Additional file 1 — State networks of $\mathcal{M}_1$ and $\mathcal{M}_2$.

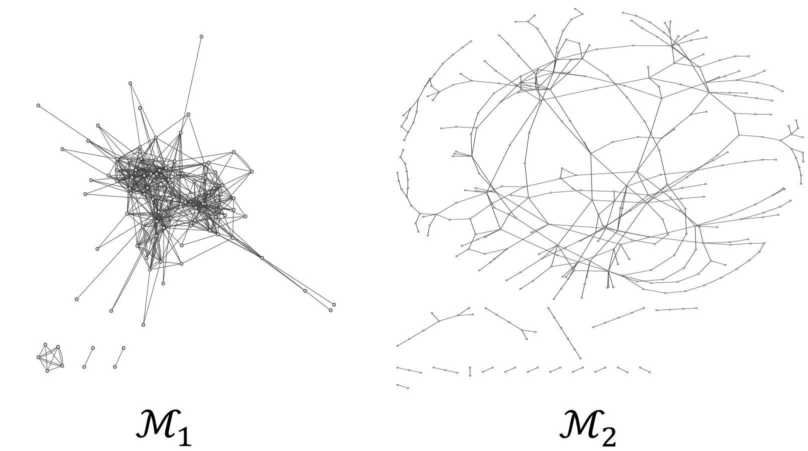


Figure 1 State networks of $\mathcal{M}_1$ and $\mathcal{M}_2$. Each node corresponds to a single state node. The first-order memory network $\mathcal{M}_1$ contains 96 state nodes with a one-to-one mapping to the 96 wards (physical nodes). $\mathcal{M}_1$ consists of four weakly connected components, one of which contains 87 out of 96 the state nodes. The second-order memory network $\mathcal{M}_2$ has 384 state nodes, in 18 weakly connected components. $\mathcal{M}_2$ consists of a single weakly component containing 329 of 384 the state nodes. Structurally, $\mathcal{M}_1$ is more connected with a clustering coefficient of 0.287 and a diameter of 6, whereas $\mathcal{M}_2$ is less connected with a clustering coefficient of 0.003 and a larger diameter of 31.

Additional file 2 — Cross validation ranking significance

Table 1 Cross validation fold ranked correlation p-values (computed using Kendall's tau measurement) for the models of order k=1,2,3,4.

Data Fold	k=1	k=2	k=3	k=4
1	8.22E-08	1.09E-15	4.90E-15	4.15E-15
2	2.03E-07	1.25E-12	1.12E-13	1.12E-13
3	1.72E-08	7.40E-13	1.27E-15	2.90E-15
4	1.06E-05	6.18E-12	3.95E-14	2.81E-14
5	4.27E-09	1.64E-13	2.30E-14	1.69E-14

Additional file 3 — PageRank difference between $\mathcal{M}_1$ and $\mathcal{M}_2$ over specialities and buildings.

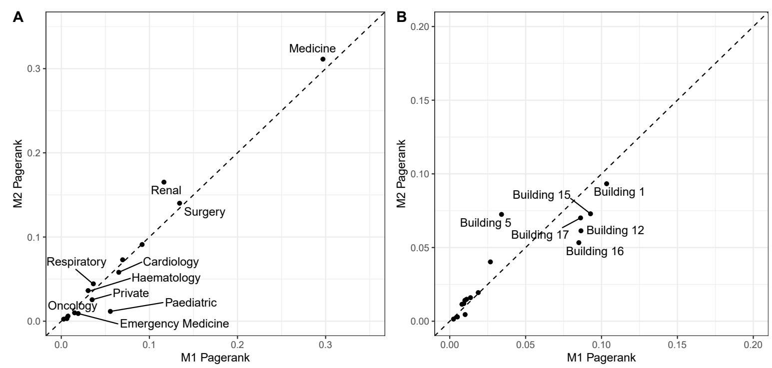


Figure 2 PageRank difference between $\mathcal{M}_1$ and $\mathcal{M}_2$ over specialities and buildings. Additional to analysing ward PageRanks between $\mathcal{M}_1$ and $\mathcal{M}_2$, we summed up the PageRanks of wards belonging to specialities and buildings to arrive at their visitation probabilities. Figure 2 A & B show the comparative results, and whilst their is less dispersion when compared to ward PageRanks, this makes sense given that specialities are more coarse groupings, and likely hide the ward variations seen previously.

