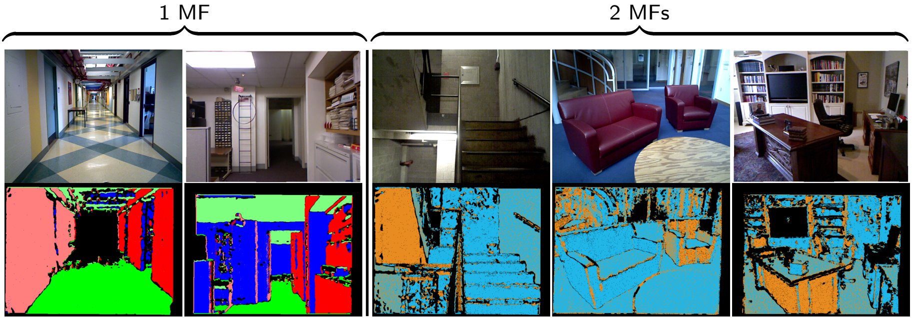




Results: MMF Models from Depth Images

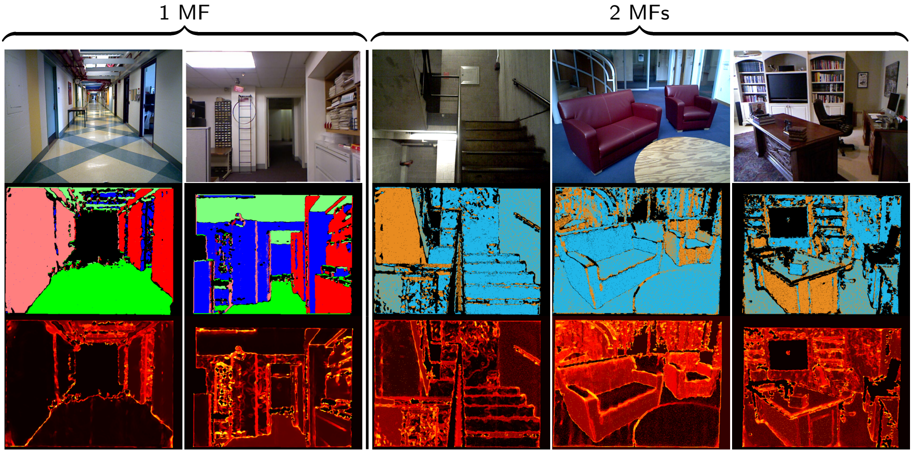


Results: MMF Models from Depth Images



Black designates areas without depth data due to sensor limitations.

Results: MMF Models from Depth Images



Black designates areas without depth data due to sensor limitations.

Statistics over NYU V2.0 dataset

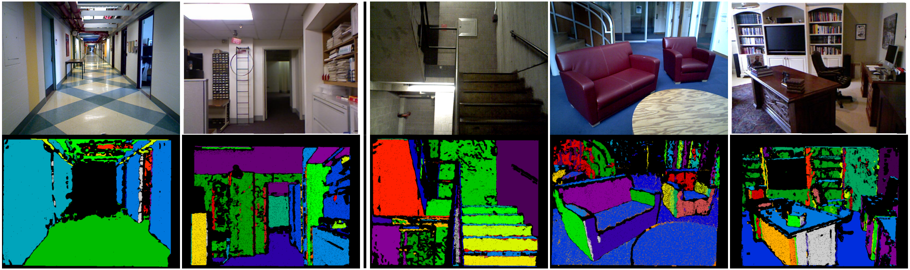
Over all 1449 scenes from the NYU V2.0 depth dataset:

- **Portion of the times inferring K correctly**: 80.5% (K=6)
- **Sensitivity to initial K**: convergence repeatedly to the same K in 95.3% of the scenes.

