nature portfolio

Peer Review File

Spatiotemporal asymmetries on brain energy landscape uncover system entrapment related to depression severity

Corresponding Author: Dr Ülgen Kilic

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
This manuscript integrates co-activation pattern (CAP) analysis with network control theory (NCT) to characterize an “energy landscape” of resting brain states in MDD. The core claims—higher recurrence yet shorter dwell of a SN/SOM/SUB-engaged state (“State 3”), asymmetric transitions relating to anhedonia and depression severity, and energetic asymmetries linking source-like and sink-like attractors—are intriguing and potentially impactful. However, several issues require further clarification and improvement.

The Introduction should be revised for clarity. In particular, the sentence “Despite decades of neuroimaging research, current first-line treatments fail in over 30% of cases underscoring the need for reliable neural biomarkers to guide diagnosis, predict treatment response, and parse clinical heterogeneity” lacks clear logic and needs to be rewritten.

The sample size is relatively small for the statistical analyses performed (clustering, correlations, etc.). Moreover, running multiple tests across scales and measures increases the risk of false positives, reducing persuasiveness. Please provide effect sizes and 95% confidence intervals for all group comparisons and regression analyses.

The rationale for choosing 7T MRI as the primary modality should be explained more clearly. All main data come from 7T scans, yet in clinical and population studies 3T MRI is far more common. Although reproducibility in an independent dataset is mentioned, I strongly recommend validating the proposed framework using larger 3T MRI datasets, ideally from public repositories.

The authors should elaborate on the novelty and clinical significance of the proposed framework. What unique insights does it offer beyond prior CAP/NCT work, and what potential impact could it have in clinical practice?

In the Results, the interpretation of State 3’s high recurrence and short dwell time as “compulsive re-engagement and premature disengagement of salience–somatosensory–subcortical processing” may be overstated. Given the cross-sectional resting-state design, such causal terms (“compulsive,” “premature”) may be too strong. Please temper this interpretation.

Medication and confounds: Beyond including “medication status” as a covariate, please report medication classes, illness duration, and comorbidities, and verify whether key findings hold in the unmedicated subgroup.

Please check the use of abbreviations throughout the manuscript and ensure consistency.

(Remarks on code availability)

Reviewer #2

(Remarks to the Author)
In this paper authors investigate the dynamics of the organization of resting-state brain activity in patients with major depressive disorder (MDD) compared to healthy controls (HC). Using k-medoids algorithm on regional activation vectors

across all subjects (HC+MDD) and all time points, 4 clusters (representing brain states) were identified. A particularly prominent brain state, marked by suppressed default mode and frontoparietal networks and heightened salience system engagement, occurred more frequently and with shorter dwell times in MDD patents and correlated with greater anhedonia severity. They also used a network control theory (NCT) framework to understand the interplay between brain structure, as measured by diffusion weighted imaging (DWI), and resting-state brain function, as measured by co-activation patterns. NCT revealed the energy landscape of the identified whole-brain functional co-activation patterns. They found that MDD individuals preferentially transition along energetically costly trajectories, particularly from salience-reactive to introspective states, despite the presence of structurally facilitated alternatives, implicating inefficient structure-function coupling. This manuscript is well-written and presents a carefully conducted analysis to provide some new insight into the relationship between structural connectivity, functional dynamics, and symptomatology in depression. Here are some points that can strengthen the paper:
-Was there any temporal filtering (band-pass filter) applied when preprocessing the resting state data?
-state-function coupling analysis can be sensitive to the brain parcellation used. What was the rationale for using Lausanne parcellation?
-Defining and identifying spatiotemporal co-activation patterns (clusters) is a critical part of this study's analysis plan. While authors have done a great job showing robustness and stability of the optimal number of clusters in their method of choice (k-medoids), they have not explained the rational for this choice. Will other methods like ICA also lead to similar results (i.e., MDD patients preferring transitioning along energetically costly trajectories)?
-The authors are generally careful about applying and reporting correction for multiple comparisons; however, there are few places that it is not clear if the results are corrected or not. For example, in lines 218-220 when reporting the correlations between exit and enter probabilities to S3 and clinical symptoms.
Minor points:
-Figure 1. Caption. "…and individuals with 38 Major Depressive Disorder.." doesn't make sense.
-Figure 5. Part b. color labels are missing.

(Remarks on code availability)

Reviewer #3

(Remarks to the Author)
This manuscript investigated the dynamic brain state transitions in patients with major depressive disorder by integrating co-activation pattern analysis of resting-state fMRI and network control theory applied to diffusion-weighted imaging. A series of analyses were conducted to examine whether transitions between fMRI-derived states are differentially constrained by the structural connections in MDD from the DWI dataset and related to symptomatology. However, the crucial NCT analysis, which is the most novel and complex part of the paper, was conducted on a subsample of only 27 HC and 26 MDD, in which ~45% of the MDD participants were on psychotropic medications. This questions the reliability of the null findings. Other concerns are as follows:

1. The choice of four brain states (k=4) is central to the entire paper. While the authors provide multiple lines of evidence, some justification is weak. The spatial correlation between State 3 in the main dataset and the replicate cluster in the independent dataset was only r=0.30, which is quite low. Furthermore, the rationale that k>5 leads to states not represented in all subjects is not a strong statistical argument for choosing a lower k; it could simply mean those rarer states are clinically informative. Alternative approaches, such as hidden Markov model, can be supplemented to uncover the dynamic properties of spatial states.
2. The brain states are defined from the fMRI data of all subjects (HC+MDD). These states are then used as the target configurations for the NCT analysis, which is run on each subject's individual structural connectome. There is a potential circularity: the states are a group-level construct, and the analysis tests how each individual's structure facilitates transitions to these group-defined states. A compelling supplementary analysis would be to repeat the NCT analysis using subject-specific states.
3. The introduction criticizes static connectivity measures. However, the discussion does not fully reconcile the dynamic findings with the vast literature on DMN hyperactivity and DMN-FPN hypoconnectivity in MDD. For instance, how does the frequently visited but unstable State 3 (with suppressed DMN) relate to the classic finding of hyperactive DMN?
4. The discussion interprets the results through the lens of complex dynamical systems theory (attractors, basins, energy landscapes). However, the direct neural basis of "control energy" is abstract. The authors correctly note it is not a direct measure of metabolic cost but is more analogous to an external input like neuromodulation (line 435-437). The discussion should include a specific paragraph hypothesizing the neurobiological instantiation of "control energy" , or explicitly state that the energy landscape is a computational theory that provides a formal framework for designing experiments to test these biological hypotheses. Above all, the calculated "control energy" is not a direct, biological measure but a theoretical metric derived from a specific simplified model. The absolute energy values and even some relative comparisons could be sensitive to these modeling choices. The conclusions about "energetic cost" are therefore model-dependent.

(Remarks on code availability)

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
This manuscript integrates co-activation pattern (CAP) analysis with network control theory (NCT) to characterize a state-transition “energy landscape” in major depressive disorder (MDD), linking specific dynamic properties to symptom dimensions (e.g., anhedonia, depression severity). The revision substantially strengthens the work through multiple targeted robustness checks and improved statistical reporting.Most major concerns have been addressed, but a few points still require clarification and/or more cautious framing to fully support the central claims.
Q1：k-medoids clustering is performed on all subjects (HC+MDD) and all time points, which is a common CAP approach, but it can still bias state definitions toward group effects in small samples.
Q2：Educational level is an important covariate for brain imaging data, but it was not taken into account in this study.。
Q3：Given that the mechanistic narrative is grounded primarily in linear NCT-derived asymmetries and energy–probability coupling, the “attractor/entrapment” framing should be clearly presented as a model-based metaphor (or operationalized explicitly via enter/exit energy asymmetry), avoiding implications of directly demonstrated nonlinear attractor dynamics.

(Remarks on code availability)

Reviewer #3

(Remarks to the Author)
The authors have addressed all my concerns. I only have a few minor suggestions:

1. In the abstract, instead of "state-transitions are perturbed", the authors may want to specify the nature of the perturbation investigated (e.g., "...how the stability and energetic landscape of state-transitions are perturbed") to sharpen the opening motivation.
2. Please consider citing the recent work by Yu et al. (2025,doi.org/10.1093/psyrad/kkaf008) on white matter functional connectome gradients in major depressive disorder. This study provides important complementary evidence of hierarchical dysfunction within the white matter's intrinsic functional architecture, which would help contextualize your structural control findings within the broader framework of network-level dysregulation in depression.

(Remarks on code availability)

Responses to reviewers' comments for the article

"Spatiotemporal asymmetries on brain energy landscape uncover system entrapment related to depression severity"

B. Ülgen Kilic*[1,2], Jenna Jubeir[1,3], Priti Balchandani[3,4], James W. Murrough[1,3,5], Laurel S. Morris[1,6,7], Yael Jacob[1,3,4]

1. Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY, USA
2. Center for Computational Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY, USA
3. Nash Family Department of Neuroscience & Friedman Brain Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA
4. BioMedical Engineering and Imaging Institute, Department of Radiology, Icahn School of Medicine at Mount Sinai, New York, NY, USA
5. VISN 2 Mental Illness Research, Education and Clinical Center (MIRECC), James J. Peters VA Medical Center, Bronx, NY, USA
6. Nuffield Department of Clinical Neurosciences, University of Oxford, UK
7. Department of Experimental Psychology, University of Oxford, UK

We thank the editors and reviewers for their time and constructive criticisms. We are now submitting an updated version of our manuscript that has taken these items into careful consideration. We have added 5 new Supplementary Materials sections with additional computational experiments remedying reviewer remarks:

- Supplementary Materials: "Medication confounds and replication of key results in the unmedicated MDD subgroup" and Supplementary Figure 7.
- Supplementary Materials: "NCT analysis with subject-specific states" and Supplementary Figure 8.
- Supplementary Materials: "Influence of brain parcellation size" and Supplementary Figure 10.
- Supplementary Materials: "Justification of the usage of k-medoids algorithm" and Supplementary Figure 11.
- Supplementary Materials: "Assessing specificity using geometry preserving null networks" and Supplementary Figure 12.

We have highlighted new changes with a blue font in the manuscript to make them easy to find.

Below, we provide a point-by-point response to these comments.

Sincerely,

B. Ülgen Kilic

Reviewer #1 (Remarks to the Author):

This manuscript integrates co-activation pattern (CAP) analysis with network control theory (NCT) to characterize an "energy landscape" of resting brain states in MDD. The core claims—higher recurrence yet shorter dwell of a SN/SOM/SUB-engaged state ("State 3"), asymmetric transitions relating to anhedonia and depression severity, and energetic asymmetries linking source-like and sink-like attractors—are intriguing and potentially impactful. However, several issues require further clarification and improvement.

We appreciate the reviewer's support and encouragement and thank them for raising these important issues.

The Introduction should be revised for clarity. In particular, the sentence "Despite decades of neuroimaging research, current first-line treatments fail in over 30% of cases underscoring the need for reliable neural biomarkers to guide diagnosis, predict treatment response, and parse clinical heterogeneity" lacks clear logic and needs to be rewritten.