Additional file 4 — Optimisation of clustering rate for state lumping

Figure 3 Optimisation of clustering rate r for lumping state nodes. In order to select a clustering rate r, which parameterises lumping, we investigated it's affect on (1) the number of states in the model, and (2) how well structures of patient movement can be detected in communities from the MS framework. We refer to this as model *fitness*, which is aggregate amount structures (hospital sites, specialities, and buildings) found significantly *over-represented* (computed using a Fisher's exact test) in MS communities for $t > 0.316$ (threshold in t corresponding to the point of 20 partitions regardless of the clustering rate). Figure 3A shows the linear relationship between the number of states and r, whereby increases in r lead to a greater number of state nodes (i.e. less lumping). Whereas Figure 3B, the fitness curve, shows that for the same parameter range in r that the model fitness does not change linearly. In fact, we observe a local peak in fitness around $r = 0.35$ whereby the total number of state nodes has reduced substantially to 171 state nodes. We hypothesise this point retains important structure of patient movement in its communities whilst removing redundant state nodes.

Additional file 5 — Markov stability run statistics.

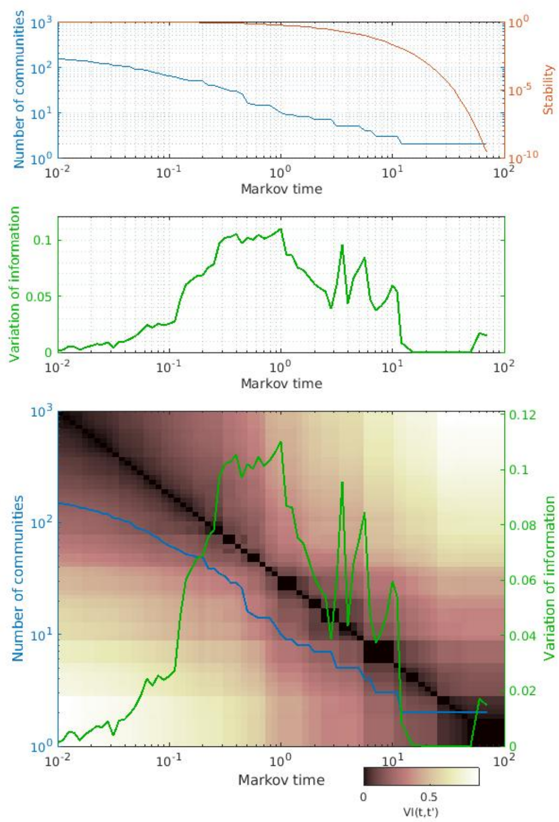


Figure 4 Markov Stability Analysis over the lumped state network $\hat{\mathcal{M}}_2$.. Top: the number of communities and the Markov Stability as a function of Markov time. Middle: The Variation of Information computed over the set of Louvain optimisations at each Markov time, whereby a low VI corresponds to a robust partition. Bottom: The combined Variation of Information and number of communities. The heatmap represents the Variation of Information computed between the optimal partition at each Markov time, where the diagonal is zeros, and we look for blocks of low VI that indicate robust partitions.

Additional file 6 — Markov stability community partitions.

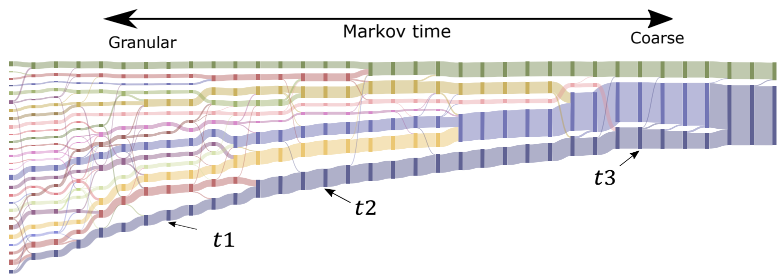


Figure 5 Sankey diagram showing full MS community partitions over t with granular partitions captured towards the left, and coarse partitions captured towards the right. Where $t_1 = 0.708$, $t_2 = 1.585$, and $t_3 = 8.913$ are the stable partitions of interest. We find that MS produces a hierarchy of community partitions across Markov time t. When t is smaller the resultant partitions are granular, and consequently more numerous, then as t increases partitions become coarse by merging granular communities together.

Additional file 7 — Variation of Information between hospital structures in community partitions.