MMF Inference on Cambridge LiDAR Dataset

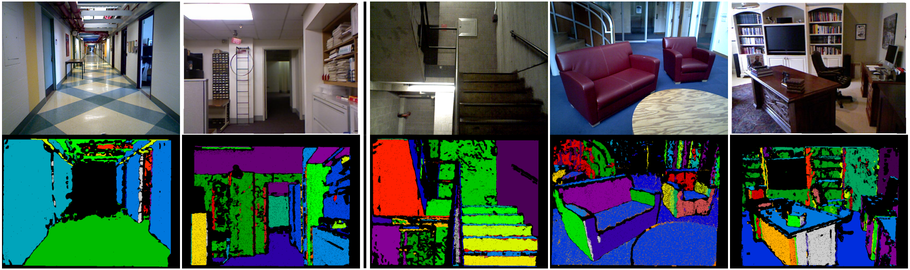
Applications

Plane segmentation: histogram based - straight forward using MMF



Applications

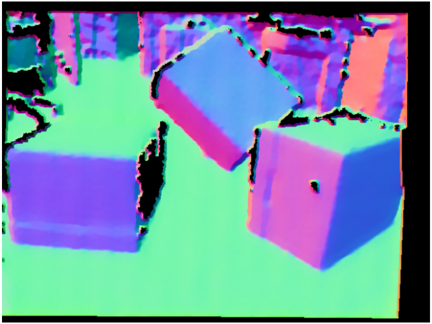
Plane segmentation: histogram based - straight forward using MMF



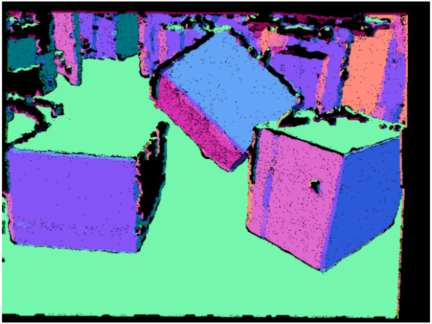
Focal length calibration of depth cameras [by co-author Guy Rosman]

Applications

Normal Correction through pooling of normals across the whole scene



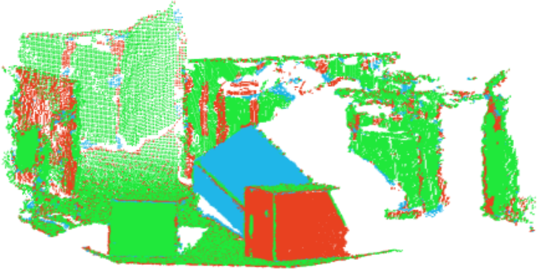
(a) Original Normals

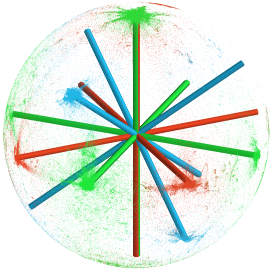


(b) Corrected Normals

Conclusion

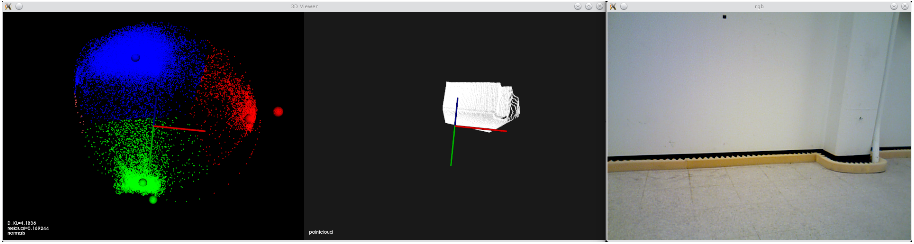
- **Novel probabilistic model** for describing complex man-made scenes
- **Full 3D rotation estimation** for all MFs
- Adaptive model complexity through **split and merge steps**
- Better normal estimates by **pooling observations** across the scene





Future Work

- **"Boxworld"** - put everything in a scene into 3D bounding boxes
- Realtime implementation utilizing GPU - **aid orientation estimation for robots**: "Compass" for major directions



