Multiscale Statistical Models for Understanding Tumor Microenvironment Heterogeneity

by

Joel Nathan Eliason

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Bioinformatics)
in the University of Michigan
2025

Doctoral Committee:

Professor Arvind Rao, Chair
Professor Jared K. Burks, MD Anderson Cancer Center
Associate Professor Gen Li
Professor Marina Pasca di Magliano
Assistant Professor Michele Peruzzi

Joel Nathan Eliason

joelne@umich.edu

ORCID iD: 0000-0003-2227-8727

© Joel Nathan Eliason 2025

DEDICATION

To Mom, who raised me as an idealist

To Dad, for giving me a love of science and math

To Corie, who supported me day in and day out

To Louisa, Jasper and baby number three, who bring me childlike joy every day

ACKNOWLEDGEMENTS

There are so many people I would like to acknowledge in helping me finish this dissertation!

First, my committee members. Dr. Gen Li initially brought me on as a research assistant when I was a masters student in the Biostatistics department at Michigan, and was instrumental in helping me start to become more comfortable with developing statistical methodology. Dr. Marina Pasca di Magliano also served on my prelim committee and has always been able to help clarify the biological interpretability in my work, something that I have had to become more comfortable with doing myself. Dr. Jared Burks (and Dr. Suhendan Ekmekcioglu!) were both such pleasant collaborators to work with at MD Anderson Cancer Center - always interested and focused on the biological implications of my analyses and always willing to hear out whatever strange analysis ideas I came up with. Dr. Michele Peruzzi has helped me continue the journey of statistical methodology development that Dr. Li started, and his feedback has been invaluable at many points in the development of these projects. Finally, many thanks to my advisor, Dr. Arvind Rao, both for granting me the autonomy and support to work on the research directions I found most compelling as well as permitting me to work remotely on my PhD, which was huge for my wife and I. Moving closer to our families has been really great not only for us but our children as well.

To my in-laws Paul and Kelli - thanks for your support as well! You always asked me how I was doing and tried to understand my work, and that always made me feel seen. Thank you.

To Gpa! It has been so wonderful living close to you now. Watching you play with our kids has been wonderful. Thanks for helping us out with childcare and other odds and ends - that has been crucial during our busy times.

Mom and Dad - how can I thank you enough? You raised me and motivated me in school, you gave me a love for science that just wouldn't quit, you have always been there for me when I needed to call. Dad - it's always been fun to talk math and science with you. I can still remember you showing me the proof to the Pythagorean Theorem in chalk on our driveway. Thank you both so much for your support.

Finally to my wife, Corie - you have been an absolute rock during the turbulent times of my academic life. Thank you for sticking with me, for tirelessly supporting me day in

and day out, for being my biggest cheerleader and always wanting to understand what I was working on. I love you so much, and thank you so much for taking a chance on me.

PREFACE

This dissertation reflects my PhD work at the intersection of statistical modeling and cancer biology, shaped by a long-standing interest in the complex ecology of the spatial tumor microenvironment. I focused on developing spatial statistical methods for analyzing high-dimensional tissue imaging data, with the broader aim of understanding how the spatial organization of cells in the tumor microenvironment influences disease progression and patient outcomes. The chapters span both applied and methodological contributions, guided by a belief that careful and structured modeling can help bring clarity to biological questions that are often messy and data-limited.

Several chapters are based on published or in-progress collaborative research. Chapter 2 is based on a publication titled *Characterizing spatial immune architecture in metastatic melanoma using high-dimensional multiplex imaging* in *Frontiers in Immunology* (April 2025). This work was conducted in collaboration with several coauthors; in particular, I would like to acknowledge Santhoshi Krishnan and Su Ekmekcioglu for their contributions to the statistical and biological analysis. Chapter 3 is based on the paper *Investigating Ecological Interactions in the Tumor Microenvironment Using Joint Species Distribution Models for Point Patterns*, coauthored with Arvind Rao and published in the *New England Journal of Statistics in Data Science* (June 2024).

Chapters 4 and 5 represent ongoing collaborative projects with Michele Peruzzi and Arvind Rao. Chapter 6 and Chapter 7 reflect more independent work–Chapter 6 focuses on the development of a Bayesian survival model with functional spatial covariates, and Chapter 7 explores simulation-based inference in mechanistic models of tumor–immune dynamics, representing a more exploratory and forward-looking direction in my research.

While the methods vary in scope and application, they are all motivated by a desire to connect spatial statistical modeling with real biological questions concerning the tumor microenvironment. I hope this work contributes in some small way to the broader effort of making analyses of spatial tissue data more interpretable, reproducible, and clinically meaningful.

All work presented here was conducted during my doctoral studies at the University of Michigan.

TABLE OF CONTENTS

DEDICATION ii

ACKNOWLEDGEMENTS iii

PREFACE v

LIST OF FIGURES x

LIST OF TABLES xviii

LIST OF APPENDICES xix

LIST OF ACRONYMS xx

ABSTRACT xxiii

CHAPTER

1 Introduction 1

1.1 Motivation and Background 1
1.2 Gaps in Existing Approaches 2
1.3 Dissertation Aims and Research Questions 3
1.4 Overview of Contributions 4
1.4.1 Overview of a Typical Multiplex Imaging Data Structure 6
1.5 Organization of the Dissertation 7

2 Spatial Analysis of Immune Neighborhoods in Melanoma 9

2.1 Introduction 9
2.2 Materials and Methods 10
2.2.1 Patient Samples and Datasets 10
2.2.2 Acquisition and Integration of IMC and mIF Imaging Data 12
2.2.3 Statistical Analysis 12
2.3 Results 13
2.3.1 Summary of Stage III Patient Characteristics 13
2.3.2 Clinical and Demographic Characteristics of Stage IV TMA Patients 14
2.3.3 Differences in Immune Cell Composition between Stages 14
2.3.4 Shifts in Inflammatory Neighborhoods 16
2.3.5 Immune Cell Distribution Within Inflammatory Neighborhoods . . 17

2.3.6 Neighborhood Representation Within Immune Cell Populations . . 21
2.3.7 Spatial Immune Interactions and Their Association with Patient Survival . . . 22
2.3.8 Architectural Features and Cellular Neighborhoods Associated with Inflammatory Nodes . . . 25
2.4 Discussion . . . 28
2.4.1 Spatial Immune Reorganization in Stage IV Melanoma . . . 28
2.4.2 Immune Interactions and Survival Outcomes . . . 29
2.4.3 TAM and CTL Dynamics in the TIME . . . 30
2.4.4 Clinical and Therapeutic Implications . . . 30
2.4.5 Study Strengths and Limitations . . . 31
2.4.6 Conclusions and Future Directions . . . 31

3 Multitype Gibbs Modeling of Cell–Cell Interactions and Survival Association . . . 32

3.1 Introduction . . . 32
3.2 Methods . . . 36
3.2.1 Dataset . . . 36
3.2.2 Multitype Gibbs Point Process (MGPP) Models . . . 37
3.2.3 An Example Analysis . . . 41
3.3 Results from Example Analysis . . . 46
3.3.1 Interpretation of Model Parameters and Predictions . . . 46
3.3.2 Heterogeneity in Model Parameter Estimates Within Patient Groups and Across Patients . . . 50
3.3.3 Predictive Modeling of Patient Outcomes . . . 51
3.3.4 Evaluation of Model Fit . . . 52
3.3.5 Comparison with Previous Analysis . . . 53
3.4 Discussion . . . 55

4 The SHADE Model for Asymmetric Multilevel Spatial Interactions . . . 59

4.1 Introduction . . . 59
4.2 Methods . . . 62
4.2.1 Hierarchical Modeling of Conditional Spatial Point Processes . . . 62
4.2.2 Multilevel Bayesian Model . . . 66
4.2.3 Logistic Regression Approximation for Computational Efficiency . . . 67
4.2.4 Summary of Estimation Procedure . . . 68
4.3 Simulation Study . . . 68
4.3.1 Number of cells of each type and number of dummy points . . . 69
4.3.2 Number of patients and images per patient . . . 70
4.3.3 Hierarchical vs non-hierarchical modeling . . . 71
4.4 Results: Multiscale inference of directional spatial interactions in CRC . . . 73
4.4.1 Description of colorectal cancer dataset . . . 73
4.4.2 How are specific cell types spatially organized around others in the tumor microenvironment? . . . 74

4.4.3 How do spatial interaction patterns vary across patients and tissue sections? . . . 75
4.4.4 How do immune and stromal spatial interactions differ between hot and cold tumors? . . . 77
4.4.5 How well can spatial interactions predict the organization of individual cell types across tumor regions? . . . 79
4.5 Discussion . . . 80

5 The BIPPS Model for Shared Spatial Dependencies Across Patients . . . 83

5.1 Introduction . . . 83
5.2 Many-subject multiplexed imaging datasets . . . 84
5.2.1 Pancreatic cancer dataset . . . 84
5.2.2 Colorectal cancer dataset . . . 86
5.3 Multi subject Bayesian latent spatial factor model . . . 88
5.3.1 Data preprocessing workflow . . . 91
5.4 Simulation studies . . . 91
5.4.1 Recovering cross-correlations when the number of latent factors is misspecified . . . 91
5.4.2 The effect of grid size on cross-correlation recovery . . . 94
5.5 Application to pancreatic cancer data . . . 95
5.5.1 Spatial Organization of the Pancreatic Cancer TME Across Patient Groups . . . 97
5.6 Application to CRC data . . . 99
5.6.1 Spatial Organization of the CRC TME Across Patient Groups . . . 100
5.7 Discussion . . . 103

6 Bayesian Survival Modeling with Functional Spatial Covariates . . . 105

6.1 Introduction . . . 105
6.2 Methods . . . 106
6.2.1 Functional Covariate Construction and Uncertainty Quantification . 106
6.2.2 Modeling the Spatially-Integrated Functional Covariate . . . 107
6.2.3 Survival Modeling and Bayesian Inference . . . 108
6.2.4 Joint Posterior Distribution . . . 108
6.3 Simulation Studies . . . 109
6.3.1 Effect of Sample Size on Model Efficiency . . . 110
6.4 Results . . . 110
6.4.1 Basel Breast Cancer Dataset . . . 110
6.4.2 Estimation of Functional Covariates . . . 111
6.4.3 Model Fitting and Posterior Inference . . . 112
6.4.4 Visualization of Spatial Covariate and Effect Function Estimates . . 112
6.5 Discussion . . . 114

7 Case Studies in Simulation-Based Inference for Mechanistic Tumor–Immune Modeling . . . 117

7.1 Simulation-Based Inference for Simple Models of Tumor-Immune Dynamics: A Proof of Concept . . . 118

7.2 Simulation-Based Inference for a Mechanistic QSP Model of PD-1/PD-L1 Inhibition . . . 121

8 Discussion . . . 128

8.1 Summary of Contributions . . . 128
8.2 Methodological Themes and Lessons Learned . . . 129
8.3 Biological and Clinical Insights . . . 130
8.4 Limitations and Open Questions . . . 132
8.5 Future Directions . . . 134
8.6 Closing Thoughts . . . 135

APPENDICES . . . 136

A.1 Supplementary Figures . . . 136
A.1.1 The impact of dummy points on inference . . . 136
A.1.2 Number of patients and images per patient . . . 139
A.1.3 CRC data analysis . . . 141
A.2 Priors for Variance Parameters . . . 144
A.3 Simulation Study Details . . . 144
A.3.1 Generating Target Points . . . 144
A.3.2 Default Parameter Settings . . . 144
A.4 Quantifying Variability in Spatial Interactions . . . 145
B.1 Simulation Details . . . 147
B.2 Supplemental Results . . . 148
B.2.1 Recovering Cross-Correlations When the Number of Latent Factors is Misspecified . . . 148
B.2.2 The Effect of Grid Size on Cross-Correlation Recovery . . . 159
B.2.3 Choosing k for PANC Analysis . . . 164
B.2.4 Analysis of PANC Dataset . . . 169
B.2.5 Choosing k for CRC Analysis . . . 171
B.2.6 Analysis of CRC Dataset . . . 177
B.3 Markov-chain Monte Carlo for posterior sampling . . . 179
B.4 Implementation details . . . 180
C.1 Figures . . . 181
C.2 Tables . . . 185
C.3 Methods . . . 186
C.3.1 Simultaneous Credible Bands for Functional Coefficients . . . 186

BIBLIOGRAPHY . . . 188

LIST OF FIGURES

FIGURE

1.1 A multitype point pattern representation of a multiplex image of a tissue section of the TME in a colorectal cancer patient. 8

2.1 Analysis pipeline to characterize the spatial immune landscape of Stage III and IV melanomas. Schematic of the IMC data acquisition of two consecutive slices of two TMAs containing 2 biopsy cores from a total of 157 patients with stage III and 393 samples from 248 patients with stage IV metastatic melanoma. Samples were stained with a protein panel, segmented for tumor/stroma regions detection, followed by cellular identification and quantifications. t-distributed stochastic neighbor embedding (t-SNE) of data was derived from CyTOF of tumor samples, labeled by cell type and the signal intensities of individual markers. 11

2.2 Kaplan-Meier survival curves comparing overall survival (OS) probability between Stage III and Stage IV patients. The shaded regions represent 95% confidence intervals. The p-value indicates the statistical significance of the difference between the survival distributions (log-rank test). The risk table below the plot shows the number of patients at risk at each time point. 17

2.3 Distribution of cell type proportions across Stage III and Stage IV melanoma. Each boxplot represents the proportion of a given cell type per image, grouped by stage. Statistical significance of difference in mean proportion was assessed using a beta regression model, with significance markers (*, **, ***) indicating significant differences between stages after false discovery rate correction. Significance levels are denoted as follows: $***(p < 0.001)$, $**$ $(p < 0.01)$, and $*(p < 0.05)$. 18

2.4 Distribution of neighborhood proportions across Stage III and Stage IV melanoma. Each boxplot represents the proportion of a given inflammatory neighborhood per image, grouped by stage. Statistical significance of difference in mean proportion was assessed using a beta regression model, with significance markers (*, **, ***) indicating significant differences between stages after false discovery rate correction. Significance levels are denoted as follows: $***(p < 0.001)$, $**$ $(p < 0.01)$, and $*(p < 0.05)$. 19

2.5 Mean proportions of different cell types within inflammatory neighborhoods in Stage III and Stage IV melanoma. Each tile is annotated with the mean proportion of the corresponding cell type among cells in that neighborhood. In the Stage IV column, a forward slash ("/") separates the mean proportion from statistical significance markers (*, **, ***), indicating significant differences in mean proportion between stages based on a beta regression model with false discovery rate correction. Significance levels are denoted as follows: $***(p < 0.001)$, $**$ $(p < 0.01)$, and $*(p < 0.05)$. Text color is adjusted for readability. 20
2.6 Mean proportions of inflammatory neighborhoods as fractions of cells within each cell type in Stage III and Stage IV melanoma. Each tile is annotated with the mean proportion of the corresponding neighborhood among cells within that cell type. In the Stage IV column, a forward slash ("/") separates the mean proportion from statistical significance markers (*, **, ***), indicating significant differences in mean proportion between stages based on a beta regression model with false discovery rate correction. Significance levels are denoted as follows: $***(p < 0.001)$, $**$ $(p < 0.01)$, and $*(p < 0.05)$. Text color is adjusted for readability. 21
2.7 Heatmap of log-hazard ratios estimated using Cox regression, where the G-cross function at 40 μm between each pair of cell types in Stage III patients serves as a predictor of survival time. Significance levels of hazard ratios are denoted as follows: $***(p < 0.001)$, $**$ $(p < 0.01)$, and $*(p < 0.05)$. 23
2.8 Heatmap of log-hazard ratios estimated using Cox regression, where the G-cross function at 40 μm between each pair of cell types in Stage IV patients serves as a predictor of survival time. Significance levels of hazard ratios are denoted as follows: $***(p < 0.001)$, $**$ $(p < 0.01)$, and $*(p < 0.05)$. 24
2.9 Plots illustrating the significance of hazard ratios ($p < 0.05$) estimated from Cox proportional hazards models, where the G-cross function between each pair of cell types serves as a predictor of survival in stage III patients. Significant associations with worse outcomes are shown in red, while those with better outcomes are shown in blue. Within each subplot, estimates are presented relative to a focal cell type located within a specific inflammatory neighborhood, at radii of 20, 40, and 60 μm. 26
2.10 Plots illustrating the significance of hazard ratios ($p < 0.05$) estimated from Cox proportional hazards models, where the G-cross function between each pair of cell types serves as a predictor of survival in stage IV patients. Significant associations with worse outcomes are shown in red, while those with better outcomes are shown in blue. Within each subplot, estimates are presented relative to a focal cell type located within a specific inflammatory neighborhood, at radii of 20, 40, and 60 μm. 27

3.1 A point pattern representation of a multiplex image of a tissue section of the TME in a CRC patient 37
3.2 Examples of potential curves commonly used for ϕ and ψ in Equations 3.4 and 3.5. 40
3.3 Spatial covariate of "distance to nearest blood vessel", as imaged in spot 59_A. . 43

3.4 Interaction parameters α_{m_i,m_j} and γ_{m_i,m_j} as estimated from spot 59_A. Stars indicate that the interaction parameter was found to be significant within that image at a significance level of 0.05. 47
3.5 Difference in the median of the short-range interaction parameters α_{m_i,m_j} and γ_{m_i,m_j} between patient groups. 48
3.6 The short-range, medium-range and combined potentials between B cells and memory CD4+ T cells. From model fit on spot 59_A. 49
3.7 The short-range, medium-range and combined potentials between a subset of cell types. From model fit on spot 59_A. 50
3.8 The top contributor to the predicted intensity of tumor cells. From model fit on spot 59_A. 51
3.9 Estimates of the median absolute deviation (MAD) for α parameter estimates, across all patients. 52
3.10 Estimates of the median absolute deviation (MAD) for α parameter estimates, in both patient groups. 53
3.11 Log-hazard ratios estimated from short- and medium-range interaction parameters that were found to be significant after FDR adjustment. 54
3.12 The predicted log-conditional intensity of plasma cells from the model fit for spot 59_A. The AUC for this cell type was 0.835, indicating high predictive ability for plasma cells. 55