We thank the reviewer for this important feedback. We have revised the introduction for clarity, including rewriting the opening paragraph **(lines:35-38)**.

The sample size is relatively small for the statistical analyses performed (clustering, correlations, etc.). Moreover, running multiple tests across scales and measures increases the risk of false positives, reducing persuasiveness. Please provide effect sizes and 95% confidence intervals for all group comparisons and regression analyses.

We thank the reviewer for this important comment. In the revised manuscript, we have listed the small sample size of our study as a potential limitation in the Discussion section **(lines: 506-508).** Additionally, for group comparisons (e.g., dwell times, fractional occupancy, and transition/persistence probabilities and transition energies), we now report the group means, with raw 95% confidence intervals, and effect sizes (Cohen's *d*) alongside the t-statistics and p-values. Similarly, for regression analyses (e.g., relationships between state metrics and clinical scores), we report both standardized coefficients (β^*) and standardized 95% confidence intervals (CI*), together with their p-values, and model fit indices (R^2). These details are included **throughout the Results sections,** and the changes are highlighted and the procedures for calculating and reporting effect sizes and confidence intervals are updated in Statistical Analysis section in the Methods **(lines:720-732).** Importantly, inclusion of these additional statistics strengthens the interpretability of our results.

The rationale for choosing 7T MRI as the primary modality should be explained more clearly. All main data come from 7T scans, yet in clinical and population studies 3T MRI is far more common. Although reproducibility in an independent dataset is mentioned, I strongly recommend validating the proposed framework using larger 3T MRI datasets, ideally from public repositories.

We thank the reviewer for this important point and agree that replication in independent datasets is essential for scientific rigor. However, as noted by the reviewer, ultra-high-field 7T MRI remains rare in clinical populations, and to the best of our knowledge, this study is the first to employ combined CAP + NCT framework to 7T data in major depression. Therefore, true 7T replication is not currently feasible in larger public datasets. We would like to highlight that we intentionally used 7T data because it offers significantly higher spatial resolution and improved signal-to-noise ratio afford an advantageous compromise between spatial and temporal resolution, thus, it substantially increases sensitivity to instantaneous coactivation patterns (CAPs). Importantly, our team has systematically investigated functional connectivity differences between 3T and 7T MRI in depression, demonstrating the advantages of ultra-high-field imaging for resolving both small, anatomically specific circuits and large-scale whole-brain patterns. Notably, this work also showed that certain group differences between MDD and controls, which can be subtle and difficult to detect at 3T, were reliably identifiable at 7T. We now cite this study (Morris, L. S. et al. 2019), along with other comparison papers (Okada, T. et al. 2022, Pohmann, R. et al. 2016), in the main manuscript for the reviewer's reference.

Overall, the present work is the first, basic mechanistic study linking fMRI and DWI to characterize energy-landscape asymmetries in depression, which lays important groundwork for future translational research.

Therefore, validating our work in 3T MRI data constitutes a distinct future project rather than a direct replication. For this reason, we treated larger 3T replication as an explicitly stated future direction (see the Discussion section, **lines 522-526,** for clarifications we added).

The authors should elaborate on the novelty and clinical significance of the proposed framework. What unique insights does it offer beyond prior CAP/NCT work, and what potential impact could it have in clinical practice?

We thank the reviewer for this remark. We aimed to emphasize in the Introduction section of our original submission that the novelty of our framework lies in coupling empirically derived CAPs with subject-specific structural connectomes in MDD. In summary, prior work has typically reported either CAPs alone or limited NCT analysis without incorporating empirical states found from the rs-fMRI. While some studies have combined CAP and NCT methods, they did not target MDD population. In contrast, our pipeline applies this previously established CAP analysis to map dynamic state transitions and then uses NCT on each subject's white-matter connectome to quantify transition energy that explicitly link observed functional dynamics to underlying structural constraints. These points are described in detail in the Introduction of our original submission **(lines: 41-43, 50-56, 67-75 and 76-83 in the new submission).**

In response to the reviewer's comment, we have also expanded the Discussion to make the clinical implications clearer and to state concrete, testable hypotheses. In summary, by producing individualized transition-energy estimates, the framework yields candidate mechanistic biomarkers that could (for example) predict stimulation dose-response, help localize optimal stimulation targets, stratify patients by likely treatment response, and track early treatment-related changes. These are explicit, falsifiable predictions that require validation. We added a more detailed explanation in the Discussion **(lines:488-505),** and new citations (Zhou, D. et al. 2025, Kringelbach, M. L. 2020, Beynel, L. et al. 2020, He, X. et al. 2025) that will help clinically oriented readers connect these theoretical concepts to translational themes.

In the Results, the interpretation of State 3's high recurrence and short dwell time as "compulsive re-engagement and premature disengagement of salience–somatosensory–subcortical processing" may be overstated. Given the cross-sectional resting-state design, such causal terms ("compulsive," "premature") may be too strong. Please temper this interpretation.

We thank the reviewer for their thoughtful feedback. After carefully re-examining the Results sections and the full manuscript, we were unable to identify any instances in the original submission where the terms "compulsive" or "premature" were used. As such, we found this comment somewhat difficult to interpret. Based on our best understanding of what the reviewer may have been referring to, we identified several sentences in the Discussion **(lines: 404, 411-428)** where we used causal language and softened this tone now.

Additionally, in the Discussion (**lines: 370-371, 398-400**), we have streamlined descriptions of State 3 ("reactive, salience-driven," "salience network dominated," or "characterized by suppression of the default mode and frontoparietal control networks alongside engagement of salience, subcortical, and somatomotor systems") by removing repetitive phrasing or replacing them with more quantitative characterizations (e.g., low-amplitude or above-baseline coactivations).