Round Objects

S^2

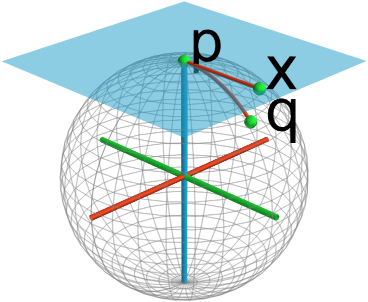
Geodesic Distance between two points p and q on the sphere defined as the angle θ between them:

$$d_G(p, q) = \theta = \arccos(a^T b) \qquad (2)$$

Riemannian logarithm $\mathrm{Log}_p : S^2 \to T_p S^2$ from a point $q \in S^2$ to the tangent space $T_p S^2$ around p:

$$\mathrm{Log}_p(q) = (q - p \cos\theta) \frac{\theta}{\sin\theta}, \qquad (3)$$

where θ is the geodesic distance between p and q

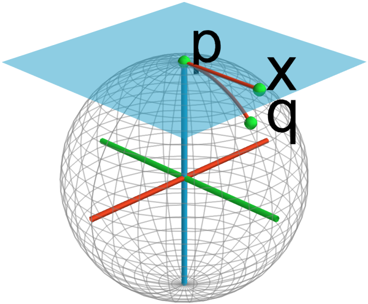


S^2

Riemannian exponential map

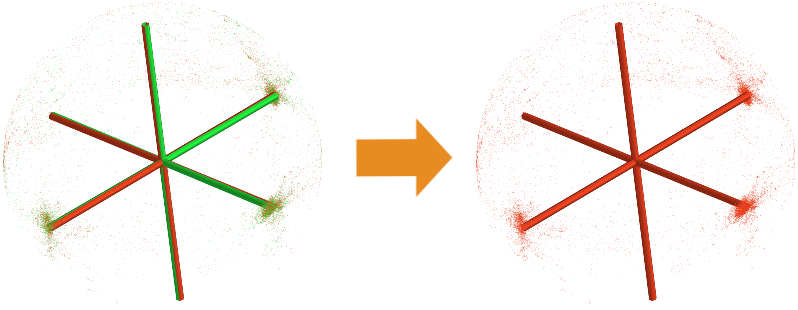
$\mathrm{Exp}_p : T_pS^2 \rightarrow S^2$ a point x in the tangent space around p back onto the sphere

$$\mathrm{Exp}_p(q) = p\cos(||x||) + \frac{x}{||x||}\sin(||x||)$$



Merge Step: MF A + MF B $\rightarrow$ MF C

1. Assign all normals of MF A and MF B to MF C
2. Sample axis assignments for MF C: $\mathbf{z}_{\{\mathbf{c}=C\}}|\mathbf{c}, w_A, R_A, \mathbf{\Sigma}_A, \mathbf{q}$
3. Sample MF rotation given assignments and normals
4. Sample MF axis covariances



Gibbs Sampling Posteriors

The posteriors for both Categorical distributions are:

$$p(\pi|\mathbf{c};\alpha) = \mathrm{Dir}(\alpha_1 + N_1, \ldots, \alpha_K + N_K) \quad (4)$$

$$p(w_k|\mathbf{c},\mathbf{z};\gamma) = \mathrm{Dir}(\gamma_1 + N_{k1}, \ldots, \gamma_{k6} + N_{k6})\ . \quad (5)$$

Evaluating the probability of q_i in $T_{[M_{c_i}]_j}S^2$, the posterior distributions for labels c_i and z_i are given as:

$$p(c_i = k|\pi, q_i, \Theta) = \pi_k \sum_{j=1}^{6} w_{kj}\, p(q_i; [M_k]_j, \Sigma_{kj}) \quad (6)$$

$$p(z_i = j|c_i, q_i, \Theta) = w_{c_i j}\, p(q_i; [M_{c_i}]_j, \Sigma_{c_i j})\ , \quad (7)$$