4.1 Summary of the SHADE framework. **A)** Multiplexed imaging data is structured hierarchically across cohorts, patients, and images. Multiplex images were adapted from Schurch et al. (2020) and are used under a Creative Commons CC BY 4.0 license. **B)** Images are processed into spatial point patterns with cell type annotations. **C)** SHADE estimates Spatial Interaction Curves (SICs) that capture directional associations between cell types across spatial scales. **D)** SICs are estimated at cohort, patient, and image levels, enabling multilevel analysis of spatial heterogeneity. **E)** Posterior distributions provide uncertainty quantification. **F)** SICs can be compared across cohorts to assess differences in spatial organization. 61
4.2 An example spatial interaction curve showing the effect of a source cell type A_k on a target cell type B. At each distance s, the curve value represents the change in log-intensity of type B associated with the presence of a type A_k cell at distance s. 64
4.3 Estimated SICs for two source cell types. Type 2 shows a strong positive association with the target population at short distances, while Type 1 exhibits negative association at short range. These contrasting patterns highlight the flexibility of the SIC framework for capturing biologically meaningful spatial structure. . . . 65
4.4 Spatial interaction fields implied by SICs from two source cell types (Panels 1 and 2) and their additive combination (Panel 3). The combined field governs the spatial distribution of target cells, which are more likely to occur in regions of elevated log-intensity. 66

4.5 RMSE for spatial interaction coefficients across spatial scales, shown separately for each hierarchical level: $\delta^{(m,p)}_{t_1\to t_2}$, $\gamma^{(n,p)}_{t_1\to t_2}$, and $\psi^{(g,p)}_{t_1\to t_2}$. Results are shown for varying cell counts and dummy point ratios. 70
4.6 Average RMSE across spatial scales for spatial interaction coefficients at the image ($\delta^{(m,p)}_{t_1\to t_2}$), patient ($\gamma^{(n,p)}_{t_1\to t_2}$), and cohort ($\psi^{(g,p)}_{t_1\to t_2}$) levels, across varying numbers of patients and images per patient. 72
4.7 Examples of estimated image-level SICs from hierachical modeling simulations, demonstrating better estimation of SICs when accounting for multilevel structure. 73
4.8 SIC showing the directional association of CAFs (source) with CTLs (target), stratified by patient group. Solid lines indicate cohort-level estimates for CLR and DII groups, while dashed lines represent patient-level SICs, highlighting within-group variability. The plot illustrates how CTLs tend to slightly cluster around CTLs in CLR patients at mid-to-long ranges (positive log-intensity), whereas a relative neutral to avoidant behavior is observed in DII patients—demonstrating a group-specific spatial pattern in CAF-CTL organization. 76
4.9 Heatmaps show the median absolute deviation (MAD) of SICs across spatial distances, summarizing two sources of variability: between patients (within cohort) and between images (within patient). Higher values indicate greater heterogeneity in spatial structure. 77
4.10 Examples of patient- and image-level SICs for cell type pairs with high within-patient MAD values. Each dotted line represents a SIC estimated from a single tissue section, while each solid line indicates a patient-level SIC. 78
4.11 The difference in SICs between the two patient groups. Curves above 0 (in the blue area) indicate that the log-intensity is greater in CLR patients than DII patients; curves in the red area show the reverse. 79
4.12 Predicted spatial distributions for target cell types in a representative image (47_B), based on the estimated conditional intensity functions. 80

5.1 Overview of the imaging and statistical analysis workflow. Biopsy samples are first collected from patients across different cohorts. These biopsies are then processed and imaged using multiplex imaging techniques, allowing for high-dimensional spatial characterization of the tumor microenvironment. Following imaging, cell phenotyping is performed to classify cells based on marker expression, generating spatial point patterns for each sample. These point patterns are subsequently converted into multivariate binned count data, which serves as the structured input for the statistical method presented in this study. This figure was developed primarily by Michele Peruzzi. 85
5.2 Point patterns derived from selected images in the pancreatic cancer dataset, including both IPMN and PDAC cases. **A)** Two images from each cancer type. **B)** IPMN image 909-1 is transformed into the binned count representation. The visualized areas in each image are approximately 1300 μm (x-axis) by 1000 μm (y-axis). 87

5.3 Point patterns derived from selected images in the colorectal cancer dataset, including both CLR and DII cases. **A)** Two images from each cancer type. **B)** DII image 59-A is transformed into the binned count representation. The visualized areas in each image are approximately 1800 μm (x-axis) by 1400 μm (y-axis). 88
5.4 Example simulation from study on misspecification of k. a) Simulated data for a single image. b) Underlying latent factors of simulated data. c) Recovered latent factors. 92
5.5 MAD of cross-correlation curves at various distances between observed and predicted when fitting with various k. Here the generative model has $k^* = 3$. . . . 93
5.6 Simulated data for a single image with a) $n_x = 12$, b) $n_x = 24$ and c) $n_x = 48$. . 95
5.7 MAD of cross-correlation curves at various distances between observed and predicted when fitting with various numbers of pixels. 96
5.8 The estimated spatial cross-correlations between all pairs of cell-type combinations from patients in the IPMN group and the PDAC group. The colored line in each plot indicates the posterior expectation, while the shaded region represents the pointwise 95% credible interval. 98
5.9 The estimated spatial cross-correlations between select pairs of cell-type combinations from patients in the CLR group and the DII group. The colored line in each plot indicates the posterior expectation, while the shaded region represents the pointwise 95% credible interval. 101

6.1 RMSE for recovery of the spatially-varying regression coefficient $\beta(s)$ and subject-specific latent functions $X_i(s)$ across varying numbers of patients and repeated measurements per patient. Each panel shows the distribution of RMSE across 100 simulated datasets. Estimation of $\beta(s)$ improves with increasing cohort size, while estimation of $X_i(s)$ benefits primarily from additional repeated measurements. 111
6.2 For each pair of cell types, posterior means and simultaneous 95% credible intervals were calculated for each spatial location using a maximum standardized residual approach. For a subset of cell type pairs, posterior summaries were plotted as red curves with ribbons. For comparison, we fit standard Cox models using three sampled spatial locations from the G-cross curves ($s = 20, 40, 60$ microns) as covariates, estimating pointwise log-hazard ratios and confidence intervals. These are overlaid as black point-ranges. Facets are organized by neighboring (Type 2) and focal (Type 1) cell type pairs. 113
6.3 For a subset of patients, the posterior mean and 95% simultaneous credible interval for $X_i(s)$ were plotted alongside the observed noisy measurements. Observed curves are shown in light color, with dashed lines indicating ± 2 standard deviations based on bootstrap-estimated sampling error. The estimated posterior function for each patient is overlaid as a solid line with a light ribbon indicating the credible interval. Facets are arranged by interaction (e.g., Macrophages $\rightarrow$ T cells) and patient ID. 116

7.1 Simulated trajectories of tumor and immune populations across multiple parameter settings. . . . 119
7.2 Posterior means versus ground-truth parameter values across held-out validation simulations. Each point represents the average posterior estimate from amortized inference using BayesFlow. . . . 121
7.3 Simulation-based calibration using rank ECDFs for each inferred parameter. . . . 122
7.4 Posterior predictive distribution over time. . . . 123
7.5 Posterior distributions as estimated via Stan. . . . 124
7.6 Training and validation loss over epochs during model training. . . . 125
7.7 Simulation-based calibration using ECDF rank plots for each inferred parameter from the QSP model. . . . 126
7.8 Posterior means versus true parameter values on held-out validation data in the QSP model. . . . 127
7.9 z-score contraction plot comparing prior and posterior distributions. . . . 127

A.1 Average RMSE for spatial interaction coefficients across hierarchical levels: image-level ($\delta^{(m,p)}_{t_1\to t_2}$), patient-level ($\gamma^{(n,p)}_{t_1\to t_2}$), and cohort-level ($\psi^{(g,p)}_{t_1\to t_2}$), under different numbers of cells per type and dummy-to-real point ratios. . . . 137
A.2 Example of estimated cohort-level SIC from dummy point simulations. . . . 138
A.3 Average RMSE for spatial interaction coefficients as a function of dataset size, varying the number of patients and number of images per patient. Results are shown for image-level ($\delta^{(m,p)}_{t_1\to t_2}$), patient-level ($\gamma^{(n,p)}_{t_1\to t_2}$), and cohort-level ($\psi^{(g,p)}_{t_1\to t_2}$) parameters. . . . 139
A.4 Example of an estimated cohort-level SIC from the dataset size simulation study, illustrating how increasing the number of patients and images per patient reduces the variance in SIC estimates. . . . 140
A.5 Radial basis functions $\phi_p(s)$ used to compute distance-based interaction features $\boldsymbol{q}_{A_k}(v)$ in the CRC analysis. These basis functions define the spatial resolution of the SICs. . . . 141
A.6 Cohort-level SICs (parameterized by $\psi^{(g,p)}_{t_1\to t_2}$) estimated for all source–target cell type pairs in the CRC dataset, stratified by CLR and DII patient groups. . . . 142
A.7 Distribution of AUCs by cell type and patient group, evaluating prediction performance of SHADE's conditional intensity model when predicting the spatial organization of each target cell type based on the spatial distribution of all other types. . . . 143

B.1 An example of cross-correlation curves estimated using $k = 2$, when $k^* = 3$. . . . 150
B.2 An example of cross-correlation curves estimated using $k = 3$, when $k^* = 3$. . . . 151
B.3 An example of cross-correlation curves estimated using $k = 6$, when $k^* = 3$. . . . 152
B.4 An example of the difference in two sets of cross-correlation curves that were both estimated using $k = 2$, when $k^* = 3$. . . . 153
B.5 Effective sample sizes (ESS) of cross-correlation estimates at different spatial distances, over varying k. . . . 154
B.6 $\hat{R}$ of cross-correlation estimates at different spatial distances, over varying k. . . . 155
B.7 The distribution of WAIC for different k, when k^*=3. . . . 156

B.8 Example of MCMC trace plot for estimated mean-centered cross-correlations at $h = 767.6\mu m$, using $k = 2$. 157
B.9 Model fitting times for simulations at various k. 158
B.10 A comparison of cross-correlation curves estimated using $n_x \in \{12, 24, 48\}$ . . . 160
B.11 Effective sample sizes (ESS) of cross-correlation estimates at different spatial distances, over varying n_x. 161
B.12 $\hat{R}$ of cross-correlation estimates at different spatial distances, over varying n_x. . 162
B.13 Model fitting times for simulations at various n_x. 163
B.14 $\hat{R}$ values to determine convergence of MCMC sampling of cross-correlations at various distances, under various k, in the PDAC patient group. Values closer to 1 indicate better convergence. 165
B.15 $\hat{R}$ values to determine convergence of MCMC sampling of cross-correlations at various distances, under various k, in the IPMN patient group. Values closer to 1 indicate better convergence. 166
B.16 Estimated cross-correlations at various distances, estimated at various k, in the PDAC patient group. 167
B.17 Estimated cross-correlations at various distances, estimated at various k, in the IPMN patient group. 168
B.18 Estimated posterior mean of the zero-distance ($h = 0$) cross-correlation matrix between the different cell types in both patient groups from the pancreatic cancer dataset. 169
B.19 The estimated difference in cross-correlation between patient groups for the select cell types. The black line in each plot indicates the posterior expectation, while the shaded region indicates the pointwise 95% credible interval. 170
B.20 $\hat{R}$ values to determine convergence of MCMC sampling of cross-correlations at various distances, under various k, in the CLR patient group. Values closer to 1 indicate better convergence. 172
B.21 $\hat{R}$ values to determine convergence of MCMC sampling of cross-correlations at various distances, under various k, in the DII patient group. Values closer to 1 indicate better convergence. 173
B.22 Estimated cross-correlations at various distances, estimated at various k, in the CLR patient group. 174
B.23 Estimated cross-correlations at various distances, estimated at various k, in the DII patient group. 175
B.24 Difference in estimated cross-correlations at various distances, estimated at various k, between the two patient groups. 176
B.25 The estimated spatial cross-correlations between select pairs of cell-type combinations from patients in the CLR group and the DII group. The colored line in each plot indicates the posterior expectation, while the shaded region indicates the pointwise 95% credible interval. 177
B.26 The estimated difference in cross-correlation between patient groups for the same selected cell types as in the main article. The black line in each plot indicates the posterior expectation, while the shaded region indicates the pointwise 95% credible interval. 178

C.1 Examples of posterior estimates of patient-level functional covariates $X_i(s)$ with 95% simultaneous credible bands, for a simulation with $N = 50$ patients and $M = 2$ images per patient. Each facet displays one simulated patient. The true curve $X_i(s)$ (blue) is overlaid with the posterior mean estimate (red) and shaded band denoting the 95% uniform credible interval. 181
C.2 Posterior estimate of the spatially-varying effect $\beta(s)$ with 95% simultaneous credible band, for a simulation with $N = 50$ patients and $M = 2$ images per patient. The true effect (blue) is overlaid with the posterior mean estimate (red) and a shaded region denoting the 95% uniform credible interval. 182
C.3 Average width of the simultaneous 95% credible intervals for the spatially-varying coefficient function $\beta(s)$. Both increased patient count and repeated measurements lead to narrower intervals, indicating improved precision in estimating population-level spatial effects. 183
C.4 Average width of the simultaneous 95% credible intervals for patient-specific functional covariates $X_i(s)$, across varying numbers of patients and repeated measurements. Increasing the number of images per patient substantially reduces interval width, reflecting improved uncertainty quantification for individual-level functions. The number of patients has minimal additional effect on this quantity. 184

LIST OF TABLES

TABLE

2.1 Summary of key clinical and demographic characteristics for stage III TMA patients. . . . 15
2.2 Summary of key clinical and demographic characteristics for stage IV TMA patients. . . . 16

3.1 Examples of questions that can be asked using MGPP models as JSDMs. . . . 34
3.2 Coefficients β_{1,m_j} as estimated from spot 59_A. . . . 46
3.3 Median AUC for each cell type, across model fits from all images within each group and across the whole patient cohort . . . 56
3.4 Abundance of each cell type in each patient group. . . . 57
3.5 Correlation between AUC score and abundance for each cell type. . . . 58

4.1 Average RMSE of $\delta_{t_1 \to t_2}^{(m,p)}$ coefficients in aggregate and across spatial scales, from both a model that fully accounts for the hierarchical structure as well as a model that does not account for hierarchy. . . . 72

5.1 The WAIC of the fitted model for various k, for both patient groups in the pancreatic cancer dataset. . . . 97
5.2 The WAIC of the fitted model for various k, for both patient groups in the CRC dataset. . . . 100

B.1 Model fitting times on images from pancreatic cancer patient groups for various k. 164
B.2 Model fitting times on images from colorectal cancer patient groups for various k. 171

C.1 Prior distributions used in the Bayesian survival model. . . . 185
C.2 Simultaneous credible interval coverage across dataset sizes. Median 95% simultaneous credible interval coverage across 100 simulations for both the spatial effect $\beta(s)$ and the patient-level covariate $X_i(s)$. . . . 185

LIST OF APPENDICES

A Supplementary Material for The SHADE Model for Asymmetric Multi-level Spatial Interactions . . . 136

Supplementary Figures . . . 136
Priors for Variance Parameters . . . 144
Simulation Study Details . . . 144
Quantifying Variability in Spatial Interactions . . . 145

B Supplementary Material for The BIPPS Model for Shared Spatial Dependencies Across Patients . . . 147

Simulation Details . . . 147
Supplemental Results . . . 148
Markov-chain Monte Carlo for posterior sampling . . . 179
Implementation details . . . 179

C Supplementary Material for Bayesian Survival Modeling with Functional Spatial Covariates . . . 181

Figures . . . 181
Tables . . . 185
Methods . . . 186

LIST OF ACRONYMS

APC antigen-presenting cell

AUC area under the curve

BIPPS Bayesian Inference for Point Processes in Space

CAF cancer-associated fibroblast

CI confidence/credible interval

CLR Crohn's-like reaction

CODEX CO-Detection by Indexing

CRC colorectal cancer

CTL cytotoxic T lymphocyte

CyCIF cyclic immunofluorescence

CyTOF cytometry by Time-of-Flight

DAG directed acyclic graph

DFS disease-free survival

DII diffuse inflammatory infiltration

ECDF empirical cumulative distribution function

ECM extracellular matrix

E/M epithelial/mesenchymal

EMT epithelial-to-mesenchymal transition

ESS effective sample size

FDR false discovery rate

FFPE formalin-fixed paraffin-embedded

GP Gaussian process

HMC Hamiltonian Monte Carlo

HR hazard ratio

ICI immune checkpoint inhibitor

IHC immunohistochemistry

IMC imaging mass cytometry

IQR interquartile range

IPMN intraductal papillary mucinous neoplasm

JSDM joint species distribution model

LGCP log-Gaussian Cox process

MAD median absolute deviation

MCMC Markov chain Monte Carlo

MGP meshed Gaussian process

MGPP multitype Gibbs point process

MIBI multiplexed ion beam imaging

mIF multiplex immunofluorescence

NK cell natural killer cell

ODE ordinary differential equation

OS overall survival

PDAC pancreatic ductal adenocarcinoma

QSP quantitative systems pharmacology

RFS recurrence-free survival

RMSE root mean square error

SBC simulation-based calibration

SBI simulation-based inference

SHADE Spatial Hierarchical Asymmetry via Directional Estimation

SIC spatial interaction curve

TAM tumor-associated macrophage

TAN tumor-associated neutrophil

Th cell helper T cell

TMA tissue microarray

TIME tumor immune microenvironment

TLS tertiary lymphoid structure

TME tumor microenvironment

Treg regulatory T cell

t-SNE t-distributed stochastic neighbor embedding

WAIC Watanabe-Akaike information criterion

ABSTRACT

The tumor microenvironment (TME) is a complex and spatially structured system of immune, stromal, and tumor cells whose interactions play a central role in cancer progression and treatment response. Advances in multiplexed imaging have enabled detailed spatial profiling of the TME at single-cell resolution, but extracting interpretable, reproducible insights from these data remains a major challenge. This dissertation develops and applies a set of spatial statistical models for analyzing spatial tissue data in a way that reflects biological complexity while remaining statistically coherent and computationally tractable.

The first part of the dissertation focuses on applied spatial analysis, characterizing immune architecture in metastatic melanoma and colorectal cancer using both descriptive and model-based approaches. These studies reveal spatial patterns associated with disease stage, immune suppression, and patient survival. The second part introduces new multilevel spatial models, including SHADE, a hierarchical extension of the Gibbs point process that captures directional cell–cell interactions across multiple spatial and biological scales, and BIPPS, a latent spatial factor model that estimates shared patterns of cellular organization across subjects. Chapter 6 develops a Bayesian survival model that integrates functional spatial covariates while accounting for sampling uncertainty and hierarchical data structure. Finally, Chapter 7 explores simulation-based inference in mechanistic models of tumor–immune dynamics, demonstrating the feasibility of applying amortized neural inference methods to recover parameters from dynamic biological systems.