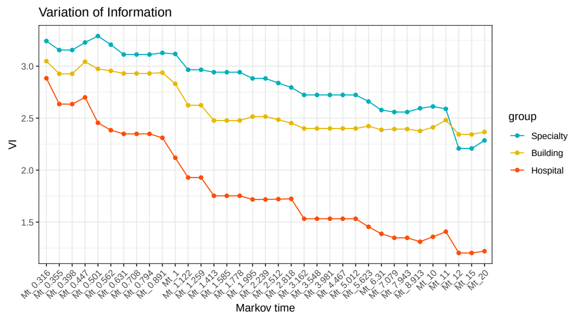


Figure 6 The Variation of Information computed between each hospital structure partition and the community partitions found at each Markov time. Similar to MS we can compute the Variation of Information (VI) to assess distance between clustering, except here we can turn to how well over t the resultant partitions confer to our known structures in the hospital (sites, buildings and specialities). As t increases all structures become more aligned with MS communities, however, hospital sites seems to confer far better across t, even with an initial high VI the rate. Furthermore, Hospital exhibits a faster decrease rate when compared to Speciality or Buildings, and suggests that coarser communities confer most to hospital sites. However, the comparatively smaller VI for Hospitals across more granular MS communities also suggests presence of within hospital structures of patient movement, not bound solely by buildings or specialities.

Additional file 8 — 2-way community partition to hospital site.

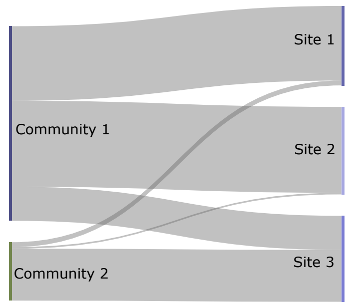
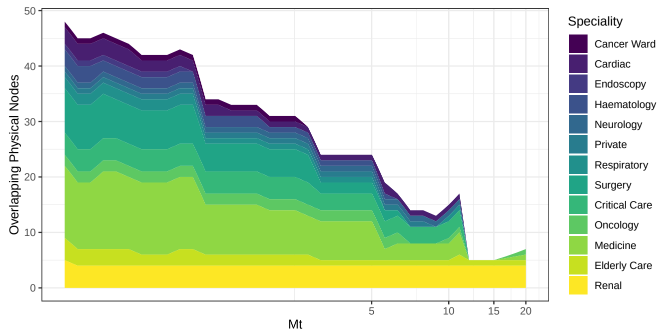


Figure 7 The Markov Stability community partition at Markov time $t = 20$ and their assignments to hospital sites.

Additional file 9 — Hospital wards overlapping communities across Markov stability partitions.



Wards over Mt.pdf

Figure 8 The frequency of physical wards that are members of more than one MS community as a function of Markov time t. For example, the Renal speciality has four wards that overlap between different communities for the majority of Markov time.

Additional file 10 — Examining the overlap between Markov Stability community partitions and known hospital structures

Table 2 Significance of overlaps between the state node in the MS communities at t_1, t_2, and t_3, and the hospital sites. Overlaps with a non-significant p-value > 0.05 (determined using a Fisher's exact test) are denoted "ns".

Time-Community	Site 1	Site 2	Site 3
t_1-C_1	4.26E-07	ns	ns
t_1-C_2	ns	ns	ns
t_1-C_3	ns	ns	ns
t_1-C_4	ns	2.80E-02	ns
t_1-C_5	ns	ns	2.77E-05
t_1-C_6	ns	3.98E-03	ns
t_1-C_7	4.42E-08	ns	ns
t_1-C_8	ns	1.22E-04	ns
t_1-C_9	ns	2.80E-06	ns
t_1-C_{10}	ns	ns	ns
t_1-C_{11}	ns	ns	6.64E-06
t_1-C_{12}	ns	ns	9.35E-04
t_1-C_{13}	ns	ns	ns
t_1-C_{14}	ns	ns	2.64E-03
t_2-C_1	1.81E-17	ns	ns
t_2-C_2	ns	ns	2.67E-10
t_2-C_3	ns	4.16E-02	ns
t_2-C_4	ns	5.60E-09	ns
t_2-C_5	ns	ns	ns
t_2-C_6	1.20E-02	ns	ns
t_2-C_7	ns	2.80E-06	ns
t_2-C_8	ns	ns	8.68E-09
t_3-C_1	ns	ns	4.48E-15
t_3-C_2	1.85E-17	ns	ns
t_3-C_3	ns	1.59E-14	ns

Table 3 Significance of overlaps between the state node in the MS communities at t_1, t_2, and t_3, and the hospital specialties. Overlaps with a non-significant p-value > 0.05 (determined using a Fisher's exact test) are denoted "ns".