Medication and confounds: Beyond including "medication status" as a covariate, please report medication classes, illness duration, and comorbidities, and verify whether key findings hold in the unmedicated subgroup.

We appreciate the reviewer's suggestion to further clarify the role of medication and related clinical confounds. In the revised **Supplementary materials section "Medication confounds and replication of key results in the unmedicated MDD subgroup", Supplementary Figure 7** and **Supplementary Tables 2 and 3**, we now provide detailed information on medication classes, illness onset age, and comorbidities for all MDD participants. In addition, we conducted follow-up analyses restricted to the unmedicated subgroup to verify whether the main effects remained consistent when excluding medicated participants. The results show that the principal findings remain qualitatively unchanged within the unmedicated subgroup, although several associations were attenuated and did not reach statistical significance ($p > 0.05$), while still mostly showing

trends in the same direction ($p < 0.1$). This attenuation is likely due to the smaller and unmatched group sizes (n(HC) = 38; n(unmedicated-MDD)= 21), which reduced statistical power. We added appropriate remarks throughout the main text regarding this additional analysis. Overall, these results suggest that the reported effects are not primarily driven by medication status.

Please check the use of abbreviations throughout the manuscript and ensure consistency.

We now define every abbreviation at first use throughout the manuscript.

Reviewer #2 (Remarks to the Author):

In this paper authors investigate the dynamics of the organization of resting-state brain activity in patients with major depressive disorder (MDD) compared to healthy controls (HC). Using k-medoids algorithm on regional activation vectors across all subjects (HC+MDD) and all time points, 4 clusters (representing brain states) were identified. A particularly prominent brain state, marked by suppressed default mode and frontoparietal networks and heightened salience system engagement, occurred more frequently and with shorter dwell times in MDD patents and correlated with greater anhedonia severity. They also used a network control theory (NCT) framework to understand the interplay between brain structure, as measured by diffusion weighted imaging (DWI), and resting-state brain function, as measured by co-activation patterns. NCT revealed the energy landscape of the identified whole-brain functional co-activation patterns. They found that MDD individuals preferentially transition along energetically costly trajectories, particularly from salience-reactive to introspective states, despite the presence of structurally facilitated alternatives, implicating inefficient structure-function coupling. This manuscript is well-written and presents a carefully conducted analysis to provide some new insight into the relationship between structural connectivity, functional dynamics, and symptomatology in depression. Here are some points that can strengthen the paper:

We appreciate the reviewer's support and positive comments about our paper.

-Was there any temporal filtering (band-pass filter) applied when preprocessing the resting state data?

We did not use any bandpass filtering. We employed 7T rs-fMRI data preprocessing via multi-echo independent component analysis (ME-ICA), which denoises the signal by identifying and removing non-BOLD components based on TE-dependence, obviating the need for band-pass filtering. We explain these details of ME-ICA preprocessing in the Methods section "MRI Data Processing" **(lines:575:582).**

-state-function coupling analysis can be sensitive to the brain parcellation used. What was the rationale for using Lausanne parcellation?

We thank the reviewer for this insightful comment. We chose the Lausanne parcellation based on prior literature demonstrating its advantages for CAP + NCT analysis (Cornblath, E. J. et al. 2020, Singleton, S. P. et al. 2022 and Muldoon, S. F. et al. 2016). Specifically, the Lausanne atlas defines brain regions in each participant's native anatomical space, similar to the Desikan-Killiany atlas, but with the added benefit of providing a hierarchical set of nested parcellations across five spatial scales. This multiscale framework allows for analyses at varying levels of spatial resolution while maintaining anatomical correspondence across scales. Starting from scale 1 (85 ROIs, analogous to the Desikan-Killiany atlas), each cortical region is progressively subdivided to produce finer parcellations with 131, 233, 463, and 1019 ROIs at scales 2–5, respectively. This hierarchical and anatomically grounded approach has been shown to enhance reproducibility and interpretability in both structural and functional connectomics studies (Tourbier, S. et al. 2022, Daducci et al., 2012). We now highlight this description in the Methods section **(lines: 572-575)**, together with new citations.

In addition, to utilize this advantage, we now provide additional analyses demonstrating that our results remain quantitatively consistent across scales. Our new results show similar trends to the main findings at scale 3 (with 233 ROIs) compared to scale 1 (85 ROIs, reported in the main text), as presented in **Supplementary Figure 10 and described in the Supplementary Materials section "Influence of brain parcellation size".**

-Defining and identifying spatiotemporal co-activation patterns (clusters) is a critical part of this study's analysis plan. While authors have done a great job showing robustness and stability of the optimal number of clusters in their method of choice (k-medoids), they have not explained the rational for this choice. Will other methods like ICA also lead to similar results (i.e., MDD patients preferring transitioning along energetically costly trajectories)?

We thank the reviewer for highlighting this important point. As the reviewer mentioned, we have extensively studied our choice of k=4 clusters in different angles (robustness, stability, optimal number of clusters). Yet, it is known that choosing the true number of clusters is a limitation of CAP-based analysis of rs-fMRI. We highlighted this point in the Discussion (**line:517**) and added a citation of a review paper (Liu, X. et al. 2018).

The other piece of this is explaining our choice of k-medoids algorithm, which we mentioned at the end of the first paragraph of the Methods Section "Clustering of fMRI volumes" in our original submission **(lines: 629-637 in the new submission)**. Moreover, we now added this criteria more carefully at the beginning of the Results section "Canonical RSNs described by spontaneous coactivation patterns" **(lines: 117-129)**.