Together, these contributions offer a flexible statistical framework for studying spatial organization in cancer tissues and connecting it to clinical outcomes. By combining spatial statistics, hierarchical modeling, and Bayesian inference, the dissertation provides tools for more principled and interpretable analysis of high-dimensional tissue imaging data.

CHAPTER 1

Introduction

1.1 Motivation and Background

The spatial organization of cells within the tumor microenvironment (TME) is a key determinant of cancer progression and treatment response (Bożyk et al., 2022; N. M. Anderson and Simon, 2020). Patterns of immune infiltration, stromal arrangement, and tumor-immune interactions are shaped not only by cell type composition but also by the spatial proximity and contextual relationships among cells (Y. Yuan, 2016; Fu et al., 2021). These spatial features can inform our understanding of tumor architecture, immune evasion, and prognostic variation across patients, yet they are not easily captured by bulk assays or conventional transcriptomic profiling.

Recent advances in multiplexed tissue imaging technologies—such as imaging mass cytometry (IMC) and multiplex immunofluorescence (mIF)—enable single-cell resolution measurement of dozens of protein markers across entire tumor sections (Sheng et al., 2023; Bodenmiller, 2016; Goltsev et al., 2018; Dam, Baars, and Vercoulen, 2022). These platforms have opened the door to detailed characterization of the TME across different cancer types and patient subgroups. With these data, researchers can now quantify immune cell distributions, identify recurring architectural features, and study spatial interactions in situ, rather than relying on surrogate measures or in vitro approximations.

Despite this progress, current analytic approaches remain limited in several important respects. Many spatial analyses are conducted at the level of individual tissue images, treating each sample independently and failing to account for the hierarchical structure of imaging studies, where multiple images are nested within patients and patients may be stratified into clinically distinct cohorts. Descriptive spatial statistics—such as cell-type proximity counts or co-occurrence indices—are commonly used to summarize spatial patterns in tissue (Palla et al., 2021; Schapiro, Jackson, et al., 2017; Summers, Wills, and Rees, 2022). While they can highlight local structure and are computationally accessible, they

often lack a formal probabilistic interpretation, making it difficult to assess uncertainty, perform model comparison, or directly relate observed patterns to underlying biological processes. In addition, standard pipelines often separate the spatial analysis stage from outcome modeling, making it difficult to assess uncertainty or attribute prognostic relevance to specific spatial features.

There is a need for statistical models that can rigorously quantify spatial interactions in tissue, account for the multilevel structure of imaging data, and link these spatial patterns to clinical outcomes. Importantly, such models should allow for uncertainty quantification and propagation, support inference across multiple spatial scales, and provide interpretable outputs that reflect biologically meaningful variation.

This dissertation develops a set of Bayesian hierarchical models and spatial statistical tools to meet these needs. The work begins by characterizing spatial immune architecture across disease stages in melanoma using descriptive spatial statistics. It then applies an existing class of multitype Gibbs point process models to quantify local cell–cell interactions in multiplexed tissue images from colorectal cancer patients and assess their association with patient outcomes. Building on this foundation, the dissertation introduces a novel multilevel spatial point process model to estimate asymmetric spatial interactions between cell types, accounting for image-, patient-, and cohort-level variation. In parallel, it introduces a Bayesian latent spatial factor model to capture shared patterns of cell–cell interaction across patients in large imaging datasets. A subsequent model incorporates spatial functional covariates, such as the G-cross function ubiquitous in spatial point pattern analysis, into survival analysis while accounting for measurement uncertainty and multilevel dataset structure. Together, these contributions provide a coherent framework for modeling spatial organization in the TME, accounting for the multilevel structure of imaging data, and for linking spatial features to clinical and biological variation. Finally, the dissertation concludes with a pivot toward simulation-based Bayesian inference applied to mechanistic models of tumor–immune dynamics, with the broader aim of rigorously characterizing immune interactions in cancer in the temporal domain.

1.2 Gaps in Existing Approaches

Although the availability of high-dimensional spatial imaging data has expanded rapidly, the analytical tools used to study tumor architecture have not kept pace. Many current approaches rely on descriptive spatial statistics—such as co-localization scores, neighborhood clustering, or proximity-based enrichment metrics—to summarize cellular organization in tissue (Bull et al., 2020; Schürch et al., 2020; Carstens, S. N. Krishnan, et al., 2024). While

these methods offer intuitive visualizations and are relatively easy to compute, they lack a probabilistic foundation. As a result, they do not yield likelihood-based estimates, making it difficult to evaluate model fit, quantify uncertainty, produce predictions or perform principled hypothesis testing. Moreover, such summary statistics may be sensitive to arbitrary thresholding and do not always have clear biological interpretations in terms of the strength or nature of spatial interactions between cell types.

A second limitation is the widespread assumption of independence across images and patients. In many multiplexed imaging studies, multiple images are collected from the same patient, and patients are often grouped into biologically or clinically meaningful cohorts. Ignoring this nested structure can lead to biased or otherwise miscalibrated inferences, as it fails to account for shared variation at the patient or cohort level. In particular, spatial features observed in a single image may not generalize across patients, and averaging across images without modeling this structure may obscure meaningful heterogeneity.

Third, spatial analysis and downstream clinical modeling are often treated as separate steps, with spatial features used as fixed inputs in survival or outcome models. In many cases, the uncertainty in these estimates is not explicitly accounted for, which can affect the interpretability and reliability of downstream predictions. This separation also makes it more difficult to assess how robust or generalizable spatial biomarkers may be in different settings.

To address these challenges, there is a need for statistical models that are not only spatially informed but also capable of accounting for hierarchical structure and uncertainty. Ideally, such models would support probabilistic inference and allow spatial dependencies to be linked to clinical outcomes in a coherent and interpretable way.

1.3 Dissertation Aims and Research Questions

The goal of this dissertation is to develop statistical models for analyzing spatial patterns in the TME using multiplexed tissue imaging data. The models aim to better account for multilevel data structure—images nested within patients, patients grouped into cohorts—and to provide interpretable summaries of spatial organization that can be related to patient-level outcomes. Throughout, the focus is on combining flexibility with statistical rigor: capturing spatial interactions in a way that supports inference, quantifies uncertainty, and respects the structure of the data.

The dissertation is organized around three main aims:

1. **Characterize spatial immune architecture and its association with clinical outcomes.**

This aim uses descriptive spatial statistics and an established multitype Gibbs point process model to investigate immune cell composition and organization across disease stages and tumor types. By applying these methods to melanoma and colorectal cancer datasets, the goal is to identify spatial features associated with progression and survival, and to evaluate the strengths and limitations of existing approaches for modeling immune architecture in the TME.

2. **Develop statistical models for spatial interaction that account for multilevel structure.**
   This aim introduces two complementary frameworks for analyzing cell–cell interactions. The first is a multilevel Gibbs point process model that estimates directional, asymmetric interactions between cell types, incorporating variation across images, patients, and cohorts. The second is a Bayesian latent spatial factor model that captures shared spatial intensity patterns among multiple cell types by treating the point pattern as a multivariate outcome. Conceptually similar to a multilevel multivariate log-Gaussian Cox process, the model accounts for both global structure and patient-level heterogeneity, offering an alternative approach to modeling spatial dependence in high-dimensional tissue imaging data.

3. **Integrate spatial covariates into survival modeling while accounting for uncertainty.**
   The third aim focuses on building a Bayesian survival model that incorporates functional covariates derived from spatial data, such as the G-cross function. The model includes an explicit representation of measurement error, allowing uncertainty in the spatial summaries to be propagated through to survival estimates.

This work is motivated by the realities of spatial tissue data: heterogeneous, structured, and difficult to summarize with conventional tools. The models developed here aim to meet those challenges with statistical approaches that reflect both complexity and biological meaning.

1.4 Overview of Contributions

This dissertation is organized into six main research chapters, each addressing a different aspect of modeling spatial interactions in the tumor microenvironment. Together, these chapters contribute a set of methods and applications for understanding immune architecture in cancer using high-dimensional spatial imaging data.

Chapter 2: Spatial Analysis of Immune Neighborhoods in Melanoma This chapter presents a spatial analysis of immune cell organization across disease stages in metastatic melanoma. Using imaging mass cytometry and multiplex immunofluorescence data from tissue microarrays, the analysis quantifies immune composition, defines inflammatory neighborhoods, and uses G-cross function estimates to assess spatial interactions between key immune populations. Associations with patient survival are examined using Cox proportional hazards models. The findings highlight stage-specific shifts in immune structure and identify spatial biomarkers associated with patient survival.

Chapter 3: Multitype Gibbs Modeling of Cell–Cell Interactions and Survival Association This chapter applies an existing multitype Gibbs point process (MGPP) model to multiplexed colorectal cancer data. This model is used to estimate local cell–cell interactions and assess their relationship with clinical covariates and outcomes. Framed as a model derived from the ecological literature, the *joint species distribution model*, the MGPP approach enables formal uncertainty quantification and prediction while remaining interpretable in ecological terms. This chapter serves as a case study in applying interpretable and probabilistic spatial models to characterize cell–cell interactions and relate them to clinical outcomes.

Chapter 4: The SHADE Model for Asymmetric Multilevel Spatial Interactions This chapter develops the SHADE model (Spatial Hierarchical Asymmetry via Directional Estimation), a multilevel extension of Gibbs process models that assumes spatial interactions between cell types are directional and asymmetric. The model also accounts for variation across images, patients, and cohorts through a hierarchical structure. The model is implemented in a Bayesian hierarchical framework and fitted using a logistic regression approximation for computational efficiency. Simulation studies and application to colorectal cancer data demonstrate how SHADE captures multiscale spatial structure and variation in interaction patterns across patient groups.

Chapter 5: The BIPPS Model for Shared Spatial Dependencies Across Patients This chapter presents BIPPS (Bayesian Inference for Point Processes in Space), a hierarchical model for analyzing multitype spatial point patterns across multiple subjects. BIPPS models cell locations as a multivariate log-Gaussian Cox process, with latent Gaussian processes capturing spatial intensity functions for each cell type. By pooling information across patients, the model estimates within- and cross-type spatial correlations that are shared across patients, offering interpretable summaries of tissue organization. Applications to pancre-

atic and colorectal cancer datasets reveal distinct spatial structures across disease subtypes and highlight tumor–immune features linked to immune-permissive and immune-excluded microenvironments.

Chapter 6: Bayesian Survival Modeling with Functional Spatial Covariates This chapter introduces a Bayesian hierarchical survival model that incorporates spatial summary functions as functional covariates. Rather than extracting values at a few predefined spatial radii, the model includes the entire interaction curve—such as a G-cross function—as a covariate with a spatially varying coefficient. To account for sampling variability, the model uses bootstrap-based estimates of measurement error, allowing uncertainty in the covariates to be propagated through to survival predictions. The framework also supports repeated measurements per patient by modeling latent patient-level covariates within a multilevel structure. Simulation studies and application to breast cancer imaging data demonstrate the model's ability to recover spatial effects and improve interpretability in survival analysis.

Chapter 7: Case Studies in Simulation-Based Inference for Mechanistic Tumor–Immune Modeling Lastly, this chapter explores simulation-based inference (SBI) as a flexible, likelihood-free approach to parameter estimation in mechanistic models of tumor–immune dynamics. Unlike the spatial models in earlier chapters, this work focuses on systems of ordinary differential equations, where traditional likelihood-based inference is often intractable. Two case studies are presented: one using a simplified tumor–immune model to compare amortized SBI with Hamiltonian Monte Carlo for Bayesian inference, and another applying SBI to a quantitative systems pharmacology model of immune checkpoint inhibition. Both studies assess the ability of neural SBI methods to recover biologically meaningful parameters from simulated time-series data, using calibration diagnostics to evaluate accuracy and uncertainty. This work serves as a proof of concept for integrating mechanistic modeling with simulation-based inference in systems oncology.

1.4.1 Overview of a Typical Multiplex Imaging Data Structure

The multiplexed imaging datasets analyzed in this dissertation are hierarchically structured, reflecting the biological and experimental design. Each cohort represents a group of patients with shared clinical or biological characteristics. Within each cohort, individual subjects contribute one or more tumor tissue sections, typically derived from core biopsy samples. These tissue sections are analyzed using high-parameter multiplexed imaging technologies—including Imaging Mass Cytometry (IMC) (Giesen et al., 2014) and CO-Detection

by Indexing (CODEX) (Black et al., 2021)—which capture spatially-resolved measurements of dozens of protein markers at single-cell resolution.

Within each image, thousands of individual cells are identified and segmented based on nuclear or cytoplasmic markers. For each cell, the following features are extracted:

- **Spatial coordinates**: precise (x, y) locations indicating the cell's position within the tissue section.
- **Cell type annotations**: categorical labels assigned based on combinations of marker expression, categorizing cells into biologically relevant populations such as antigen-presenting cells, cytotoxic T lymphocytes, T helper cells, regulatory T cells, epithelial cells, and others.
- **Marker intensities**: continuous measurements reflecting the expression levels of individual markers used to infer cell types and states.

The result of this imaging and segmentation process is a *multitype spatial point pattern*, where each point corresponds to a cell, located at a specific position in the tissue and assigned a type based on marker expression. The hierarchical structure thus encompasses cells within images, images within subjects, and subjects within cohorts. An example of a multitype point pattern derived from multiplexed imaging is provided in Figure 1.1, illustrating the mapping from high-dimensional imaging data to the analytical frameworks applied and developed in this dissertation.

Such a hierarchical structure is typical for multiplexed imaging datasets and provides a natural framework for statistical modeling. It supports the development of models that account for variability at different levels, including within images, between subjects, and across cohorts. The subsequent chapters use this structure to motivate and implement methods for analyzing spatial interactions, tissue organization, and clinical outcomes.

1.5 Organization of the Dissertation

This dissertation reflects both a practical response to challenges in spatial tissue data analysis and a broader interest in how statistical models can help make sense of complex biological systems. The chapters are ordered roughly by level of abstraction: beginning with applied analyses of immune architecture (Chapters 2–3), moving into new hierarchical models for multilevel spatial inference and survival analysis (Chapters 4–6), and concluding with an exploratory application of simulation-based inference to mechanistic tumor–immune models (Chapter 7). While the chapters differ in focus and methodology, they are united by a central aim: to model the tumor microenvironment with statistical rigor and biological relevance.

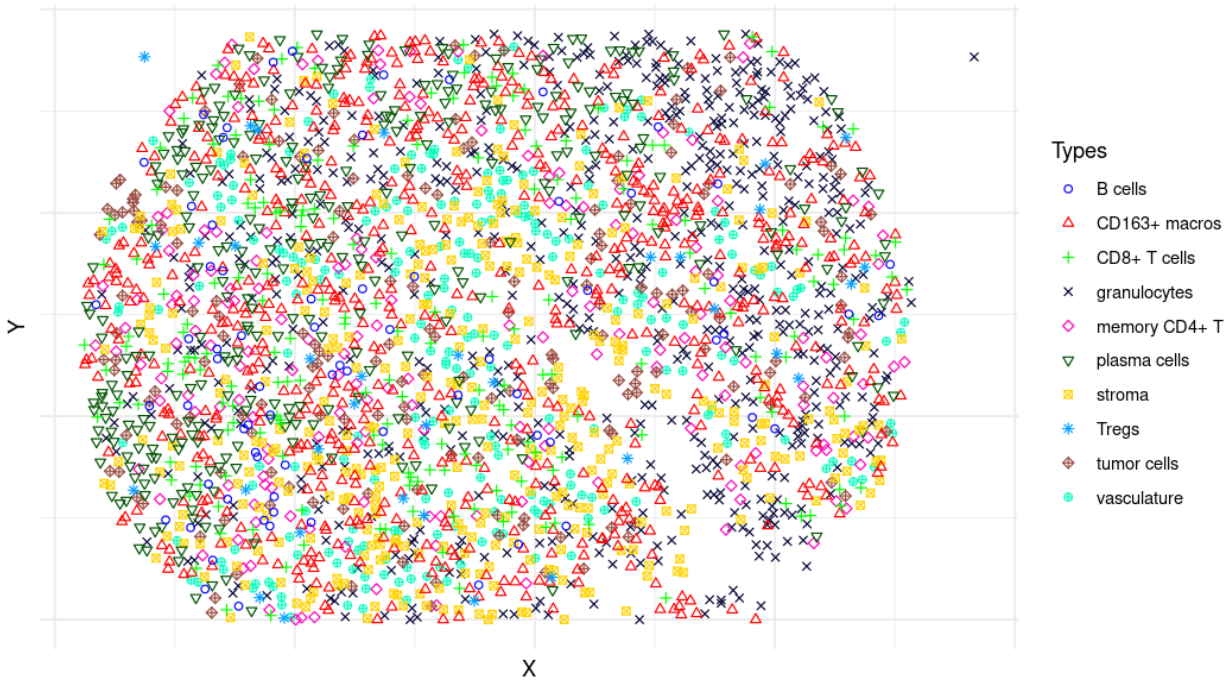


Figure 1.1: A multitype point pattern representation of a multiplex image of a tissue section of the TME in a colorectal cancer patient.

CHAPTER 2

Spatial Analysis of Immune Neighborhoods in Melanoma

2.1 Introduction

Invasive melanoma is the least common but deadliest form of skin cancer. It is estimated that 200,340 new cases of melanoma will be diagnosed in 2024 (American Cancer Society, 2024)[1]. In the United States, melanoma is the fifth most common cancer in both men and women across all age groups and the third most common cancer among individuals aged 20–39 (American Cancer Society, 2021). Over the past decade, there has been a significant increase in the overall five-year survival rate for patients with metastatic melanoma. Between 2014 and 2018, the melanoma mortality rate declined by approximately 7% per year in adults younger than 50 and by about 5% per year in older adults (American Cancer Society, 2022). This improvement is largely attributed to earlier diagnosis, advancements in surgical techniques, and the development of novel therapeutic approaches. In particular, molecularly targeted therapies, vaccines, adoptive T-cell therapy, and immune checkpoint inhibitors (ICIs) have revolutionized melanoma treatment (Knight and Luke, 2024).

Among these, ICIs have had the most significant impact by inducing durable responses and conferring long-term survival benefits through enhanced tumor immunity (Chang et al., 2020; Huang and Zappasodi, 2022; Dummer et al., 2020). However, despite their efficacy, a subset of patients experiences either primary resistance, where they fail to respond initially, or acquired resistance, where tumors progress after an initial response to ICIs (R. W. Jenkins, Barbie, and Flaherty, 2018). These resistance mechanisms are primarily driven by immune evasion strategies that alter tumor-immune interactions, including T cell exhaustion, regulatory T cell activation, recruitment of myeloid-derived suppressor cells and macrophage polarization toward the immunosuppressive M2 phenotype (Lim et al., 2023; Karasarides et

[1]This chapter is based on the publication Eliason, S. Krishnan, et al. (2025)

al., 2022; Chow et al., 2022). To develop more effective therapeutic strategies that enhance immunotherapy outcomes, it is critical to first gain a comprehensive understanding of both the phenotypic and functional spatial organization of the tumor immune microenvironment (TIME).