Time-Community	Oncology	Medicine	Renal	Respiratory	Surgery	Elderly Care	Private	Critical Care	Haematology	Cardiology	Cancer	endoscopy	Emergency Medicine	Neurology	Gynaecology
t_1-C_1	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
t_1-C_2	9.98E-03	ns	ns	ns	ns	ns	1.06E-03	ns	ns	ns	ns	ns	ns	ns	ns
t_1-C_3	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
t_1-C_4	ns	ns	ns	ns	ns	ns	ns	ns	ns	6.94E-06	ns	ns	ns	ns	ns
t_1-C_5	ns	2.82E-02	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
t_1-C_6	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
t_1-C_7	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
t_1-C_8	ns	ns	1.68E-05	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
t_1-C_9	ns	ns	ns	ns	ns	ns	ns	ns	4.22E-06	ns	ns	ns	ns	ns	ns
t_1-C_{10}	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
t_1-C_{11}	ns	ns	ns	ns	4.04E-04	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
t_1-C_{12}	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
t_1-C_{13}	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
t_1-C_{14}	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
t_2-C_1	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
t_2-C_2	ns	4.56E-03	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
t_2-C_3	ns	ns	ns	ns	ns	ns	ns	ns	ns	4.57E-06	ns	ns	ns	ns	ns
t_2-C_4	ns	ns	2.74E-05	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
t_2-C_5	9.98E-03	ns	ns	ns	ns	ns	1.06E-03	ns	ns	ns	ns	ns	ns	ns	ns
t_2-C_6	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
t_2-C_7	ns	ns	ns	ns	ns	ns	ns	ns	4.22E-06	ns	ns	ns	ns	ns	ns
t_2-C_8	ns	ns	ns	ns	2.22E-03	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
t_3-C_1	ns	2.11E-02	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
t_3-C_2	ns	ns	3.76E-06	ns	ns	ns	ns	ns	1.92E-06	ns	ns	ns	ns	ns	ns
t_3-C_3	3.46E-03	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns

Table 4 Significance of overlaps between the state node in the MS communities at t_1, t_2, and t_3, and the hospital buidlings. Overlaps with a non-significant p-value > 0.05 (determined using a Fisher's exact test) are denoted "ns".

Time-Community	Building 1	Building 2	Building 3	Building 4	Building 5	Building 5	Building 7	Building 8	Building 9	Building 10	Building 11	Building 12	Building 13	Building 14	Building 15	Building 16	Building 17	Building 18
t_1-C_1	ns	2.27E-04	ns	ns	2.62E-03	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
t_1-C_2	2.43E-02	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	2.01E-02	ns	ns	ns	ns	ns	ns
t_1-C_3	ns	ns	1.49E-02	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
t_1-C_4	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	8.50E-04	ns	ns	ns	ns	ns	ns
t_1-C_5	ns	ns	ns	ns	ns	ns	ns	ns	8.50E-03	ns	ns	ns	ns	ns	ns	ns	ns	ns
t_1-C_6	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
t_1-C_7	9.61E-03	7.45E-03	ns	4.31E-02	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
t_1-C_8	ns	ns	2.78E-04	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
t_1-C_9	ns	ns	ns	ns	ns	7.53E-03	ns	5.25E-03	ns	ns	ns	ns	ns	ns	2.14E-02	ns	ns	ns
t_1-C_{10}	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
t_1-C_{11}	ns	ns	ns	ns	ns	ns	ns	ns	2.58E-03	ns	ns	ns	ns	ns	ns	ns	ns	ns
t_1-C_{12}	ns	ns	ns	ns	ns	ns	ns	ns	5.28E-03	ns	ns	ns	ns	ns	ns	ns	ns	ns
t_1-C_{13}	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
t_1-C_{14}	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
t_2-C_1	1.46E-02	3.68E-08	ns	2.23E-02	4.94E-04	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
t_2-C_2	ns	ns	ns	ns	ns	ns	ns	ns	2.52E-05	2.29E-02	2.05E-02	ns	ns	5.89E-03	ns	ns	ns	ns
t_2-C_3	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	1.99E-05	ns	ns	ns	ns	ns	ns
t_2-C_4	ns	ns	8.77E-07	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
t_2-C_5	2.43E-02	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	2.01E-02	ns	ns	ns	ns	ns	ns
t_2-C_6	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
t_2-C_7	ns	ns	ns	ns	ns	7.53E-03	ns	5.25E-03	ns	ns	ns	ns	ns	ns	2.14E-02	ns	ns	ns
t_2-C_8	ns	ns	ns	ns	ns	ns	ns	ns	3.96E-04	ns	ns	ns	ns	ns	ns	ns	ns	ns
t_3-C_1	ns	ns	ns	ns	ns	ns	ns	ns	7.71E-08	1.16E-02	ns	ns	ns	ns	ns	ns	ns	ns
t_3-C_2	1.18E-04	3.27E-06	ns	2.92E-02	5.66E-03	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
t_3-C_3	ns	ns	6.69E-07	ns	ns	4.37E-02	ns	3.29E-04	ns	ns	ns	ns	ns	ns	8.79E-03	ns	ns	ns