Additionally, in response to the reviewer comment, we provide **Supplementary Materials Section "Justification of the usage of k-medoids algorithm" and Supplementary Figure 11**, and additional citations (Liu, X. 2013, Smith, S. M. et al. 2012) which will help contextualize this choice.

In this new analysis, we justify our choice of k-medoids algorithm by creating 3 synthetic toy examples of time signals and comparing different clustering approaches: k-means, k-medoids, ICA followed by k-means and ICA followed by k-medoids. Note that an ICA-only approach is not a suitable method for our main goal of studying state transitions. Because this method extracts temporally independent modes, or brain states, each defined by a time course and a node-weight vector that maps its spatial pattern, it produces continuous component loadings rather than temporal state sequences, i.e., discrete cluster labels corresponding to each time point. Therefore, to derive subject-level state sequences of discrete cluster labels, we applied a distance-based clustering to the time-resolved component activations, i.e., clustering in component-activation space rather than on raw instantaneous whole-brain BOLD volumes. Our findings indicate medoid-based clustering is better suited for identifying meaningful, physiologically plausible instantaneous coactivation patterns in BOLD data. Because k-medoids selects real volumes as representative centers it is resilient to outliers and amplitude shifts and produces temporally coherent state blocks that align with visible amplitude regimes in our simulations. By contrast, centroid-based k-means often yields synthetic centers located in low-density regions and gives noisier temporal assignments. Although clustering in the component activation space yields temporally coherent blocks of cluster assignments, these regimes are less obvious to interpret, complicating interpretation. We noted in the newly added limitations paragraph in Discussion section **(lines: 515-522)** that future work should benchmark these different alternatives of clustering in large fMRI datasets. Nevertheless, this additional experiment inspires the future work and justifies our main state definitions and downstream analyses using k-medoids with a correlation-based distance.

-The authors are generally careful about applying and reporting correction for multiple comparisons; however, there are few places that it is not clear if the results are corrected or not. For example, in lines 218-220 when reporting the correlations between exit and enter probabilities to S3 and clinical symptoms.

We thank the reviewer for this comment, we now added all the necessary remarks showing FDR correction throughout the text whenever applicable. We also highlighted number of corrections made in the Methods section "Statistical Inference" (**lines: 724-725**).

Minor points:
-Figure 1. Caption. "…and individuals with 38 Major Depressive Disorder.." doesn't make sense.

We thank the reviewer for catching this mistake, it is now fixed.

-Figure 5. Part b. color labels are missing.

We now added the group labels to the panel.

Reviewer #3 (Remarks to the Author):

This manuscript investigated the dynamic brain state transitions in patients with major depressive disorder by integrating co-activation pattern analysis of resting-state fMRI and network control theory applied to diffusion-weighted imaging. A series of analyses were conducted to examine whether transitions between fMRI-derived states are differentially constrained by the structural connections in MDD from the DWI dataset and related to symptomatology. However, the crucial NCT analysis, which is the most novel and complex part of the paper, was conducted on a subsample of only 27 HC and 26 MDD, in which ~45% of the MDD participants were on psychotropic medications. This questions the reliability of the null findings.

We thank the reviewer for their important remarks. In our primary analysis, all the tests are controlled for subjects' medication status. However, in line with reviewer comments on a subset of the MDD cohort being on psychotropic medications, we now added a new **Supplementary materials section "Medication confounds and replication of key results in the unmedicated MDD subgroup" and Supplementary Figure 7** where we report follow-up analyses restricted to the unmedicated subgroup. The core results are qualitatively consistent within the unmedicated sample, although several effects are attenuated and no longer reach conventional significance ($p > 0.05$), while many show the same directional trends ($p < 0.1$). This attenuation is most likely attributable to reduced power from smaller, unmatched group sizes (n(HC)=38; n(unmedicated-MDD)=21). We have added remarks about these analyses in the main text. Taken together, the supplemental analyses indicate that medication status does not appear to be the primary driver of the reported effects.

Additionally, we fully agree that the smaller DWI cohort could reduce power. Because this study is completed and missing tractography scans cannot be recovered, we adopted a principled sensitivity strategy to evaluate whether the NCT results reflect true structure-function relations rather than idiosyncrasies of the reduced DWI sample. This analysis is added as a **Supplementary Materials section, "Assessing specificity using geometry-preserving null networks", and Supplementary Figure 12**.

In this new sensitivity analysis, we generated 100 geometry-preserving surrogate structural connectomes for each subject with a DWI scan under two established null models (weight–distance preserving, WP; and strength-sequence–distance preserving, SSP) by rewiring that subject's structural connectivity matrix, and computed subject-level surrogate transition-energy (TE) distributions. For subjects without DWI, we created M = 1,000 imputed TE datasets by sampling plausible TE values from the pooled surrogate TE values. Sampling was stratified to reduce bias (k-nearest-neighbor matching with k = 4 on age, sex, medication use, and QIDS scores in 90% of imputations; 10% of imputations were drawn at random from the pool to account for noise). For each imputed dataset we re-ran the same group-level regressions as in the main analysis and grouped parameter estimates, standard errors, t-stats, p-values, and 95% CIs across imputations using Rubin's rule.

According to this analysis, the direction of effects observed in the primary analyses was preserved across both WP and SSP nulls and across the imputed datasets: State 4 to State 1 transitions occurred more frequently with decreased transition energy in healthy controls, while State 3 to State 2 occurred less frequently with lower transition energy. These results were robust and significant under both WP and SSP, except the transition from State 3 to State 2 result was trend-level under SSP ($p \approx 0.08$), indicating model-dependent sensitivity for that contrast. The exit/enter asymmetry pattern replicated in the imputed data. Because medication status was included as a matching variable during imputation, the imputed TE samples reflect the observed distribution of medication use and mitigate, but do not eliminate, bias due to medication.