The complexity and heterogeneity of the TIME pose significant challenges in identifying molecular biomarkers in metastatic melanoma and understanding how the TIME influences responses to ICI therapy. Tumor cells, as major constituents of the TIME, play a pivotal role in shaping the immune landscape by secreting tumor antigens and modulating immune cell function, often creating an unfavorable microenvironment that hinders immune responses. Previous studies by our group have demonstrated that inflammatory marker expression in metastatic melanoma is associated with prognostic indicators such as recurrence-free survival (RFS) and overall survival (OS) Fukuda et al., 2023; Hailemichael et al., 2022; Ding et al., 2021; Ogata et al., 2020; Grimm, Sikora, and Ekmekcioglu, 2013.

The present study aims to provide an extensive analysis of the TIME in advanced melanoma, with a particular focus on the TIME's cellular components and their spatial organization in relation to inflammatory marker expression within the immediate cellular neighborhood.

2.2 Materials and Methods

2.2.1 Patient Samples and Datasets

Two independent tissue microarrays (TMAs) were utilized, representing non-primary metastatic melanoma lesions: Stage III (regional metastases) and Stage IV (distant metastases). Under IRB-approved protocols, cases with available archival material and clinical follow-up exceeding five years were included. Clinical records were reviewed for key events, including local recurrence, distant metastasis patterns, and overall survival (OS).

- **the Stage III TMA**, developed at the Saint John's Cancer Institute (SJCI), was prepared from formalin-fixed paraffin-embedded (FFPE) lymph node tumors. It contained two core samples per patient from 157 Stage III patients, with annotated clinical outcomes and follow-up data.

- **the Stage IV TMA**, also developed at SJCI, comprised 393 core samples from 248 Stage IV patients. These samples, derived from FFPE melanoma tissues obtained from Saint John's Health Center (SJHC), represented a range of distant metastatic sites. Clinical outcomes and follow-up data were sourced from the SJCI melanoma database.

A hematoxylin and eosin (H&E) slide was prepared for each sample, and all specimens were reviewed by a dedicated dermatopathologist to ensure tumor viability. Tumor regions of interest were mapped to FFPE blocks before TMA construction.

To enable simultaneous staining and data acquisition across the cohort, the Stage IV TMA was constructed using 1 mm cylindrical cores, distributed in duplicate and randomized to minimize spatial bias. After accounting for core loss during sectioning, 393 evaluable samples remained, representing metastatic sites such as lung, gastrointestinal tract, lymph nodes, skin, and other distant locations.

Staging for both TMAs followed AJCC guidelines at the time of melanoma diagnosis (Amin et al., 2017). Clinicopathological characteristics are summarized in Tables 2.1 and 2.2. A diagram of our analysis pipeline is shown in Figure 2.1.

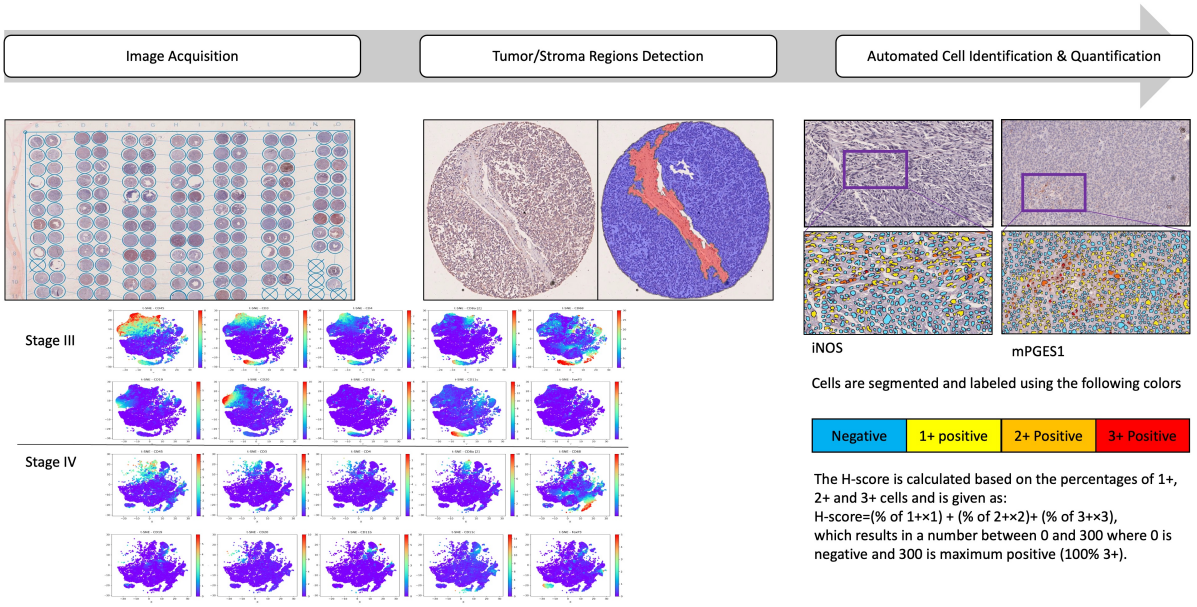


Figure 2.1: Analysis pipeline to characterize the spatial immune landscape of Stage III and IV melanomas. Schematic of the IMC data acquisition of two consecutive slices of two TMAs containing 2 biopsy cores from a total of 157 patients with stage III and 393 samples from 248 patients with stage IV metastatic melanoma. Samples were stained with a protein panel, segmented for tumor/stroma regions detection, followed by cellular identification and quantifications. t-distributed stochastic neighbor embedding (t-SNE) of data was derived from CyTOF of tumor samples, labeled by cell type and the signal intensities of individual markers.

2.2.2 Acquisition and Integration of IMC and mIF Imaging Data

Multiplexed imaging data were acquired using two complementary platforms: imaging mass cytometry (IMC) and multiplex immunofluorescence (mIF). IMC data were generated using the Hyperion Imaging System (Fluidigm) from TMAs, employing a panel of 35 metal-tagged antibodies targeting phenotypic, functional, and inflammatory markers relevant to melanoma. Tissue sections were ablated at approximately 1 μm resolution, and the resulting ion signals were reconstructed into multi-channel image stacks. Poor-quality cores were excluded following visual inspection. Cells were segmented based on nuclear signals and assigned phenotypes using a combination of marker intensity thresholds and H-score calculations. High-dimensional data were visualized using t-distributed stochastic neighbor embedding (t-SNE), enabling exploratory comparisons of immune composition across disease stages.

In parallel, mIF staining was performed on adjacent TMA sections using the Opal 7-Color IHC system (Akoya Biosciences). Slides were stained for inflammatory and structural markers and imaged using the Vectra Polaris system. Following staining, mIF and IMC images were co-registered using the TissueAlign pipeline in Visiopharm. Cell boundaries were approximated by nuclear expansion, and phenotypes were defined using intensity-based classification. Cellular neighborhoods were defined as spatial regions extending 20–80 μm around each target cell type and transferred across imaging modalities. These neighborhood regions of interest served as the foundation for downstream spatial analysis, enabling joint interrogation of phenotypic and spatial features within the tumor microenvironment.

2.2.3 Statistical Analysis

To compare the proportions of cell types and inflammatory neighborhoods between Stage III and Stage IV melanoma, we used beta regression models with false discovery rate (FDR) correction for multiple comparisons. Beta regression was chosen as it appropriately models proportion data constrained between 0 and 1. Groups with zero variance were excluded from analysis to ensure valid model estimation. For each comparison, the beta regression model included stage as the independent variable and the proportion of the respective feature as the dependent variable. Beta regression models were fit using the `betareg` R package (Cribari-Neto and Zeileis, 2010). P-values were adjusted using the Benjamini-Hochberg method (Benjamini and Hochberg, 1995), and statistical significance was defined as $p < 0.05$. Boxplots and heatmaps were used to visualize the distributions of proportions, with significance markers displayed where appropriate. Cox proportional hazards models were fitted using the R package `survival` (Therneau et al., 2024), and Kaplan-Meier plots were generated using the

R package `survminer` (Kassambara et al., 2024). To account for intra-patient correlation when patients contributed more than one sample, as in the Stage IV TMA, Cox regression models were fitted with robust standard errors using the cluster argument in the `coxph` function . This approach adjusts variance estimates to account for the non-independence of observations within the same patient. Other plots were created using the `tidyverse` packages in R (Wickham et al., 2019). All statistical analyses performed on the Stage IV TMA were performed on samples from all organs pooled together. All samples from the Stage III and Stage IV TMA were used for all statistical analyses.

2.2.3.1 Spatial Analysis

For the analysis of spatial dispersion and clustering between different pairs of cell types, we estimated the G-cross nearest neighbor distance function $G_{i,j}(r)$, which is a function of the distance r, with i and j indicating the two types of cells (Barua et al., 2018; Carstens, Correa de Sampaio, et al., 2017). Briefly, $G_{i,j}(r)$ represents the cumulative distribution function of the distance from a typical point of type i to the nearest point of type j. Mathematically, this can be expressed as:

$$G_{i,j}(r) = 1 - \exp(-\alpha_j \pi r^2) \tag{2.1}$$

where the subscripts i and j indicate that the spatial distribution of cell type j relative to cell type i is being computed, r refers to the distance from the reference cell type, and α_j is the overall density of cell type j on the slide.

We estimated $G_{i,j}(r)$ using the `Gcross` function in the R package `spatstat` (A. Baddeley, Rubak, and Turner, 2015). To assess the clinical significance of the estimated $G_{i,j}(r)$ across Stages III and IV, we extracted the $G_{i,j}(r)$ estimates at various radii for each cell type pair and used them as independent variables in a univariate Cox proportional hazards model. These models were fitted using the `coxph` function in the R package `survival` (Therneau et al., 2024).

2.3 Results

2.3.1 Summary of Stage III Patient Characteristics

The Stage III TMA cohort included 157 patients with a median age of 52 years (IQR: 24 years) (Table 2.1). The cohort was 59% male and 41% female. Median Breslow thickness was 2.20 mm (IQR: 2.65 mm), though data were missing for a substantial number of patients

(n = 50). Ulceration was reported in 31% of cases, while 57 patients had unknown ulceration status. The median observation period was 100 months (IQR: 145 months). Most patients (76%) did not receive chemotherapy, while 35% received radiation therapy. Immunotherapy use was variable, with 41% receiving adjuvant treatment, 20% receiving neoadjuvant treatment, and 18% receiving both.

The median overall survival for Stage III TMA patients was 20.2 months (95% CI: 15.3–61.7). The estimated 5-year survival probability was 41.5%, while the 10-year survival probability was 27.9%.

2.3.2 Clinical and Demographic Characteristics of Stage IV TMA Patients

The Stage IV TMA cohort included 248 patients with a median age of 56 years (IQR: 18 years) (Table 2.2). The cohort was 68% male and 32% female. Metastatic sites varied, with 35% of patients having metastases at other locations, followed by skin (20%), gastrointestinal (GI) tract (17%), lung (16%), and lymph nodes (12%). Median Breslow thickness was 1.90 mm (IQR: 1.70 mm), though data were missing for 67 patients. Ulceration was present in 39% of cases, with ulceration status unknown for 95 patients. The median observation period was 56 months (IQR: 80 months), though 44 patients had missing data. The median number of positive lymph nodes was 1.00 (IQR: 1.00), with lymph node status unavailable for 48 patients.

The median overall survival (OS) for Stage IV melanoma patients was 10.2 months (95% CI: 8.6–12.4), with a 5-year survival probability of 6.8% and a 10-year survival probability of 2.6%. Median disease-free survival (DFS) was 18.7 months (95% CI: 15.0–29.2). The 5-year DFS probability was 31.8%, decreasing to 20.2% at 10 years. Kaplan-Meier curves showing the empirical overall survival probabilities for patients from both the Stage III and IV TMA are shown in Figure 2.2.

2.3.3 Differences in Immune Cell Composition between Stages

To characterize changes in the tumor immune microenvironment between Stage III and IV melanoma, we quantified the relative proportions of immune and tumor cells in tumor core biopsies. Beta regression analysis identified significant shifts in cell type distributions (Figure 2.3), where the beta coefficient (β) represents the change in log-odds of a cell type's proportion in Stage IV relative to Stage III.

Stage IV tumors exhibited significantly higher proportions of cytotoxic T lymphocytes (CTL, $\beta = 0.474$, $p < 0.001$) and tumor cells ($\beta = 0.487$, $p < 0.001$), while B cells ($\beta =$

Table 2.1: Summary of key clinical and demographic characteristics for stage III TMA patients.

Stage III Patient Characteristics	N = 157
Age (years)	52 (24)
(Missing)	10
Gender	
Female	61 (41%)
Male	86 (59%)
(Missing)	10
Breslow Thickness (mm)	2.20 (2.65)
(Missing)	50
Has Ulceration	31 (31%)
(Missing)	57
Observation Period (months)	100 (145)
(Missing)	26
Immunotherapy Timing	
No Treatment	32 (22%)
Pre-Surgery (Neoadjuvant)	29 (20%)
Post-Surgery (Adjuvant)	60 (41%)
Both Pre- & Post-Surgery	26 (18%)
(Missing)	10
Chemotherapy Timing	
No Treatment	112 (76%)
Pre-Surgery (Neoadjuvant)	10 (6.8%)
Post-Surgery (Adjuvant)	25 (17%)
Both Pre- & Post-Surgery	0 (0%)
(Missing)	10
Radiation Timing	
No Treatment	95 (65%)
Pre-Surgery (Neoadjuvant)	9 (6.1%)
Post-Surgery (Adjuvant)	40 (27%)
Both Pre- & Post-Surgery	3 (2.0%)
(Missing)	10

[1] Median (IQR); n (%)

-0.365, $p = 0.007$), NK cells ($\beta = -0.541$, $p < 0.0001$), and overall T cell populations ($\beta = -0.188$, $p = 0.044$) were reduced. Differences in tumor-associated macrophages (TAM, $p = 0.109$), M2 macrophages ($p = 0.192$), and T helper (Th) cells ($p = 0.153$) were not statistically significant. These results indicate an immune shift in Stage IV, characterized by increased CTL infiltration and decreased adaptive immune populations.

Table 2.2: Summary of key clinical and demographic characteristics for stage IV TMA patients.

Stage IV Patient Characteristics	N = 248
Age (years)	56 (18)
Gender	
Female	79 (32%)
Male	169 (68%)
Metastatic Site	
GI	42 (17%)
Skin	50 (20%)
Lymph Node	30 (12%)
Lung	39 (16%)
Other Metastatic Site	87 (35%)
Breslow Thickness (mm)	1.90 (1.70)
(Missing)	67
Has Ulceration	60 (39%)
(Missing)	95
Observation Period (months)	56 (80)
(Missing)	44
Lymph Nodes Positive	1.00 (1.00)
(Missing)	48

[1] Median (IQR); n (%)

2.3.4 Shifts in Inflammatory Neighborhoods

To examine immune spatial organization, we analyzed seven inflammatory neighborhoods based on prior studies (Ding et al., 2021; Ogata et al., 2020; Ekmekcioglu, Ellerhorst, et al., 2006; Sikora et al., 2010; Tanese, Hashimoto, et al., 2015; Ekmekcioglu, Davies, et al., 2016). Beta regression analysis identified significant differences in their prevalence between Stage III and IV melanoma (Figure 2.4).

Stage IV tumors had a higher prevalence of CD74 CD44 ($\beta = 0.0841$, $p = 0.015$), CD74 MIF ($\beta = 0.306$, $p < 0.0001$), and MIF CD44 neighborhoods ($\beta = 0.320$, $p < 0.0001$), while iNOS mPGES ($\beta = -0.772$, $p < 0.0001$), iNOS mPGES-NT ($\beta = -0.773$, $p < 0.0001$), and iNOS NT neighborhoods ($\beta = -0.455$, $p < 0.0001$) were significantly less frequent. A modest but significant reduction was also observed for CD74 CD44 MIF ($\beta = -0.121$, $p = 0.015$). These findings indicate a shift toward CD74- and MIF-enriched neighborhoods in Stage IV, accompanied by a loss of iNOS-associated regions.

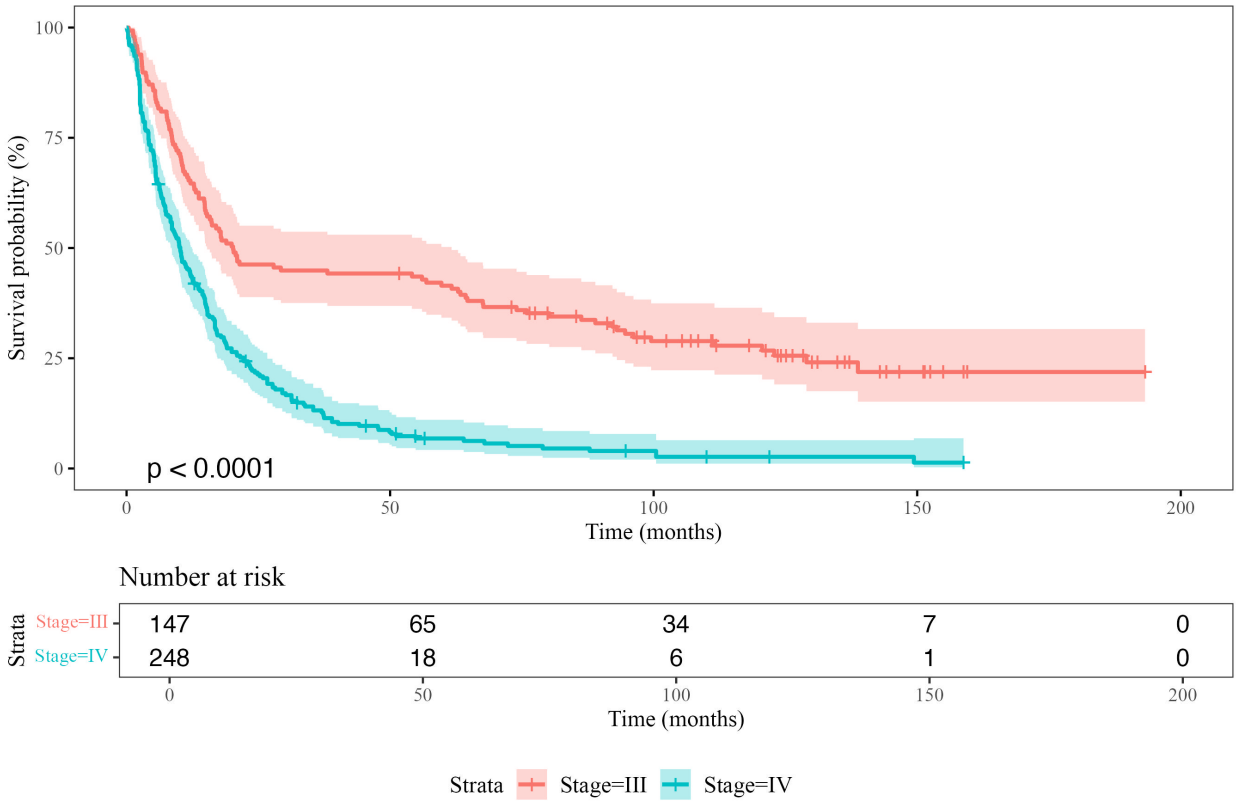


Figure 2.2: Kaplan-Meier survival curves comparing overall survival (OS) probability between Stage III and Stage IV patients. The shaded regions represent 95% confidence intervals. The p-value indicates the statistical significance of the difference between the survival distributions (log-rank test). The risk table below the plot shows the number of patients at risk at each time point.