where we collect the random variables $\mathbf{w}, \mathbf{\Sigma}$ and $\mathbf{R}$ in Θ.
$x_i = \mathrm{Log}_{[M_{c_i}]_{z_i}}(q_i)$, the mapping of q_i into $T_{[M_{c_i}]_{z_i}}S^2$, to obtain the scatter matrix $S_{kj} = \sum_i^N \mathbb{1}_{[c_i=k]}\mathbb{1}_{[z_i=j]} x_i x_i^T$ in $T_{[M_k]_j}S^2$. This allows us to evaluate the posterior distribution over covariances Σ_{kj} as:

$$p(\Sigma_{kj}|\mathbf{c},\mathbf{z},\mathbf{q},\mathbf{R};\Delta,\nu) = \mathrm{IW}\left(\Delta + S_{kj}, \nu + N_{kj}\right)\ . \quad (8)$$

Metropolis Hastings

$$\Pr\left(\text{accept merge}\right) = \min\left\{1, \frac{p(\mathbf{q}, \hat{\mathbf{c}}, \hat{\mathbf{z}}, \hat{\pi}, \hat{\mathbf{w}}, \hat{\Sigma}, \hat{\mathbf{R}}; \alpha, \gamma, \Delta, \nu)}{p(\mathbf{q}, \mathbf{c}, \mathbf{z}, \pi, \mathbf{w}, \mathbf{\Sigma}, \mathbf{R}; \alpha, \gamma, \Delta, \nu)} \frac{q(\text{split})}{q(\text{merge})}\right\} .$$

$$\frac{p(\mathbf{q},\hat{\mathbf{c}},\hat{\mathbf{z}},\hat{\pi},\hat{\mathbf{w}},\hat{\Sigma},\hat{\mathbf{R}};\alpha,\gamma,\Delta,\nu)}{p(\mathbf{q},\mathbf{c},\mathbf{z},\pi,\mathbf{w},\mathbf{\Sigma},\mathbf{R};\alpha,\gamma,\Delta,\nu)} = \frac{p(\mathbf{q}|\hat{\mathbf{c}},\hat{\mathbf{z}},\hat{\mathbf{R}},\hat{\Sigma})p(\hat{\mathbf{c}};\alpha)p(\hat{\mathbf{z}}|\hat{\mathbf{c}};\gamma)p(\hat{\Sigma};\Delta,\nu)p(\hat{\mathbf{R}})}{p(\mathbf{q}|\mathbf{c},\mathbf{z},\mathbf{R},\mathbf{\Sigma})p(\mathbf{c};\alpha)p(\mathbf{z}|\mathbf{c};\gamma)p(\mathbf{\Sigma};\Delta,\nu)p(\mathbf{R})}$$
$$= \frac{8\pi^2\left(\prod_i^N p(q_i|\hat{c}_i,\hat{z}_i,\hat{\mathbf{R}},\hat{\Sigma})\right)p(\hat{\mathbf{c}};\alpha)\prod_{k=1}^{\hat{K}} p(\hat{\mathbf{z}}_{\{\hat{\mathbf{c}}=k\}}|\hat{\mathbf{c}};\gamma)\prod_{j=1}^{6} p(\hat{\Sigma}_{kj};\Delta,\nu)}{\left(\prod_i^N p(q_i|c_i,z_i,\mathbf{R},\mathbf{\Sigma})\right)p(\mathbf{c};\alpha)\prod_{k=1}^{K} p(\mathbf{z}_{\{\mathbf{c}=k\}}|\mathbf{c};\gamma)\prod_{j=1}^{6} p(\Sigma_{kj};\Delta,\nu)} ,$$

$$\frac{q(\text{split})}{q(\text{merge})} = \frac{q(\mathbf{c}, \mathbf{z}, w_{\{l,m\}}, R_{\{l,m\}}, \mathbf{\Sigma}_{\{l,m\},\{1\ldots6\}}|\hat{\mathbf{c}}, \hat{R}_n, \hat{\mathbf{\Sigma}}_n, \mathbf{q}; \alpha, \gamma, \Delta, \nu)}{q(\hat{\mathbf{c}}, \hat{\mathbf{z}}, \hat{w}_n, \hat{R}_n, \hat{\Sigma}_{n,\{1\ldots6\}}|\mathbf{c}, R_l, \mathbf{\Sigma}_l, \mathbf{q}; \alpha, \gamma, \Delta, \nu)} .$$