In sum, these complementary sensitivity checks increase our confidence that the reported NCT findings are not simply coincidental results of the small DWI subsample or of a single null model. Nevertheless, we concur with the reviewer and explicitly note in the Discussion that the limited DWI sample and imputation uncertainty warrant cautious interpretation **(lines: 506-515).**

Other concerns are as follows:

1. The choice of four brain states (k=4) is central to the entire paper. While the authors provide multiple lines of evidence, some justification is weak. The spatial correlation between State 3 in the main dataset and the replicate cluster in the independent dataset was only r=0.30, which is quite low.

We thank the reviewer for this observation. While an r value of 0.30 represents a borderline medium effect, considering it alongside the correlations observed for the other three states (r = 0.56, r = 0.88, and r = 0.59) helps contextualize the overall level of replication. Because the study evaluates reproducibility across two independent datasets, even partial spatial correspondence provides informative evidence about the stability of the derived states.

In addition, the originally reported value reflects a conservative methodological choice for computing spatial similarity between real and replicate clusters. Specifically, after running the k-medoids algorithm for 50 iterations on the replicate dataset, we selected the states from the iteration with the highest inertia within that dataset, and computed correlation between the max inertia states in the replication dataset and max inertia states in the original dataset. To provide a more flexible and comprehensive assessment, we also examined an alternative approach in which clusters from the two datasets were optimally matched using Hungarian assignment, thereby minimizing overall dissimilarity between solutions. This refined analysis yielded r = 0.85, r = 0.74, r = 0.81, and r = 0.64, demonstrating even stronger spatial correspondence and greater robustness of the identified cluster centers across datasets. We have incorporated these results into the revised submission **(lines: 142-149)** and reorganized the supplementary materials accordingly, with these reproducibility analyses now presented in **Supplementary Figure 9.**

Furthermore, the rationale that k>5 leads to states not represented in all subjects is not a strong statistical argument for choosing a lower k; it could simply mean those rarer states are clinically informative.

We thank the reviewer for this helpful comment. We agree that our argument for choosing k=4 due to absence of k>5 states in every subject is not a very strong statistical argument, rather, it is a practical one. Our aim in selecting k was to ensure that each identified state is represented across all subjects, which facilitates interpretable and comparable estimation of state-related metrics (e.g., dwell time, fractional occupancy, transition probabilities) at the group level. For example, one can think of a scenario in which a subset of the participants is missing a state, then, all the quantities we measured (dwell time, fractional occupancy, transition probability etc..) associated to that state would be zero for those participants, simply because their temporal state sequence doesn't contain that state and they can't, for example, dwell on that state. In line with the reviewer's suggestion, we emphasize that determining the optimal number of CAPs is a broader, known, methodological limitation of the approach, which we now state explicitly in the Discussion **(line: 517)** together with an added citation (Liu, X. et al. 2018) to a review paper outlining these challenges.

Regarding the possibility that rarer states may be clinically informative, we agree this is an important consideration. While increasing the number of brain states can reduce bias, it also increases intersubject variability and the risk of overfitting, thereby limiting generalizability. As shown in Supplementary Figure 3 and Supplementary Materials section "Influence of number of clusters", when k = 5, the extra state found was very similar to one of the existing states when k = 4, showing redundancy of these extra states for k >4. The primary goal of our study is not to exhaustively search all possible brain states for each population or individual, but rather to establish a robust baseline framework, consistent with previous work, for comparing the dynamic properties of brain state time series and integrating multimodal measures to explain these dynamics in MDD. Therefore, we consider two balancing criteria for selecting an optimal number of clusters (k): maximizing between-cluster variance and minimizing within-cluster variance, while ensuring that all identified clusters are represented across subjects to enable meaningful group-level comparisons. We now explain these criteria

more carefully at the beginning of the Results section “Canonical RSNs described by spontaneous coactivation patterns” **(lines: 117-129).**

Alternative approaches, such as hidden Markov model, can be supplemented to uncover the dynamic properties of spatial states.

We thank the reviewer for the suggestion to use Hidden Markov Models (HMMs) for state characterization, and we agree that HMMs are indeed complementary and powerful for modeling temporal state transitions. Implementing and validating HMMs in a rigorous manner, however, would considerably expand the methodological scope of the present study and would most appropriately be developed as its own dedicated project. For this reason, we listed this endeavor as a limitation in the Discussion section and encourage future work to explore it **(lines: 515-522).** Our aim here is not to benchmark state-extraction methods but to apply a simple, well-established (in the form of distance-based clustering, implemented as k-means algorithm in the literature), and reproducible clustering framework to link resting-state fMRI with diffusion-derived network control theory and thereby characterize the brain’s energy landscape in depression.

Added novelty to the previous studies, herein, we are using k-medoids algorithm and we justify this choice in a new **Supplementary Materials section (“Justification of the usage of k-medoids algorithm”) and Supplementary Figure 11**. There, we compare four pipelines on three synthetically generated examples: k-means on volumes, k-medoids on volumes, ICA followed by k-means on component space, and ICA followed by k-medoids on component space. We show that medoid-based clustering best recovers physiologically interpretable instantaneous coactivation states. As we mentioned in the Methods Section “Clustering of fMRI volumes” **(lines: 629-637)** and in this new Supplementary Material section, k-medoids selects real volumes as representatives, is robust to outliers and amplitude shifts, and yields temporally coherent state blocks that match visible amplitude regimes in the simulations. By contrast, centroid-based k-means frequently locates synthetic centroids in low-density regions and produces noisier assignments. Dimensionality reduction (ICA) can improve separability but produces component mixtures rather than BOLD signal fluctuations and therefore offers less direct interpretability. In sum, for these reasons, we use k-medoids with a correlation-based dissimilarity for our primary state definitions and downstream inferences.