2.3.5 Immune Cell Distribution Within Inflammatory Neighborhoods

We next examined the distribution of immune cells as a fraction of all cells in each inflammatory neighborhood in Stage III and IV melanoma (Figure 2.5). CTLs and tumor cells were significantly more abundant in all neighborhoods in Stage IV, suggesting a broader restructuring of the tumor microenvironment.

B cells and NK cells were generally reduced in Stage IV, particularly as a proportion of CD74- and MIF-enriched neighborhoods. B cells were significantly lower in CD74 CD44 ($\beta = -0.393$, $p = 0.014$) and CD74 MIF ($\beta = -0.426$, $p = 0.012$), while NK cells were depleted in CD74 CD44 ($\beta = -0.367$, $p = 0.0018$) and MIF CD44 ($\beta = -0.463$, $p < 0.0001$). T cells were also significantly reduced as a fraction of multiple neighborhoods, including

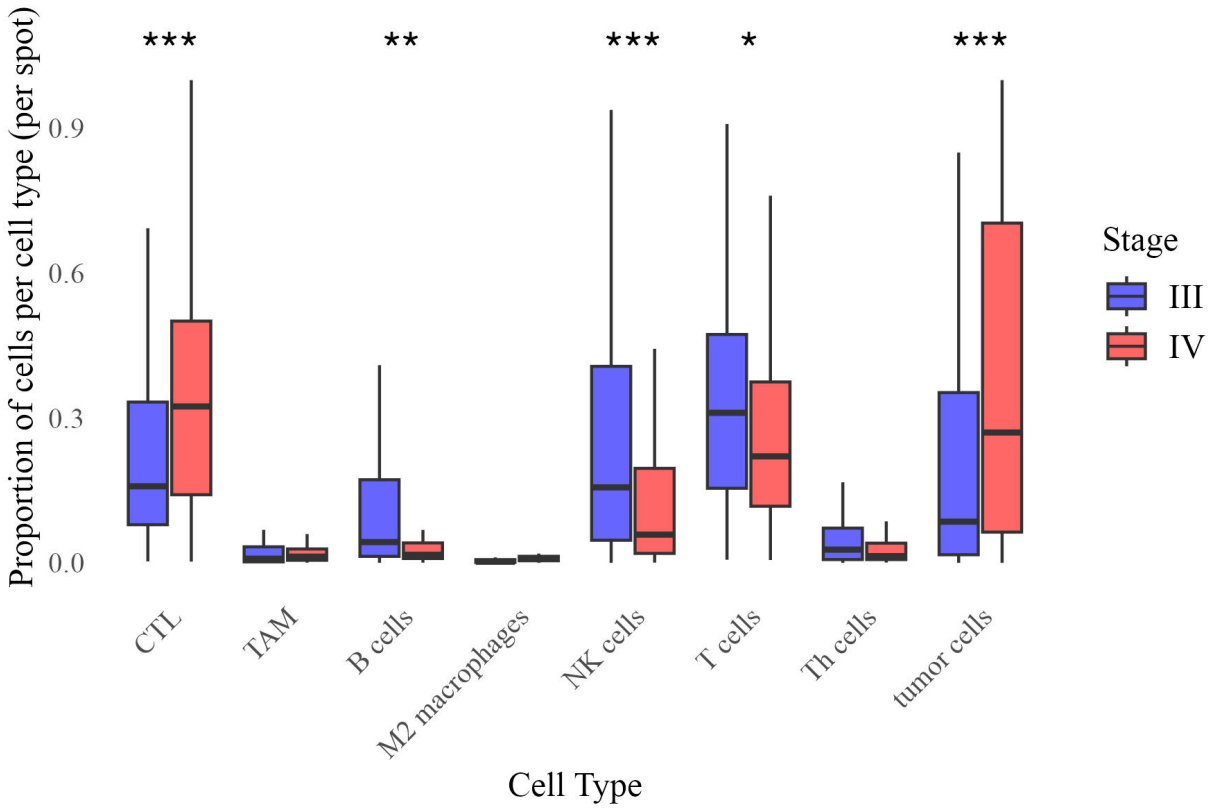


Figure 2.3: Distribution of cell type proportions across Stage III and Stage IV melanoma. Each boxplot represents the proportion of a given cell type per image, grouped by stage. Statistical significance of difference in mean proportion was assessed using a beta regression model, with significance markers (*, **, ***) indicating significant differences between stages after false discovery rate correction. Significance levels are denoted as follows: $***(p < 0.001)$, $**$ $(p < 0.01)$, and $*(p < 0.05)$.

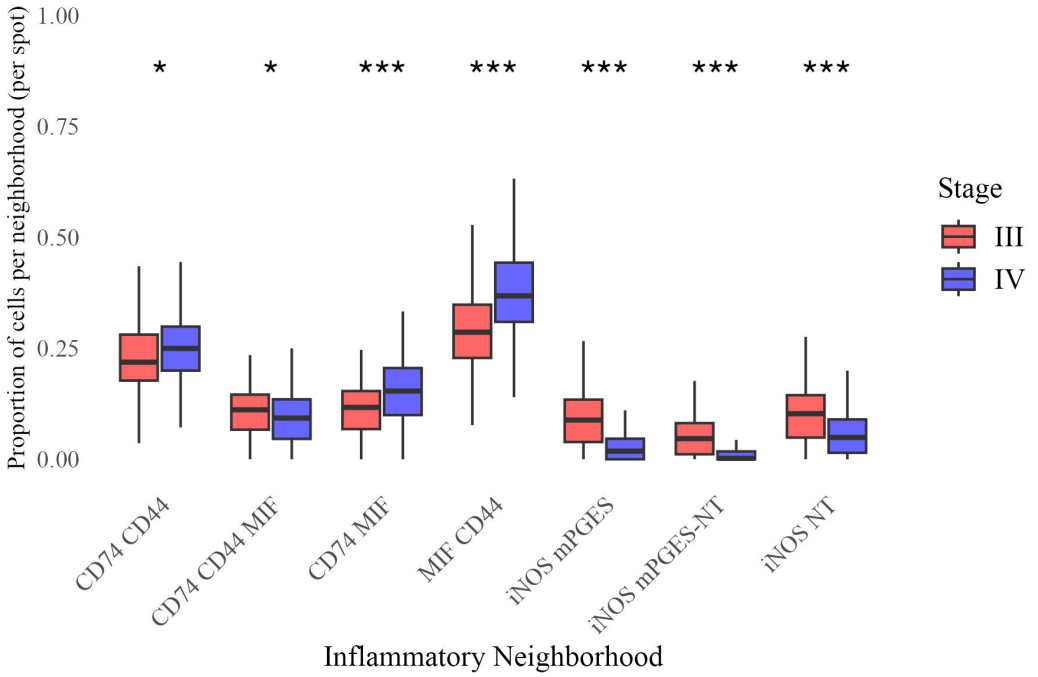


Figure 2.4: Distribution of neighborhood proportions across Stage III and Stage IV melanoma. Each boxplot represents the proportion of a given inflammatory neighborhood per image, grouped by stage. Statistical significance of difference in mean proportion was assessed using a beta regression model, with significance markers (*, **, ***) indicating significant differences between stages after false discovery rate correction. Significance levels are denoted as follows: $***(p < 0.001)$, $**$ $(p < 0.01)$, and $*(p < 0.05)$.

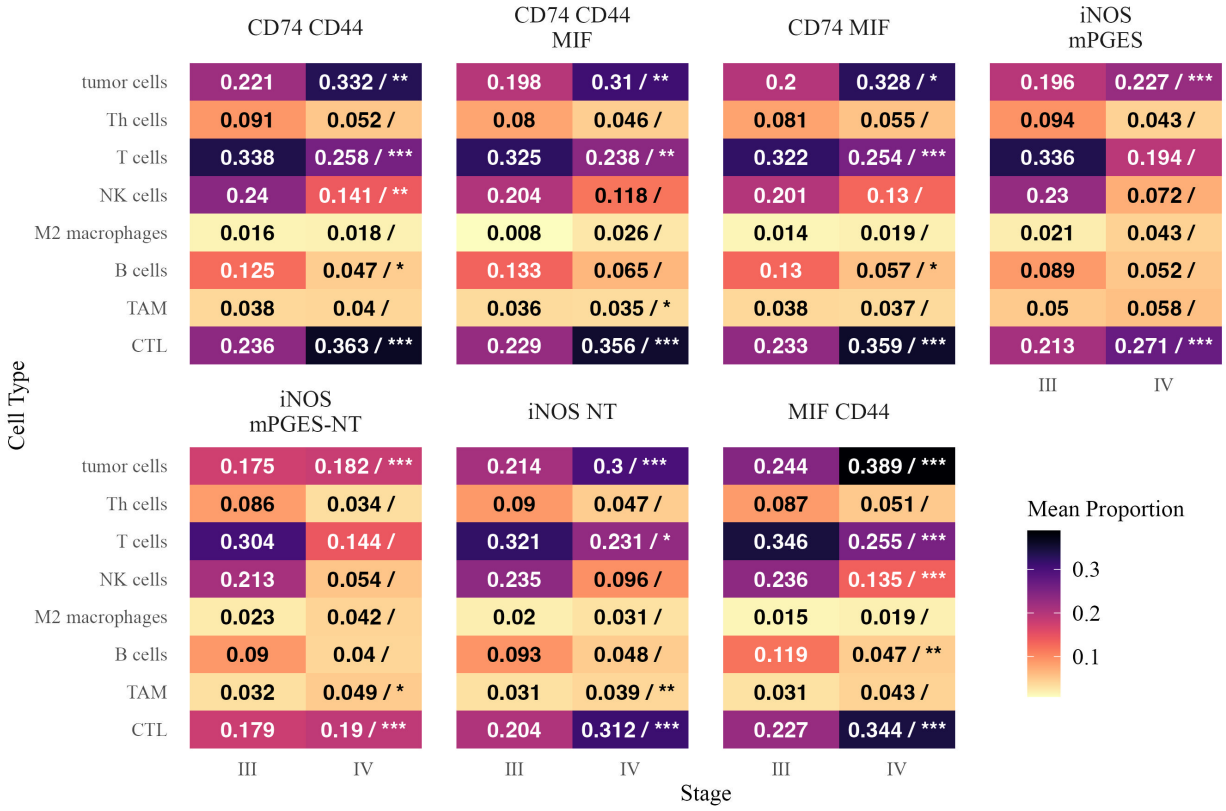


Figure 2.5: Mean proportions of different cell types within inflammatory neighborhoods in Stage III and Stage IV melanoma. Each tile is annotated with the mean proportion of the corresponding cell type among cells in that neighborhood. In the Stage IV column, a forward slash ("/") separates the mean proportion from statistical significance markers (*, **, ***), indicating significant differences in mean proportion between stages based on a beta regression model with false discovery rate correction. Significance levels are denoted as follows: $***(p < 0.001)$, $**$ $(p < 0.01)$, and $*(p < 0.05)$. Text color is adjusted for readability.

CD74 MIF ($\beta = -0.292$, $p < 0.0001$) and MIF CD44 ($\beta = -0.359$, $p < 0.0001$), despite the overall increase in CTLs. M2 macrophages and Th cells did not show significant enrichment or depletion in any neighborhood.

2.3.6 Neighborhood Representation Within Immune Cell Populations

To assess how immune and tumor cells partition across neighborhoods, we analyzed the fraction of each cell type residing in different inflammatory neighborhoods (Figure 2.6). A positive estimate indicates a greater fraction of the overall population of a given cell type was located within a specific neighborhood in Stage IV.

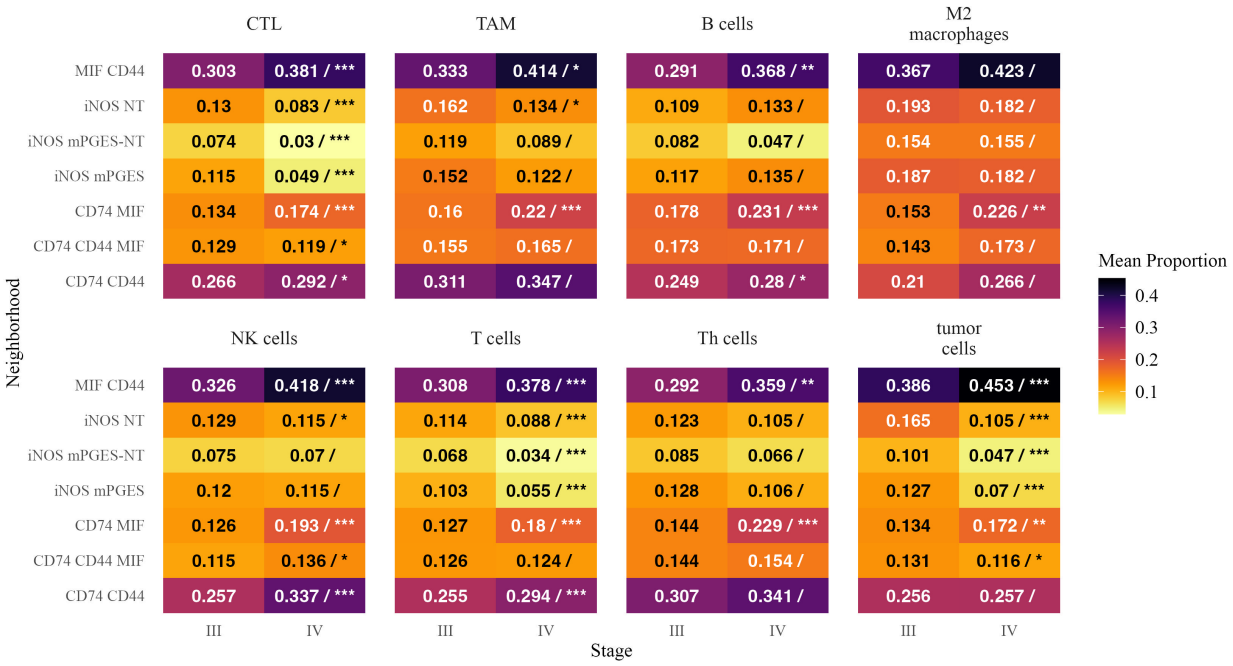


Figure 2.6: Mean proportions of inflammatory neighborhoods as fractions of cells within each cell type in Stage III and Stage IV melanoma. Each tile is annotated with the mean proportion of the corresponding neighborhood among cells within that cell type. In the Stage IV column, a forward slash ("/") separates the mean proportion from statistical significance markers (*, **, ***), indicating significant differences in mean proportion between stages based on a beta regression model with false discovery rate correction. Significance levels are denoted as follows: $***$ ($p < 0.001$), $**$ ($p < 0.01$), and $*$ ($p < 0.05$). Text color is adjusted for readability.

CTLs were more frequently found in CD74 MIF ($\beta = 0.286$, $p < 0.0001$) and MIF CD44 ($\beta = 0.241$, $p < 0.0001$), suggesting an increased presence of cytotoxic immune responses

in these neighborhoods. B cells were more localized within CD74 MIF ($\beta = 0.249$, $p < 0.0001$) and MIF CD44 ($\beta = 0.231$, $p = 0.0016$). Tumor cells followed a similar trend, increasing in CD74 MIF ($\beta = 0.219$, $p = 0.005$) and MIF CD44 ($\beta = 0.190$, $p < 0.0001$), further supporting their role as hubs of immune activation and tumor presence in advanced disease (Tanese and Ogata, 2024).

Conversely, iNOS- and mPGES-enriched neighborhoods accounted for a smaller fraction of CTLs and T cells in Stage IV. The proportion of CTLs was significantly lower in iNOS mPGES ($\beta = -0.804$, $p < 0.0001$) and iNOS mPGES-NT ($\beta = -0.750$, $p < 0.0001$). T cells showed a similar reduction in iNOS mPGES ($\beta = -0.684$, $p < 0.0001$) and iNOS mPGES-NT ($\beta = -0.642$, $p < 0.0001$), suggesting a diminished adaptive immune presence in these regions.

MIF CD44 neighborhoods encompassed a greater fraction of NK cells in Stage IV ($\beta = 0.251$, $p < 0.0001$), while iNOS NT neighborhoods accounted for a smaller proportion of NK cells ($\beta = -0.266$, $p = 0.016$). Th cells were more frequently found in MIF CD44 ($\beta = 0.234$, $p = 0.0035$) and CD74 MIF ($\beta = 0.424$, $p < 0.0001$). M2 macrophages in CD74 MIF made up a larger proportion of their overall cell populations in Stage IV than in Stage III ($\beta = 0.360$, $p < 0.0001$; $\beta = 0.478$, $p = 0.0037$).

These findings suggest that in Stage IV melanoma, immune and tumor cells preferentially localize within CD74- and MIF-enriched neighborhoods, while iNOS-associated neighborhoods see a reduced presence of CTLs, T cells, and tumor cells. This redistribution may reflect altered immune engagement and inflammatory signaling in advanced disease.

2.3.7 Spatial Immune Interactions and Their Association with Patient Survival

2.3.7.1 Spatial Interactions in Stage III Patients

We examined spatial immune interactions at 40 μm (as measured by the G-cross function) in Stage III melanoma and their association with survival (Figure 2.7). Several interactions were linked to improved outcomes, suggesting that coordinated adaptive and innate immune responses enhance tumor control (Piatakova et al., 2021). These included CTLs attracting M2 macrophages (log-HR = -6.20, $p = 0.022$) and M2 macrophages attracting B cells (log-HR = -2.68, $p = 0.023$), potentially reflecting M1 polarization that promotes anti-tumor immunity (Su et al., 2024).