Additional file 11 — Multiscale Centrality model comparison.
For further examination of the importance of higher-order modelling, we compared the MSC ranking of wards in the lumped network $\hat{\mathcal{M}}_2$ to the original state node networks of $\mathcal{M}_1$ and $\mathcal{M}_2$. We found that whilst correlated, there were a number of distinct differences between the models (Figure 10).
For instance, we found several wards, including a critical care ward that were central at all time-scales in $\mathcal{M}_2$ and $\hat{\mathcal{M}}_2$ only appeared as important at short time-scales in $\mathcal{M}_1$. We found that the MSC node ranking for $\hat{\mathcal{M}}_2$ was marginally more correlated with $\mathcal{M}_1$(Ranked Cor: 0.86 (pval <0.01)) than $\mathcal{M}_2$Ranked Cor: 0.84 (pval <0.01)), which makes sense given that the state space of the lumped state network $\hat{\mathcal{M}}_2$ is closer in size to $\mathcal{M}_1$ than $\mathcal{M}_2$.

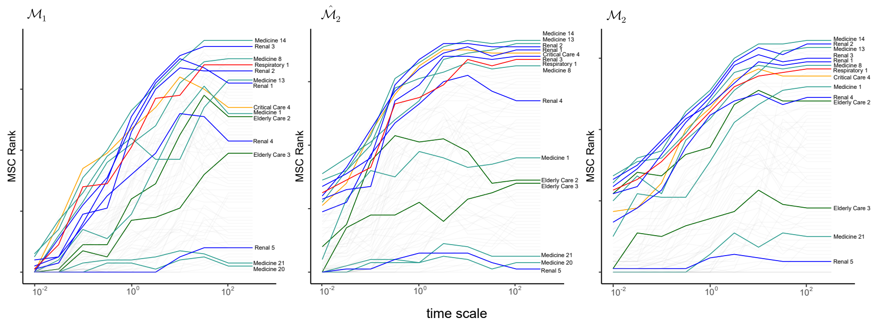


Figure 9 Multiscale centrality ranks over time for $\mathcal{M}_1$, $\hat{\mathcal{M}}_2$, and $\mathcal{M}_2$ from left to right.

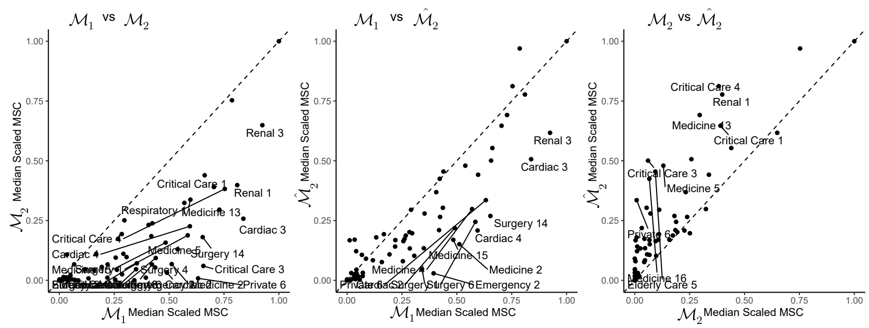


Figure 10 A comparison of the median Multiscale centrality for the first-order $\mathcal{M}_1$, second-order $\mathcal{M}_2$ and lumped $\hat{\mathcal{M}}_2$ memory networks