2. The brain states are defined from the fMRI data of all subjects (HC+MDD). These states are then used as the target configurations for the NCT analysis, which is run on each subject's individual structural connectome. There is a potential circularity: the states are a group-level construct, and the analysis tests how each individual's structure facilitates transitions to these group-defined states. A compelling supplementary analysis would be to repeat the NCT analysis using subject-specific states.

We thank the reviewer for their comment and understand the concern here. Our logic for defining group-level states was to be able to compare each state’s dynamic properties (such as dwell times, fractional occupancy and transition/persistence probabilities) across subjects, since it would be a non-trivial problem to compare these quantities if the states across subjects didn’t match. Similarly, this was also our reasoning for the choice of optimal number of clusters so that each state is represented in every subject’s temporal state sequence, avoiding discrepancies in the quantities computed. However, as the reviewer points out, we are not bound to this criteria for the NCT analysis since transition energy is not a dynamic property of the temporal state sequence, but rather it is a global quantity of the white matter connectivity. Therefore, in line with reviewer’s comments, we added a supplementary analysis in which we repeated our NCT analysis with subject-specific states, which is added to the **Supplementary Material Section “NCT analysis with subject-specific states” and Supplementary Figure 8.**

This analysis revealed that our results regarding structural connectivity facilitates (and constraints) transitions from State 4 to State 1 (and State 3 to State 2) remained qualitatively and statistically unchanged (Supplementary Figure 8c and 8d). Similarly, the characterization of the energy landscape in terms of local maximas and minimas is observed (Supplementary Figure 8b). Interestingly, the calculated transition energy

(TE) values were dispersed on a wider range for each transition, with greater standard deviations and coefficient of variation (CV). For State 4 to State 1: $TE_{subject} \in (1372,3593)$, $\sigma_{subject} = 528.1$, $CV_{subject} = 0.20$ compared to previous result $TE_{group} \in (2155,2291)$, $\sigma_{group} = 32.9$, $CV_{group} = 0.02$, and for State 3 to State 2: $TE_{subject} \in (1675,3900)$, $\sigma_{subject} = 498.3$, $CV_{subject} = 0.16$ compared to previous result $TE_{group} \in (2776,3040)$, $\sigma_{group} = 62.1$, $CV_{group} = 0.02$, with very significant difference between the two ($p < 10^{-10}$). This finding indicates that allowing individual variability in state definitions increases heterogeneity in how structural connectivity supports transitions between functional configurations. In other words, personalized state representations capture subject-specific features of the control energy landscape highlighting individual differences in structure–function coupling rather than random noise. This is similar to the idea, described in Luppi, A.I. et al. 2024, that subject-level morphological differences yielding heterogeneity in transition energies. We once again thank the reviewer for encouraging us to perform this additional analysis as it revealed this interesting result.

3. The introduction criticizes static connectivity measures. However, the discussion does not fully reconcile the dynamic findings with the vast literature on DMN hyperactivity and DMN-FPN hypoconnectivity in MDD. For instance, how does the frequently visited but unstable State 3 (with suppressed DMN) relate to the classic finding of hyperactive DMN?

We thank the reviewer for this insightful comment. We agree that the relationship between our findings and the extensive prior functional connectivity literature warrants explicit clarification. While our CAP methodology (which captures spontaneous amplitude fluctuations in BOLD signal) does not allow for a direct one-to-one comparison with static functional connectivity measures (which measures inter-regional correlations i.e., synchronous activity over a prolonged time), we have now expanded the Discussion to explicitly reconcile State 3 with prior DMN findings in MDD.

We now note in the Discussion section (**lines:373-381)** that static correlations reflect only the time-averaged coupling between regions and therefore cannot distinguish whether reduced connectivity arises from weak inter-regional activity or temporal instability in coordinated activity. Our findings suggest that reduced static DMN-FPN connectivity in MDD may result from frequent, intermittent entry into State 3, which fragments sustained periods of DMN and FPN coactivation and consequently attenuates static coupling strength.

We also clarify (**lines:405-410**) that individuals with MDD did not spend more time in State 2, the state characterized by high-amplitude DMN activity. This suggests that prior reports of DMN hyperactivity may not solely reflect increased engagement of DMN-dominant states. Instead, the increased transitions we observed between State 2 (high-amplitude DMN) and State 3 (low-amplitude DMN) point to greater instability between DMN configurations as a potentially more relevant feature of depressive pathology in our sample.

4. The discussion interprets the results through the lens of complex dynamical systems theory (attractors, basins, energy landscapes). However, the direct neural basis of "control energy" is abstract. The authors correctly note it is not a direct measure of metabolic cost but is more analogous to an external input like neuromodulation (line 435-437). The discussion should include a specific paragraph hypothesizing the neurobiological instantiation of "control energy" or explicitly state that the energy landscape is a computational theory that provides a formal framework for designing experiments to test these biological hypotheses. Above all, the calculated "control energy" is not a direct, biological measure but a theoretical metric derived from a specific simplified model. The absolute energy values and even some relative comparisons could be sensitive to these modeling choices. The conclusions about "energetic cost" are therefore model-dependent.

We thank the reviewer for these insightful suggestions. We now expanded the Discussion section **(lines: 474-487 and 488-502)** carefully explaining the above points and added new citations (Wu, Z. et al. 2024, Karrer, T. M. et al. 2020, Zhou, D. et al. 2025, Kringelbach, M. L. 2020, Beynel, L. et al. 2020, He, X. et al. 2025).