Conversely, interactions linked to worse survival included tumor cells attracting TAMs (log-HR = 1.15, $p = 0.036$), suggesting that macrophage clustering contributes to an immunosuppressive environment. Additionally, CTLs (log-HR = 1.41, $p = 0.0077$), NK cells

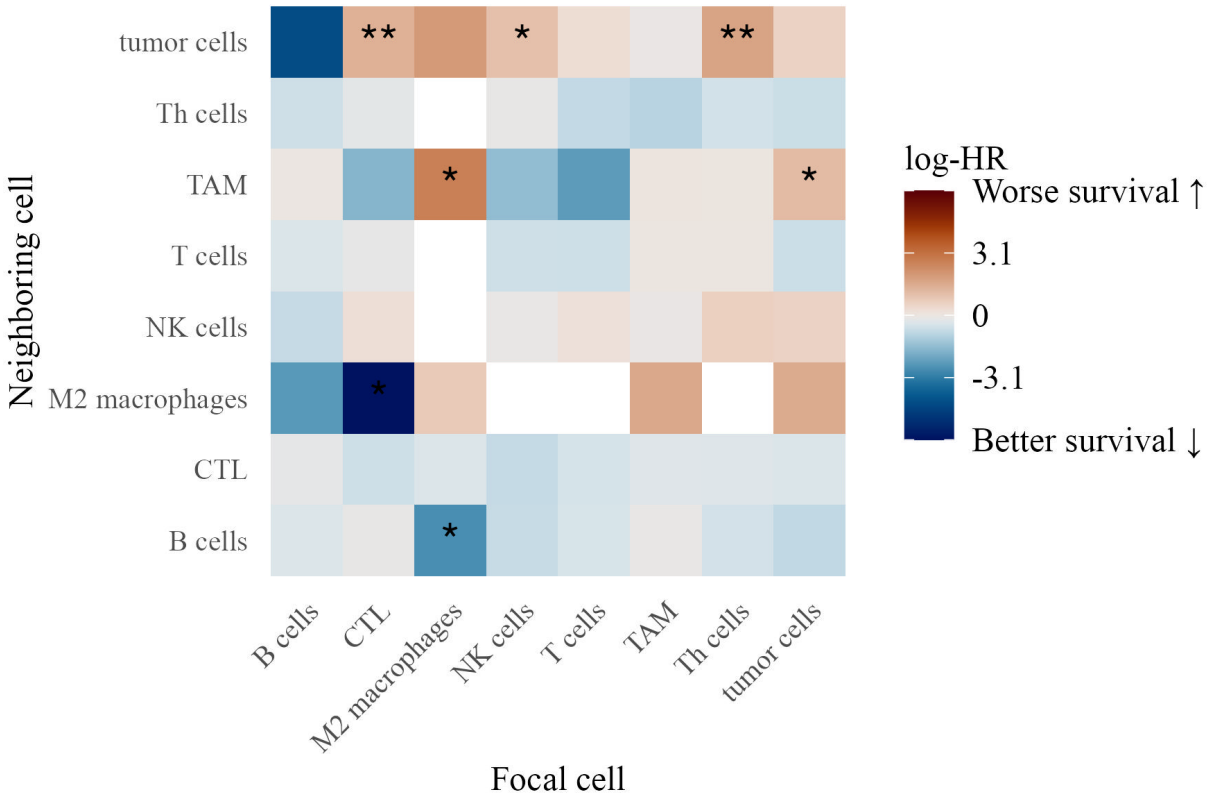


Figure 2.7: Heatmap of log-hazard ratios estimated using Cox regression, where the G-cross function at 40 μm between each pair of cell types in Stage III patients serves as a predictor of survival time. Significance levels of hazard ratios are denoted as follows: $***$ ($p < 0.001$), $**$ ($p < 0.01$), and $*$ ($p < 0.05$).

(log-HR = 1.02, $p = 0.015$), and Th cells (log-HR = 1.72, $p = 0.0011$) attracting tumor cells were associated with poorer outcomes, possibly reflecting ineffective immune surveillance or tumor-driven immune evasion (DeNardo and Ruffell, 2019).

2.3.7.2 Spatial Interactions in Stage IV Patients

In Stage IV patients, distinct spatial immune interactions correlated with survival (Figure 2.8). Th cells attracting M2 macrophages (log-HR = 3.43, $p = 0.0029$) and tumor cells (log-HR = 1.10, $p = 0.0034$) were associated with worse outcomes, suggesting that Th cells in these contexts contribute to a pro-tumor immune environment.

Conversely, several interactions were linked to better survival, highlighting the role of immune coordination. These included B cells attracting T cells (log-HR = -1.19, $p = 0.0005$), T cells attracting B cells (log-HR = -0.995, $p = 0.014$), and NK cells attracting B cells (log-

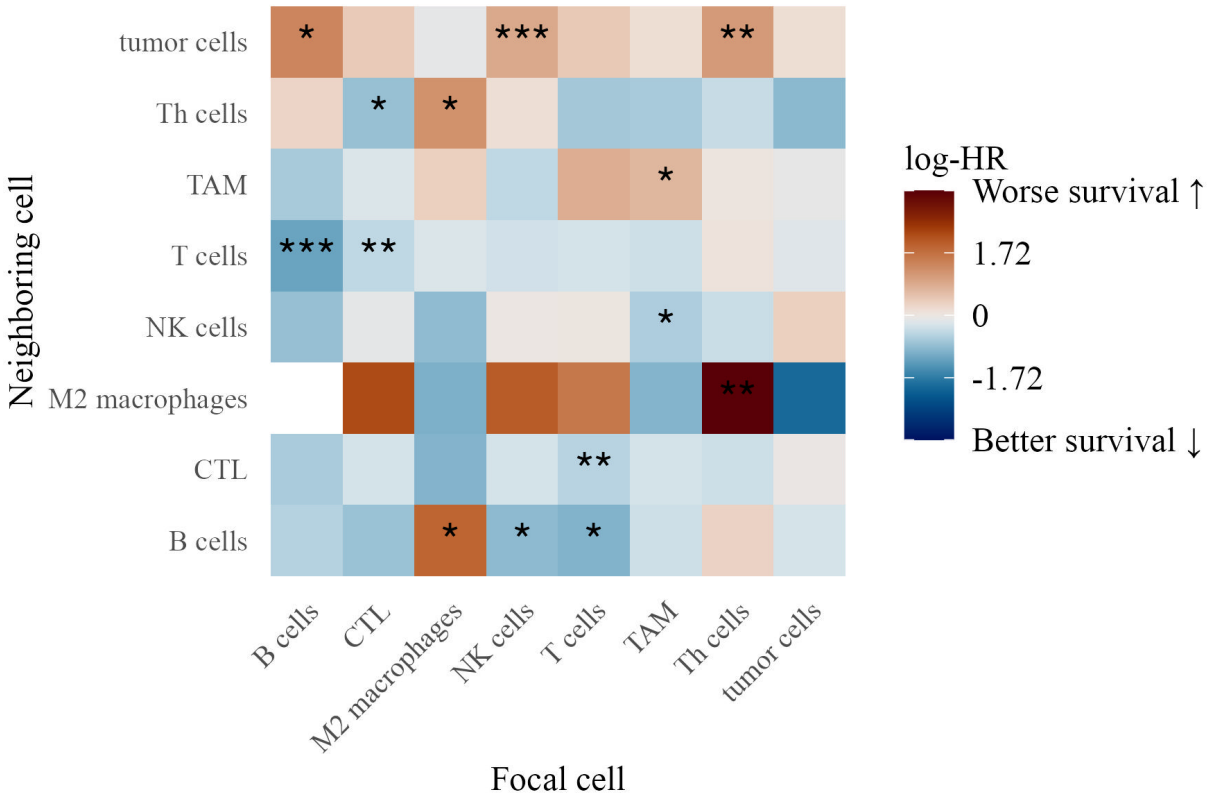


Figure 2.8: Heatmap of log-hazard ratios estimated using Cox regression, where the G-cross function at 40 μm between each pair of cell types in Stage IV patients serves as a predictor of survival time. Significance levels of hazard ratios are denoted as follows: $***$ ($p < 0.001$), $**$ ($p < 0.01$), and $*$ ($p < 0.05$).

HR = -0.888, $p = 0.045$), suggesting that reciprocal interactions among adaptive and innate immune populations support anti-tumor immunity. TAMs attracting NK cells (log-HR = -0.616, $p = 0.020$) was also associated with improved survival, potentially indicating that macrophage-NK cell interactions facilitate tumor clearance (Zhou et al., 2023).

T cells attracting CTLs (log-HR = -0.532, $p = 0.0035$) and CTLs attracting T cells (log-HR = -0.473, $p = 0.0078$) reinforced the role of T cell cross-talk in effective tumor control. Additionally, CTLs attracting Th cells (log-HR = -0.814, $p = 0.044$) was linked to improved outcomes, suggesting a coordinated cytotoxic-helper T cell response (Raskov et al., 2021).

The findings from both TMAs underscore the complexity of immune interactions in the tumor microenvironment. While certain immune clusters, such as TAM aggregation and NK cell-tumor cell proximity, were linked to poor survival, interactions among T cells, CTLs, NK cells, and B cells were associated with improved outcomes, highlighting the importance of spatial immune organization in melanoma progression.

2.3.8 Architectural Features and Cellular Neighborhoods Associated with Inflammatory Nodes

We further assessed the spatial clustering and dispersion patterns of immune cell pairs within inflammatory neighborhoods at distances of 20, 40, and 60 μm, using G-cross nearest neighbor estimates and Cox proportional hazards models. These analyses revealed distinct patterns of immune coordination and dysfunction across melanoma progression.

2.3.8.1 Spatial Inflammatory Interactions in Stage III Patients

In Stage III, B cell clustering in iNOS mPGES-NT neighborhoods at 60 μm was associated with better survival (log-HR = -1.69, $p = 0.016$) (Figure 2.9). Similarly, CTLs (log-HR = -1.37, $p = 0.037$) and NK cells (log-HR = -0.867, $p = 0.027$) clustering at 60 μm correlated with improved outcomes, suggesting that cytotoxic immune cells forming structured aggregates enhance anti-tumor responses.

Conversely, interactions associated with poor survival included Th cells attracting T cells (log-HR = 4.73, $p = 0.0055$) and TAMs (log-HR = 2.60, $p < 0.001$) at 40 μm, suggesting that these interactions contribute to an immunosuppressive environment. T cells attracting TAMs at 20 μm (log-HR = 2.78, $p = 0.0005$) were also associated with worse survival, possibly reflecting immune dysfunction or tumor-driven inflammation.

Macrophage interactions in CD74 CD44 MIF neighborhoods exhibited mixed survival associations. Increased clustering of TAMs around T cells at 20 μm (log-HR = 1.94, $p = 0.0005$) was associated with worse survival, consistent with macrophage-mediated suppression of T cell activity (Balogh, Templeton, and Cross, 2018). Similarly, tumor cells clustering around NK cells at 20 μm (log-HR = 1.10, $p = 0.0038$) and 40 μm (log-HR = 1.11, $p = 0.022$) correlated with worse outcomes, suggesting potential immune evasion mechanisms (Noe and Mitchell, 2020). Th cells clustering with tumor cells at 40 μm (log-HR = 2.22, $p = 0.0036$) and 60 μm (log-HR = 1.78, $p = 0.013$) further supported a role for these interactions in immune escape or chronic inflammation.

2.3.8.2 Spatial Inflammatory Interactions in Stage IV Patients

In Stage IV melanoma, spatial immune interactions within inflammatory neighborhoods showed significant associations with survival (Figure 2.10). Several interactions were associated with improved outcomes, suggesting that coordinated immune activity plays a role in tumor suppression.

B cells attracting NK cells in CD74 CD44 MIF neighborhoods was linked to better survival at 20 μm (log-HR = -1.48, $p = 0.033$) and 60 μm (log-HR = -1.73, $p = 0.0018$), indicating



Figure 2.9: Plots illustrating the significance of hazard ratios ($p < 0.05$) estimated from Cox proportional hazards models, where the G-cross function between each pair of cell types serves as a predictor of survival in stage III patients. Significant associations with worse outcomes are shown in red, while those with better outcomes are shown in blue. Within each subplot, estimates are presented relative to a focal cell type located within a specific inflammatory neighborhood, at radii of 20, 40, and 60 μm.

that NK cell presence close to B cells in these regions may contribute to an effective anti-tumor response. Similarly, NK cells attracting T cells at 40 μm (log-HR = -0.958, $p = 0.0001$) and 60 μm (log-HR = -1.16, $p < 0.0001$) suggests that NK-T cell interactions may play a role in immune surveillance (Lee et al., 2021; Fife et al., 2024).

Additional immune interactions in CD74 CD44 MIF neighborhoods were also linked to better survival. CTLs attracting T cells at 20 μm (log-HR = -0.587, $p = 0.011$) and 40 μm (log-HR = -0.436, $p = 0.023$) suggests that cytotoxic and helper T cell coordination contributes to positive outcomes (Raskov et al., 2021). M2 macrophages attracting NK cells at 40 μm (log-HR = -2.97, $p = 0.020$) and TAMs attracting NK cells at 20 μm (log-HR = -1.18, $p = 0.050$) and 60 μm (log-HR = -0.919, $p = 0.027$) suggests that macrophages in

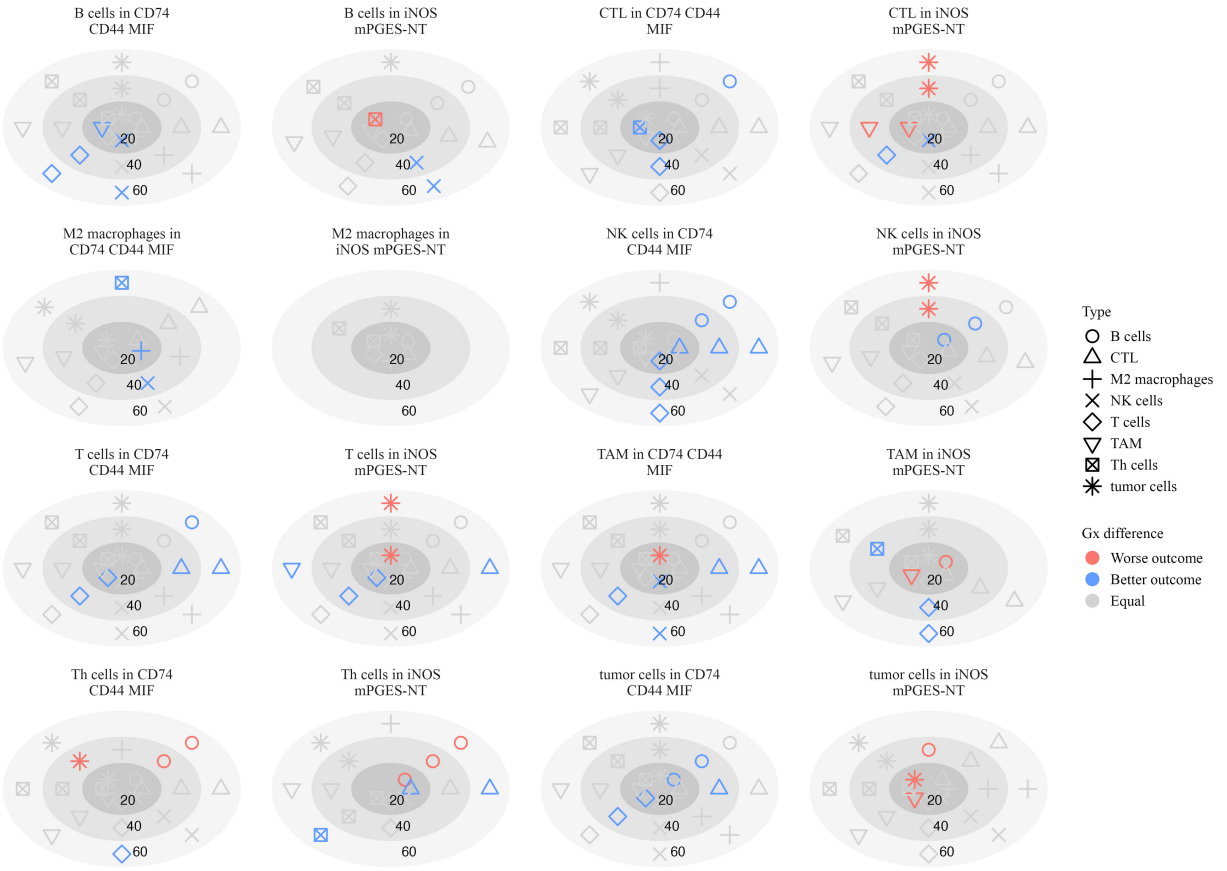


Figure 2.10: Plots illustrating the significance of hazard ratios ($p < 0.05$) estimated from Cox proportional hazards models, where the G-cross function between each pair of cell types serves as a predictor of survival in stage IV patients. Significant associations with worse outcomes are shown in red, while those with better outcomes are shown in blue. Within each subplot, estimates are presented relative to a focal cell type located within a specific inflammatory neighborhood, at radii of 20, 40, and 60 μm.

these neighborhoods may recruit NK cells to mediate tumor suppression (Kitamura, 2023).

In iNOS mPGES-NT neighborhoods, additional immune interactions correlated with improved survival. Th cells attracting Th cells at 60 μm (log-HR = -3.58, $p = 0.012$) and CTLs at 60 μm (log-HR = -3.41, $p = 0.033$) suggests a coordinated immune response rather than immune dysfunction. Additionally, CTLs attracting NK cells at 20 μm (log-HR = -0.939, $p = 0.024$) and B cells attracting NK cells at 60 μm (log-HR = -5.30, $p < 0.0001$) further support the role of cytotoxic cell recruitment in tumor control (Zhou et al., 2023).

Several immune interactions were associated with worse survival, particularly those involving excessive clustering of immune cells within inflammatory neighborhoods. In iNOS mPGES-NT neighborhoods, Th cells attracting B cells at 20 μm (log-HR = 7.53, $p = 0.028$)

and 40 μm (log-HR = 2.04, $p = 0.0007$) was linked to poor prognosis, suggesting that excessive B cell clustering around Th cells in these regions may contribute to immune dysregulation rather than an effective anti-tumor response. Similarly, TAMs attracting B cells at 20 μm (log-HR = 5.71, $p = 0.048$) and tumor cells attracting TAMs at 20 μm (log-HR = 4.03, $p = 0.0006$) were associated with poor survival, reinforcing the idea that macrophages in these settings may promote tumor progression rather than restrict it (Noy and Pollard, 2014; Wortman et al., 2021).