Additional file 12 — Distribution of node statistics.

node stats.pdf

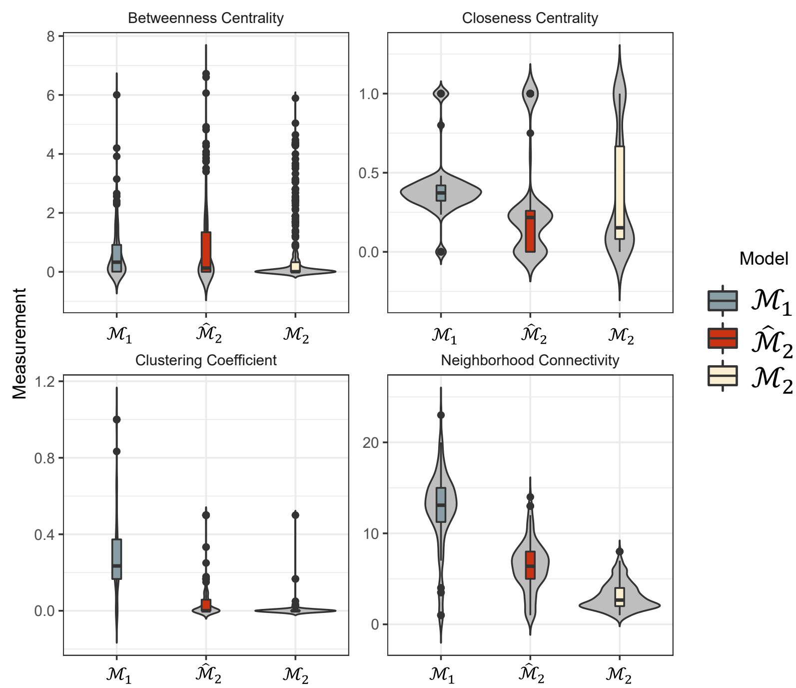


Figure 11 Comparison of network node statistics distributions for the first-order $\mathcal{M}_1$, second-order $\mathcal{M}_2$ and lumped $\hat{\mathcal{M}}_2$ memory networks. For greater comparison of the three networks, we computed node betweenness centralities, closeness centralities, clustering coefficients and neighbourhood connectivities. In terms of betweenness centrality, all distributions are left-skewed. $\mathcal{M}_2$ is the most dramatically skewed, with the lowest median=0.01, and the lowest variance=1.03. $\mathcal{M}_1$ has the highest median=0.33, for a similar variance=1.05, suggesting a higher proportion of nodes play a role in the core network structure. Notably, $\hat{\mathcal{M}}_2$ sits between $\mathcal{M}_1$ and $\mathcal{M}_2$ in terms of median=0.87, yet has the highest variance=1.37. The closeness centrality which is the measure of a nodes average farness to all other nodes decreases with the amount of memory included per model, with medium values falling from 0.37 for $\mathcal{M}_1$, to 0.22 for $\hat{\mathcal{M}}_2$, then to 0.15 for $\mathcal{M}_2$. This decreasing trend in centrality highlights the increasing sparseness and size of the networks as more memory (and states nodes) are incorporated into its structure. All three network models have a portion of nodes with a higher closeness centrality suggesting the presence of certain hub-like state nodes, regardless of the memory. Notably, $\hat{\mathcal{M}}_2$ is the only model with pronounced bi-modal distribution, interestingly this in-between $\mathcal{M}_1$ and $\mathcal{M}_2$'s distributions, which may indicate the that $\hat{\mathcal{M}}_2$ is preserving certain structural properties of both $\mathcal{M}_1$ and $\mathcal{M}_2$. Both the clustering coefficient and neighbourhood connectivity followed similar trends, with $\mathcal{M}_1$ having the highest median values (0.24 and 13.09 for clustering coefficient and neighbourhood connectivity respectively), highly the largely more connectedness of $\mathcal{M}_1$ in comparison to $\hat{\mathcal{M}}_2$ and $\mathcal{M}_2$'. Similar to previous metrics, $\hat{\mathcal{M}}_2$ also sat between $\mathcal{M}_1$ and $\mathcal{M}_2$ in terms of both clustering coefficient and neighborhood connectivity, highlighting again that $\hat{\mathcal{M}}_2$ is a mixture of properties from both $\mathcal{M}_1$ and $\mathcal{M}_2$.