Responses to reviewer comments for "Spatiotemporal asymmetries on brain energy landscape uncover system entrapment related to depression severity"

B. Ülgen Kilic[*1,2], Jenna Jubeir[1,3], Priti Balchandani[3,4,5,6], James W. Murrough[1,3,7], Laurel S. Morris[1,8,9], Yael Jacob[1,3,4,5,6]

1. The Dennis S. Charney, MD, Depression and Anxiety Discovery Center, Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY, USA
2. Center for Computational Psychiatry, Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY, USA
3. Nash Family Department of Neuroscience & Friedman Brain Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA
4. BioMedical Engineering and Imaging Institute, Department of Radiology, Icahn School of Medicine at Mount Sinai, New York, NY, USA
5. Center for Engineering and Precision Medicine, Icahn School of Medicine at Mount Sinai & Rensselaer Polytechnic Institute, New York, NY, USA
6. Department of Diagnostic, Molecular and Interventional Radiology, Icahn School of Medicine at Mount Sinai, New York, NY, USA
7. VISN 2 Mental Illness Research, Education and Clinical Center (MIRECC), James J. Peters VA Medical Center, Bronx, NY, USA
8. Nuffield Department of Clinical Neurosciences, University of Oxford, UK
9. Department of Experimental Psychology, University of Oxford, UK

We thank the editor and reviewers for their time and support in publication of our paper. We are now submitting a minimally modified version of our paper in which the editorial requests and Reviewer 1's comments are addressed. Since Reviewer 2's minor suggestions are conflicting with editorial requirements, we leave the decision to implement those to the editor.

Below, we provide responses to reviewer comments.

Sincerely,

B. Ülgen Kilic

* bengier.kilic@mssm.edu

Reviewer #1 (Remarks to the Author):

This manuscript integrates co-activation pattern (CAP) analysis with network control theory (NCT) to characterize a state-transition "energy landscape" in major depressive disorder (MDD), linking specific dynamic properties to symptom dimensions (e.g., anhedonia, depression severity). The revision substantially strengthens the work through multiple targeted robustness checks and improved statistical reporting. Most major concerns have been addressed, but a few points still require clarification and/or more cautious framing to fully support the central claims.

We are pleased that our revisions were able to address the reviewer's concerns, and we thank the reviewer for their positive comments and support.

Q1：k-medoids clustering is performed on all subjects (HC+MDD) and all time points, which is a common CAP approach, but it can still bias state definitions toward group effects in small samples.

We agree with the reviewer that clustering in modest samples can increase the possibility of bias in CAP definition. In the present study, we sought to reduce this risk by using a robust k-medoids framework for extraction of brain-states, ultra-high-field 7T MRI to maximize signal quality, and multiple sensitivity analyses that supported the stability of the main findings. In addition, prior work (An et al. 2024) has reported similar CAP descriptions, suggesting that the states identified here are not idiosyncratic to this sample. We also note that our Limitations section acknowledges concerns related to sample size and the dependence of downstream analyses on CAP definitions.

Q2：Educational level is an important covariate for brain imaging data, but it was not taken into account in this study.

We thank the reviewer for raising this point. We examined group differences in educational attainment across HC and MDD participants. Education did not significantly differ between groups across categories (fMRI dataset: p=0.259, DWI dataset: p=0.274; chi-square test). We now included population means of education level in estimated years and significance test results in the Supplementary Table 1 presented with other demographical information.

Q3：Given that the mechanistic narrative is grounded primarily in linear NCT-derived asymmetries and energy–probability coupling, the "attractor/entrapment" framing should be clearly presented as a model-based metaphor (or operationalized explicitly via enter/exit energy asymmetry), avoiding implications of directly demonstrated nonlinear attractor dynamics.

We thank the reviewer for this important point and agree that our original wording could be interpreted as implying directly demonstrated nonlinear attractor dynamics. To clarify this distinction, we revised the Discussion to use more cautious and operational language throughout.

Specifically, we changed:

- Line 388: attractor dynamics→ state transition dynamics
- Line 412: aberrant attractor dynamics → attractor-like instability
- Line 418: stable attractor basins→ stable configurations
- Line 420: two-state attractor itself→two-state motif

In addition, we added the following clarifying sentence in lines 478-481:

- "Accordingly, our linear NCT framework does not directly demonstrate nonlinear attractor dynamics, yet this language is used heuristically to describe recurrent transition motifs and asymmetries in modeled control energy."

Reviewer #3 (Remarks to the Author):

The authors have addressed all my concerns. I only have a few minor suggestions:
We are happy that we were able to remedy reviewer concerns and thank them for their suggestions which overall improved our paper.

1. In the abstract, instead of "state-transitions are perturbed", the authors may want to specify the nature of the perturbation investigated (e.g., "...how the stability and energetic landscape of state-transitions are perturbed") to sharpen the opening motivation.
We would be happy to make this modification. However, the manuscript is currently at the 150-word limit imposed by the editorial requirements. If the editor is willing to allow a minor exception to that limit, we can incorporate this change in the revised version.

2. Please consider citing the recent work by Yu et al. (2025,doi.org/10.1093/psyrad/kkaf008) on white matter functional connectome gradients in major depressive disorder. This study provides important complementary evidence of hierarchical dysfunction within the white matter's intrinsic functional architecture, which would help contextualize your structural control findings within the broader framework of network-level dysregulation in depression.
We agree that this recent white matter gradient study in MDD is a relevant paper adjacent to the methods used here. However, we are currently already above the editorial 70-reference limit. If the editor is willing to allow the addition of one more reference, we would be happy to include it in the revised manuscript.