Th cells had several interactions in CD74 CD44 MIF neighborhoods that were associated with poor patient outcomes. Specifically, Th cells attracting B cells at 60 μm (log-HR = 1.39, $p = 0.03$) suggests that B cell clustering around Th cells in this neighborhood may not contribute to an effective anti-tumor response. Similarly, Th cells attracting tumor cells at 40 μm (log-HR = 0.973, $p = 0.023$) indicates that these interactions may facilitate immune evasion or chronic inflammation.

These findings underscore the role of spatial immune organization in shaping patient outcomes. While interactions involving NK cells, CTLs, and macrophages were linked to improved survival, excessive B cell and T cell clustering within pro-inflammatory neighborhoods, as well as macrophage recruitment, were associated with immune dysfunction and worse prognosis in advanced melanoma.

2.4 Discussion

The tumor immune microenvironment (TIME) plays a crucial role in melanoma progression and treatment response. Understanding the spatial organization of immune cells within inflammatory neighborhoods can reveal mechanisms of immune suppression, evasion, and activation that influence patient outcomes. This study provides a comprehensive analysis of immune spatial architecture in metastatic melanoma, highlighting key differences between Stage III and Stage IV disease and identifying immune interactions linked to survival.

2.4.1 Spatial Immune Reorganization in Stage IV Melanoma

Our results indicate a significant reorganization of immune cell distributions in Stage IV melanoma compared to Stage III. CD74- and MIF-enriched inflammatory neighborhoods were more prevalent in Stage IV, while iNOS-associated neighborhoods were significantly reduced. These findings suggest a shift from oxidative stress-driven inflammation in Stage III toward CD74-MIF-associated immune modulation in advanced disease. Previous studies have implicated CD74-MIF signaling in promoting tumor progression by suppressing cy-

totoxic immune responses and enhancing macrophage-mediated immune evasion (Tanese, Hashimoto, et al., 2015; Ekmekcioglu, Davies, et al., 2016).

Stage IV tumors exhibited a depletion of B cells and NK cells across multiple inflammatory neighborhoods, while cytotoxic T lymphocytes (CTLs) were more frequently found in CD74-MIF regions. This suggests that despite increased CTL presence, their effectiveness may be diminished in an environment enriched for immune-modulatory signals (Grimm, Sikora, and Ekmekcioglu, 2013). Tumor cells were also significantly more localized within CD74 MIF and MIF CD44 neighborhoods, reinforcing the role of these regions as hubs of immune suppression.

2.4.2 Immune Interactions and Survival Outcomes

We identified distinct immune interactions that correlated with patient survival. In Stage III melanoma, interactions involving B cells, macrophages, and CTLs were linked to improved outcomes. Specifically, CTLs attracting M2 macrophages and M2 macrophages attracting B cells were associated with better survival, suggesting that coordinated innate and adaptive immune responses contribute to tumor control (Piatakova et al., 2021; Su et al., 2024). Conversely, interactions where tumor cells attracted TAMs, and macrophages clustered within tumors, were associated with worse prognosis, supporting prior evidence that macrophage accumulation can promote immune evasion and tumor progression (DeNardo and Ruffell, 2019).

In Stage IV melanoma, cytotoxic immune interactions remained important for survival. B cells attracting NK cells and NK cells attracting T cells were associated with improved outcomes, suggesting that NK cells play a role in sustaining immune surveillance. Similarly, CTLs attracting Th cells were linked to better survival, reinforcing the importance of coordinated adaptive immune responses (Raskov et al., 2021). These findings align with recent work demonstrating that effective tumor control relies on dynamic crosstalk between cytotoxic and helper T cell populations.

Conversely, several immune interactions were associated with worse survival, particularly those involving excessive clustering of immune cells within inflammatory neighborhoods. In iNOS mPGES-NT neighborhoods, Th cells attracting B cells at short distances correlated with poor prognosis, suggesting that excessive B cell clustering in these regions may contribute to immune dysfunction rather than anti-tumor immunity. Similarly, TAMs attracting B cells and tumor cells attracting TAMs were linked to worse survival, reinforcing the idea that macrophages in these contexts may promote tumor progression rather than restrict it (Noy and Pollard, 2014; Wortman et al., 2021).

2.4.3 TAM and CTL Dynamics in the TIME

Our spatial analysis of G-cross nearest neighbor interactions provided additional insights into the complex roles of macrophages and CTLs in melanoma. While some TAM-CTL interactions were linked to better survival, excessive TAM accumulation was associated with poorer outcomes. This asymmetry suggests that TAMs can either support or suppress anti-tumor responses depending on their activation state and spatial organization within the tumor bed. Previous work has shown that TAMs can adopt both pro- and anti-inflammatory phenotypes depending on local signaling cues (Kitamura, 2023). Our findings further emphasize the need to functionally characterize macrophage subsets in melanoma tumors to determine their precise role in immune regulation.

Another key observation was the differential immune cell composition between Stage III and Stage IV tumors. Stage III tumors exhibited a higher proportion of naive T cells, helper T cells, NK cells, and B cells, whereas Stage IV tumors were dominated by CTLs and TAMs. While increased CTL infiltration is often associated with improved prognosis, our data suggest that not all CTLs in Stage IV tumors exhibit effective cytotoxic function. Emerging studies indicate that tumor-infiltrating CTLs can become dysfunctional or exhausted in immunosuppressive microenvironments, reducing their efficacy (McKenzie and Valitutti, 2023). The prolonged interactions observed between CTLs and TAMs in Stage IV may further contribute to T cell dysfunction.

2.4.4 Clinical and Therapeutic Implications

These findings have important implications for immunotherapy strategies. Given the enrichment of CTLs in CD74- and MIF-associated neighborhoods in Stage IV tumors, interventions targeting these pathways may help restore CTL function and improve therapeutic outcomes. Previous studies have shown that blocking CD74-MIF signaling can enhance T cell infiltration and reduce tumor growth (Tanese, Hashimoto, et al., 2015). Similarly, targeting mPGES1, which we identified as a key player in immune evasion of melanoma, may enhance responses to immune checkpoint inhibitors (Fukuda et al., 2023).

The spatial organization of immune cells may also inform patient stratification for immunotherapy. Our results suggest that patients with increased NK cell-B cell interactions or structured CTL-Th cell clustering may have better responses to immune-based treatments. Conversely, patients with excessive macrophage clustering or high B cell accumulation in inflammatory regions may require combination therapies that target these immunosuppressive niches.

2.4.5 Study Strengths and Limitations

The strengths of this study lie in the size of our cohort and the high-resolution spatial analysis of immune interactions. By integrating multiplex imaging with computational modeling, we were able to identify key differences in immune cell organization between Stage III and IV melanoma. However, several limitations should be noted. First, the retrospective nature of this study introduces potential biases related to treatment variability. Due to the specifics of these TMA cohorts' establishment, we were limited in the clinical information available, and, as a result, our conclusions carefully refrain from making any statements on treatment outcomes or responses to therapy. Therefore, our focus remains on the differences in TIME immunobiology between stages, which may inform future decisions regarding IO approaches in clinical trials. Additionally, while our analyses reveal associations between spatial immune patterns and survival, they do not establish causal relationships. Further functional validation is required to determine how specific immune interactions influence tumor progression and response to therapy.

2.4.6 Conclusions and Future Directions

Our findings underscore the importance of spatial immune organization in melanoma progression and treatment response. We identified distinct immune interactions that correlate with survival, highlighting the roles of cytotoxic immune coordination and macrophage-driven immune suppression in shaping patient outcomes. Future studies should focus on validating these findings in independent cohorts and exploring therapeutic strategies that modulate immune cell spatial organization to improve clinical responses.

CHAPTER 3

Multitype Gibbs Modeling of Cell–Cell Interactions and Survival Association

3.1 Introduction

Similar to natural ecosystems, the TME comprises a dynamic web of relationships, where cancer cells, immune populations, stromal cells, and various molecular cues engage in complex communication and interdependence (Bożyk et al., 2022; Visser and Joyce, 2023)[1]. Furthermore, just as environmental conditions influence species behavior in ecological systems, the TME's microenvironmental factors, including nutrient availability, oxygen levels, and inflammatory signals, play a pivotal role in shaping cellular behaviors, proliferation, migration, and responses to therapy (Boedtkjer and Pedersen, 2020; Petrova et al., 2018; Y. Li, Zhao, and X.-F. Li, 2021). These interactions exert pivotal influences on tumor progression and therapeutic outcomes (Tan and Naylor, 2022; Sun, 2016; Ni et al., 2021; X. Li et al., 2021). This ecosystem-like framework underscores the TME's complexity and underscores the need to employ ecological modeling approaches to gain deeper insights into its intricate dynamics and therapeutic implications (Pienta et al., 2008; Adler et al., 2022; Somarelli, 2021; Thomas et al., 2020; Reynolds, M. W. Oli, and M. K. Oli, 2020).

The advent of highly multiplexed imaging technologies that allow for the simultaneous measurement of dozens of cell-surface protein biomarkers in tissue samples has allowed for the creation of datasets that have precise spatial coordinates and cell-type annotations for each cell within a tissue (Dam, Baars, and Vercoulen, 2022). Such multiplexed imaging technologies are of particular relevance to studies of the tumor microenvironment, in which multiplexed images can be used to ascertain the complex interrelationships between cell types and cell states (Schürch et al., 2020; Black et al., 2021). Such imaging technologies are similar to the remote sensing technologies employed in ecological studies, which allow

[1]This chapter is based on the publication Eliason and Rao (2024)

the localization of various species and environmental covariates in space.

Spatial descriptive statistics are fundamental tools employed in such ecological studies to gain insights into the spatial patterns and relationships within natural systems. These statistics provide a quantitative framework for characterizing the arrangement, dispersion, and clustering of species or individuals across a landscape. Metrics such as nearest neighbor distances, spatial autocorrelation, and Ripley's K-function are commonly used to assess the degree of spatial aggregation or dispersion of species within a given habitat (A. Baddeley and Turner, 2005; Fortin, Dale, and Ver Hoef, 2016; Ben-Said, 2021; Velázquez et al., 2016). By analyzing these patterns, ecologists can infer crucial information about species interactions, habitat preferences, and the influence of environmental factors on distribution. Such statistics have also been widely adopted by researchers investigating the tumor microenvironment (Canete et al., 2022; Vu et al., 2022; Creed et al., 2021). More bespoke methods have also been developed, such as those in which a "neighborhood composition" vector is calculated for each cell (Schürch et al., 2020; Behanova, Klemm, and Wählby, 2022; Z. Yuan et al., 2022; Tanevski et al., 2022). Other relatively simple statistics, such as quantifying the amount of immune infiltration based on the Immunoscore metric, in various regions of the TME, are also widely used in clinical models (Galon et al., 2012; Allam et al., 2022).

While spatial descriptive statistics offer essential preliminary insights into ecological spatial patterns, more advanced techniques, such as Joint Species Distribution Models (JSDMs), extend this analysis by capturing the complex interactions and dependencies that characterize ecological systems in a more holistic manner (Tikhonov, Opedal, et al., 2020; Tikhonov, Abrego, et al., 2017; Ovaskainen et al., 2016; Wilkinson et al., 2021; Clark et al., 2014). In particular, JSDMs are able to model the spatial distributions of multiple species simultaneously, while accounting for the associations between each of these species, as well as their association with environmental variables. Within the context of the TME, these models can enable researchers to identify cell type-cell type or cell type-environment associations while adjusting for other associations, predict cell type distributions, and quantify the uncertainty of model parameters. Such capabilities are typically not available for spatial descriptive statistics.

Point process models are another widely used tool for modeling ecological relationships (Velázquez et al., 2016; Stoyan and Penttinen, 2000; Law et al., 2009; Illian, Martino, et al., 2013; Illian, Møller, and Rasmus P. Waagepetersen, 2009; Renner et al., 2015). Such models make explicit the probabilistic data-generating process that is assumed to generate an observed pattern of points, and thus are amenable to uncertainty quantification and producing predictions (A. Baddeley and Turner, 2005). A particular class of point process model, called multitype Gibbs point process (MGPP) models, are able to model the pairwise

cell type interactions jointly (A. Baddeley, Rubak, and Turner, 2015; Rajala, Murrell, and Olhede, 2018; A. Baddeley and Turner, 2005; Flint et al., 2022), along with cell type/environment interactions and thus can be used as JSDMs in the context of the TME ecosystem. Gibbs point process models explicitly model the attraction or inhibition between different points in space; MGPPs generalize this to when there are several types of points in space, such as different cell types (A. Baddeley and Turner, 2005).

More specifically, modeling interactions via MGPP models can improve on the aforementioned spatial descriptive statistics in a number of ways (A. Baddeley and Turner, 2005): a) spatial variables, such as the concentration levels of various inflammatory factors, and non-spatial variables, such as organ of origin or patient demographics, can be easily adjusted for, allowing for the modeling of inhomogeneous distributions of cells; b) the joint modeling of pairwise interactions between cell types can be achieved, as opposed to the case of descriptive statistics, in which case only the marginal distribution of interactions can be estimated; c) uncertainty quantification of interaction effects can be performed; and d) predictions and residuals can be obtained from the model, allowing for the assessment of model fit, as well as the prediction of distributions of cells under new conditions.

Some examples of relevant questions that JSDMs (and thus MGPP models, in this context) can answer, in a general setting and specific to the TME, are found in Table 3.1. In order to motivate the utility of this model, in this paper we will demonstrate an analysis framework that could be used to answer many of the questions from Table 3.1. The questions in Table 3.1 are asked either at the image level (that is, they can be answered after a JSDM model is fit on a particular image) or at the cohort level (they are answered after fitting JSDM models for each image in a cohort, extracting parameters from the fitted models and then using those parameters as inputs for further modeling and analysis). The third column of Table 3.1 indicates at which level each question can be answered.

Table 3.1: Examples of questions that can be asked using MGPP models as JSDMs.

Example of question in a general setting	**Specific to the TME**	**Image or Cohort level**
Are different cell types *positively* or *negatively correlated* in space, after adjusting for the presence of other cell types and spatial covariates?	Are B cells and granulocytes positively or negatively correlated in this patient, after controlling for the effects of other cells?	Image level

Example of question in a general setting	**Specific to the TME**	**Image or Cohort level**
How do cell type interactions change across different spatial scales?	At which spatial scales are Tregs repulsive of one another, and at which are they attractive?	Image level
Which cell type or spatial covariate is the best predictor of another cell type at each location in an image?	Which cell type best predicts the locations of tumor cells in the tumor core? How about in stromal regions?	Image level
Given some cell types and spatial covariates in an image, how well can we predict the distribution of a new cell type, using a model fit on other images?	Can we correctly predict the distribution of M2 macrophages in an image where we did not mark for M2 macrophages?	Image level (but may require model fits from other images)
Can the estimated interaction parameters predict patient outcomes?	How well does the interaction between CD8+ T cells and granulocytes predict patient survival?	Cohort level
Can we predict which distributions of cell types are associated with the best tumor responses?	Which distributions of immune cell types are associated with a near absence of tumor cells?	Image level
How do changes in spatial covariates change the distribution of different cell types?	How might changes in oxygenation or angiogenesis affect the distribution of tumor cells?	Image level

Example of question in a general setting	Specific to the TME	Image or Cohort level
Is there evidence of two cell types coexisting in the same metabolic niche? Is the nature of their interaction competitive or mutualistic?	Do immune cells and tumor cells compete for the same metabolic factors? In which areas of the tumor?	Image level
How do cell type interactions compare across images obtained from the same patient?	How variable are the interaction parameters between B cells/CD4+ T cells across images within the same patient among patients with low survival probability?	Cohort level

3.2 Methods

3.2.1 Dataset

The dataset used to showcase the MGPP method comprises multiplexed images of tumor tissue samples obtained from a publicly available repository for colorectal cancer (Schürch et al., 2020). The images capture different cell types within the tumor microenvironment, including cancer cells, immune cells, and stromal cells. In the original study, preprocessing was conducted on the images. This involved cell type classification and determination of spatial coordinates for each cell type.

Specifically, this dataset consists of 4 multiplexed images of 56 different cell-surface markers of tumor tissue sections from each of 35 CRC patients, with patients having varying degrees of disease progression and severity. These images were obtained via the PhenoCycler imaging technology (formerly known as CODEX) (Goltsev et al., 2018; Black et al., 2021). The patients are divided into two groups, based on the histopathology of their tumors: those in which tertiary lymphoid structures (TLSs) are visible at the tumor invasive front, and those where diffuse inflammatory infiltration (DII) was present and no TLSs were visible. These patient groups will be subsequently referred to as CLR (Crohn's-like reaction) and

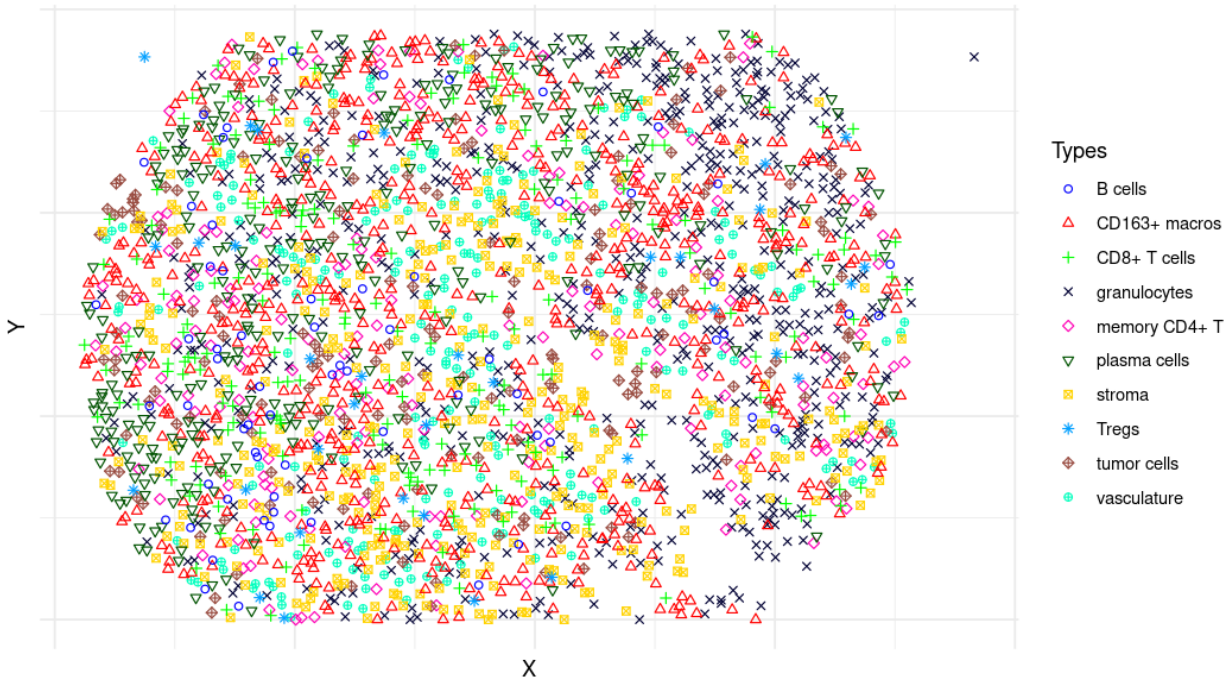


Figure 3.1: A point pattern representation of a multiplex image of a tissue section of the TME in a CRC patient

DII, with CLR patients generally having better survival prognosis (Schürch et al., 2020). 16 cell types have been annotated using the 56 different cell-surface markers and in total, the dataset consists of 244,504 annotated cells. In addition, patient metadata is included, including overall/progression-free survival times, tumor grade, and demographic variables such as age and sex. An example image can be seen in Figure 3.1.