Merge Proposal

The proposal distributions of merging MF l and m into MF n, factors into

$$
\begin{aligned}
q(\hat{\mathbf{c}}_{\{\mathbf{c}\in\{l,m\}\}}|\mathbf{c};\alpha) &= \delta(\hat{\mathbf{c}}_{\{\mathbf{c}\in\{l,m\}\}} = n) \\
q(\hat{\mathbf{z}}_{\{\hat{\mathbf{c}}=n\}}|\hat{\mathbf{c}}, w_l, R_l, \boldsymbol{\Sigma}_l, \mathbf{q}) &= \prod_{\{i|\hat{c}_i=n\}} w_{l,z_i}\, p(q_i; [M_l]_{z_i}, \Sigma_{lz_i}) \\
q(\hat{R}_n|R_l, \hat{\mathbf{z}}, \hat{\mathbf{c}}, \mathbf{q}) &= \mathcal{N}(\hat{R}_n; \hat{R}^{\star}_n(R_l, \hat{\mathbf{z}}, \hat{\mathbf{c}}, \mathbf{q}), \Sigma^P) \\
q(\hat{\Sigma}_{n,\{1\dots6\}}|\hat{\mathbf{c}}, \hat{\mathbf{z}}, \mathbf{q}, \hat{R}_n; \Delta, \nu) &= \prod_{j=1}^{6} p(\hat{\Sigma}_{nj}|\hat{\mathbf{z}}, \hat{\mathbf{c}}, \mathbf{q}, \hat{R}_n; \Delta, \nu)\ ,
\end{aligned}
$$

Split Proposal

For the split proposal, we randomly assign normals in MF n to MF l or m

$$q(\mathbf{c}_{\{\hat{\mathbf{c}}=n\}}|\hat{\mathbf{c}};\alpha) = \mathrm{DirMult}\left(\mathbf{c}_{\{\hat{\mathbf{c}}=n\}};\alpha_l,\alpha_m\right)$$

$$q(\mathbf{z}_{\{\mathbf{c}=l\}},\mathbf{z}_{\{\mathbf{c}=m\}}|\mathbf{c},\hat{w}_n,\hat{R}_n,\hat{\Sigma}_n,\mathbf{q}) =$$

$$\prod_{\{i|c_i=l\}} \hat{w}_{n,z_i}\, p(q_i;[\hat{M}_n]_{z_i},\hat{\Sigma}_{nz_i}) \prod_{\{i|c_i=m\}} \hat{w}_{n,z_i}\, p(q_i;[\hat{M}_n]_{z_i},\hat{\Sigma}_{nz_i})$$

$$q(R_l,R_m|\mathbf{z},\mathbf{c},\mathbf{q},\hat{R}_n) =$$

$$\mathcal{N}(R_l;R_l^\star(\hat{R}_n,\mathbf{z},\mathbf{c},\mathbf{q}),\Sigma^P)\,\mathcal{N}(R_m;R_m^\star(\hat{R}_n,\mathbf{z},\mathbf{c},\mathbf{q}),\Sigma^P)$$

$$q(\boldsymbol{\Sigma}_{\{l,m\},\{1\ldots6\}}|\mathbf{c},\mathbf{z},\mathbf{q},R_{\{l,m\}};\Delta,\nu) =$$

$$\prod_{j=1}^{6} p(\Sigma_{lj}|\mathbf{z},\mathbf{c},\mathbf{q},R_l;\Delta,\nu)p(\Sigma_{mj}|\mathbf{z},\mathbf{c},\mathbf{q},R_m;\Delta,\nu)\ .$$