3.2.2 Multitype Gibbs Point Process (MGPP) Models

Once cell types have been annotated, a multiplex image, such as in Figure 3.1, can be viewed as a multitype point pattern, or a realization $\mathbf{x}$ of a particular random variable $\mathbf{X}$ called a *multitype point process*. Subsequently, a MGPP model can be fit to this multitype point pattern. In the following, we will describe first a general *pairwise interaction* MGPP model, and subsequently the *saturated pairwise interaction* MGPP model.

Mathematically, a pairwise interaction MGPP model characterizes the probability density $f(\mathbf{X})$ of a multitype point process $\mathbf{X}$ in the following way (A. Baddeley, Rubak, and Turner,

2015), with i and j indexing over the points in $\mathbf{x}$:

$$f(\mathbf{X} = \mathbf{x}) = C \exp\left(\left[\sum_{i=1}^{n(\mathbf{x})} \sum_{k=1}^{K} \beta_{k,m_i} X_k(x_i) \right] + \left[\sum_{i<j}^{n(\mathbf{x})} g_{m_i,m_j}(x_i, x_j) \right] \right) \tag{3.1}$$

Here, C indicates a normalizing constant, while the first term in brackets signifies the "first-order" effects on the density of each cell type, that is, the effects from spatial and non-spatial variables. There are K such variables (the $X_k(x_i)$) per cell type. These variables are typically used to model "environmental variables"; in the context of the TME, this could include such variables as the concentration levels of various metabolic or inflammatory factors, or the distance to the nearest blood vessel. The second bracketed term contains the pairwise interaction functions $g_{m_i,m_j}(x_i, x_j)$, which are functions of two points x_i and x_j of types m_i and m_j. These interaction functions are often parameterized with interaction parameters θ_{m_i,m_j}, which typically quantify the direction and magnitude of interaction between points of type m_i and m_j. For example, in the multitype Strauss model, the interaction function $g_{m_i,m_j}(x_i, x_j)$ is written as:

$$g_{m_i,m_j}(x_i, x_j) = \begin{cases} 0, & \text{if } ||x_i - x_j|| > R_{m_i,m_j} \\ \theta_{m_i,m_j}, & \text{if } ||x_i - x_j|| \leq R_{m_i,m_j} \end{cases} \tag{3.2}$$

for user-defined *interaction radii* R_{m_i,m_j}. The multitype Strauss model assumes that interactions between cells of types m_i and m_j do not exist if cells are farther than R_{m_i,m_j} apart, and are constant if they are within that radius.

One drawback of many pairwise interaction MGPP models, such as the multitype Strauss model, is that they are only able to model repulsion (and not clustering) between points - that is, that $\theta_{m_i,m_j} < 0$. This has led to the development of MGPP models that do allow clustering, such as the *Geyer saturation* model (C. Geyer, 1999) or the *saturated pairwise interaction* MGPP model (Flint et al., 2022). For example, in the saturated pairwise interaction Gibbs point process, as in (Flint et al., 2022) and borrowing some of their notation,

we have

$$\begin{aligned} f(\mathbf{x}) =& C \exp\left[\sum_{x\in\mathbf{x}}\left(\beta_{0,t} + \sum_{k=1}^{K}\beta_{k,t(x)}X_k(x)\right)\right. \\ &+ \sum_{t_2=1}^{p}\sum_{z\in\mathbf{x}_{t_1}}\alpha_{t_1,t_2}u\left(z,(\mathbf{x}\backslash\{z\})_{t_2}\right) \\ &\left.+ \sum_{t_2=1}^{p}\sum_{z=\in\mathbf{x}_{t_1}}\gamma_{t_1,t_2}v\left(z,(\mathbf{x}\backslash\{z\})_{t_2}\right)\right] \end{aligned} \tag{3.3}$$

with

$$u(x,\mathbf{x}_{t_2}) = \max_{\eta\in S(\mathbf{x}_{t_2},N)}\sum_{z\in\eta}\phi(||x-z||;R^S_{t(x),t_2)}) \tag{3.4}$$

and v defined similarly as

$$v(x,\mathbf{x}_{t_2}) = \max_{\eta\in S(\mathbf{x}_{t_2},N)}\sum_{z\in\eta}\psi(||x-z||;R^M_{t(x),t_2)}) \tag{3.5}$$

Here, t_1 and t_2 can take on values in $\{j\}_{j=1}^p$ and denote the j^{th} cell type of p cell types, and $t(x) \in \{j\}_{j=1}^p$ denotes the type of point $x \in \mathbf{x}$. Note that $t(x_i) = m_i$, as in our notation in Equation 3.1. Functions ϕ and ψ represent user-defined short- and medium-range *potential functions* which set the shape of short- and medium-range interactions, and are dependent only on the distance between two points. ϕ and ψ are commonly functions such as the exponential or step functions (Figure 3.2). $R^S_{t_1,t_2}$ and $R^M_{t_1,t_2}$ are user-defined short- and medium-range interaction radii, and the interaction parameters of interest, α_{t_1,t_2} and γ_{t_1,t_2}, are interpreted as the strength and direction of interaction between any two types of points at short and medium ranges. $S(\mathbf{x}_{t_i}, N)$ is the set of all subsets of points from $\mathbf{x}_{t_i}$ that are of size N - thus, u and v can be interpreted as the sum of the N largest potentials that point x has with points $z \in \mathbf{x}$ of type t_2. Summing only the top N potentials (rather than all of them) is what leads to this being called a *saturated* MGPP model, and the parameter N is called the *saturation parameter* accordingly. Because of its great flexibility in being able to model clustering and repulsion, the saturated pairwise interaction MGPP model will be the one we will be using for the remainder of the paper as the JSDM. It should be noted, furthermore, that currently MGPP models, like the saturated pairwise MGPP model we describe above, are limited to modeling interactions between pairs of points. While higher-order interaction *single-type* Gibbs models do exist, such as the triplet interaction process (C. Geyer, 1999) and the area-interaction process (A. J. Baddeley and Lieshout, 1995), from the authors' perspective, there has not been any work yet to generalize higher-order interaction processes to *multitype* data.

Once the probability density $f(\mathbf{X})$ of a MGPP model has been defined, estimating the

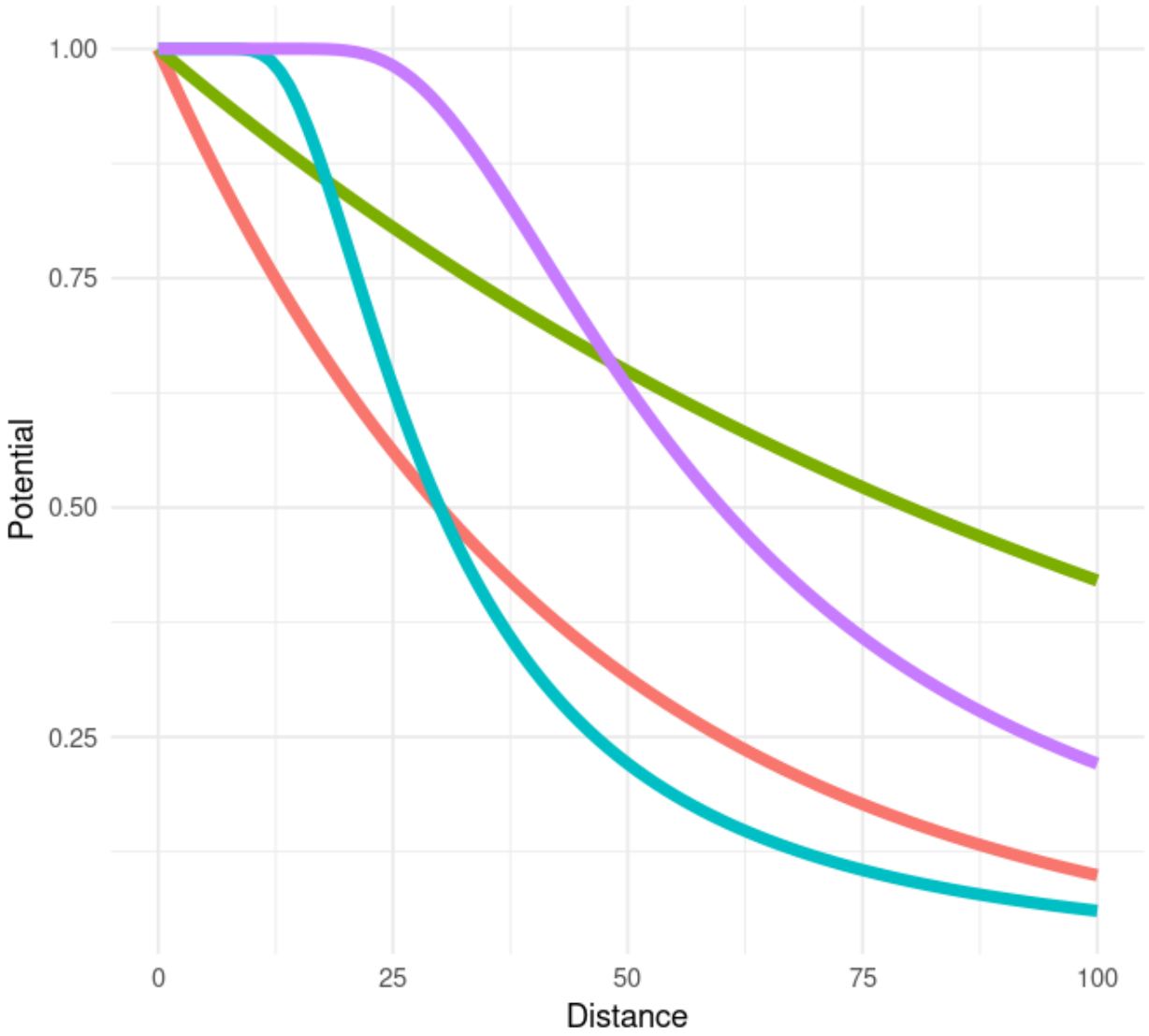


Figure 3.2: Examples of potential curves commonly used for ϕ and ψ in Equations 3.4 and 3.5.

parameters of this model can proceed via standard statistical techniques such as maximum likelihood estimation. However, because this is often quite computationally costly (due to the often intractable normalizing constant C), estimation techniques based on the *conditional intensity* $\lambda(y|\mathbf{x})$ are often preferred. The *intensity* $\lambda(y)$ of a point pattern at a location y is the expected number of points falling in a small window around y, divided by the area of that small window. The *conditional intensity* defined at a spatial location y is defined as $\lambda(y|\mathbf{x}) = \frac{f(\mathbf{x} \cup y)}{f(\mathbf{x} \setminus y)}$, and can be interpreted as the "intensity at y given the rest of the spatial pattern $\mathbf{x}$" (A. Baddeley and Turner, 2005). Using the conditional intensity, we can then produce predictions from our MGPP model and subsequently assess model fit. For the specific formulation of the conditional intensity for the saturated pairwise interaction Gibbs point process model, see (Flint et al., 2022).

In case there are cell types with very low representation, there are a few possible strategies. If an analyst is not interested in interactions involving rare cell types, those cell types can simply be removed from the pattern or lumped into a separate category, as we do in the example analysis below. However, if the analyst is interested in rare cell-type interactions, it is possible to fit a multilevel model across multiple images simultaneously, in order to pool the estimates of interaction strength across multiple images simultaneously. Such a strategy is used to model spatiotemporal point pattern data in (Illian and Hendrichsen, 2010; King et al., 2012), and similar methods could be applied in this case as well.

3.2.3 An Example Analysis

Here we demonstrate an example analysis to answer a subset of the questions from Table 3.1. In this example, we jointly model the interactions between sufficiently abundant cell types in each image. Furthermore, we include as a spatial covariate the distance from each point to the nearest blood vessel, as a proxy for angiogenesis and oxygenation found in the TME. Specifically, we set out to answer the following:

1. Which cell types interact with one another, after adjusting for the interactions of other cell types and distance to blood vessels? Which cell types stay away from one another, again, after adjusting for the interactions of other cell types and the distance to blood vessels? What confidence can I have in individual parameter estimates?
    (a) Do these interactions occur at close or far spatial scales?
    (b) Which cell type most strongly uniquely interacts with tumor cells, on average, across the whole image?

2. Which cell types best predict the distribution of tumor cells at each location in space?

3. How much is the distribution of each cell type uniquely predicted by its distance to the nearest blood vessel?

4. What is the difference in interaction strength between each pair of cell types, comparing between the two patient groups (CLR and DII patients)?

5. What levels of intra- and inter-patient heterogeneity are there in how different cell types interact?

6. Are any of these parameter estimates useful clinically for predicting patient outcomes?

7. How much can I trust the results from this model? i.e., how well does it predict spatial distributions of cells?

The code to reproduce the example analysis that follows can be found at https://github.com/jeliason/mgpp_TME/.

3.2.3.1 Fitting MGPP models

We employed a systematic approach in fitting the MGPP models to the images from the CRC dataset. We focused our analysis on cell types that exhibited a minimum of 20 cells within a given image, ensuring that we focused on sufficiently populated cell populations to enable reliable modeling of interactions. Next, we identified regions devoid of cells and incorporated this information as a spatial covariate within the model. By accounting for these regions, we hoped to avoid any bias in model fitting from assuming all areas in the spatial window were equally likely to contain cells. We further incorporated a spatial covariate of "distance to nearest blood vessel", as a proxy of angiogenesis and oxygenation throughout the TME. An example of this spatial covariate can be seen in Figure 3.3.

The fitting of the MGPP models was performed using the `ppjsdm` package, an R package tailored for fitting saturated pairwise interaction Gibbs point process models (Flint et al., 2022). Specifically, we employed an exponential interaction function to model the pairwise interactions between cell types at both short and medium ranges (that is, the functions ϕ and ψ in Equations 3.4 and 3.5). The interaction radii were set at 30, 70, and 150 microns, reflecting short-, medium- and long-range interaction distances, respectively. A saturation parameter of $N = 5$ was set. Lastly, $N_d = 5,000$ dummy points were included per cell type. The dummy points are related to the inferential procedure for parameter estimation. Briefly,

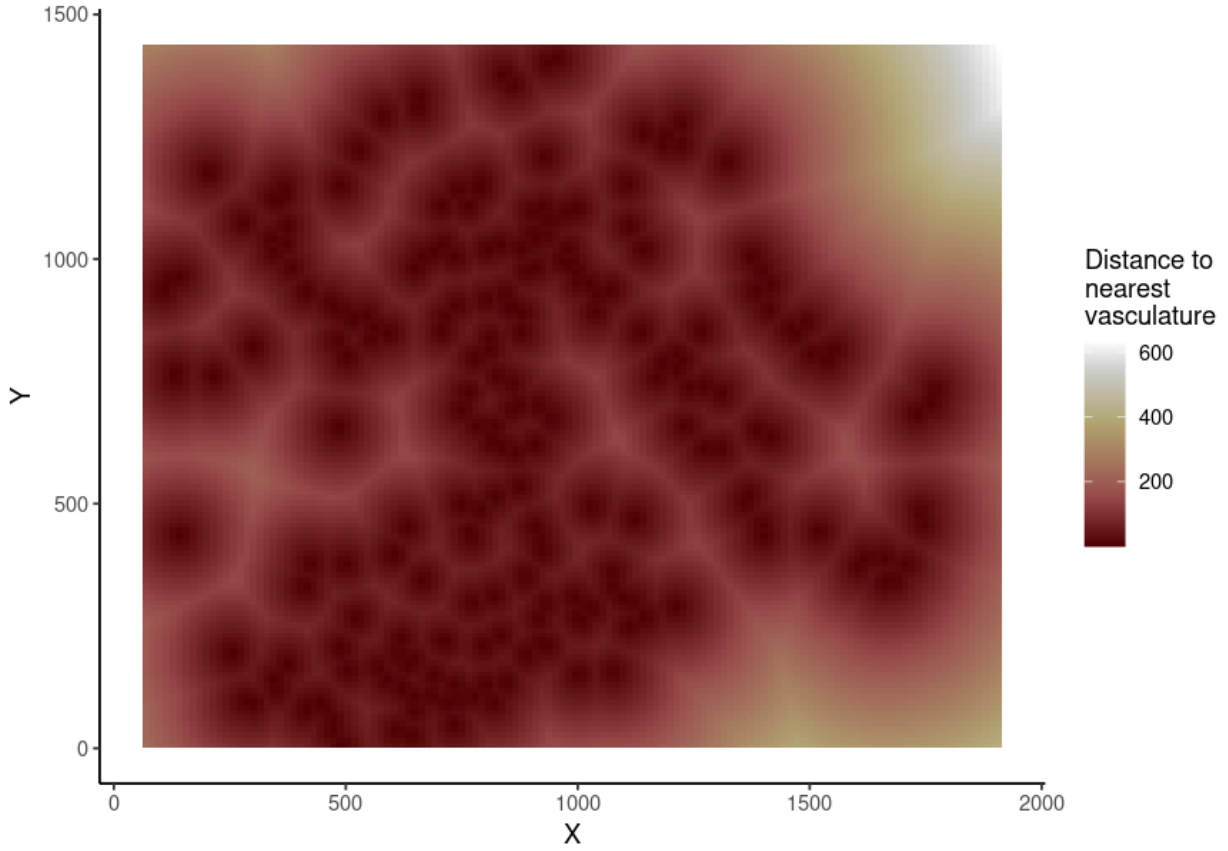


Figure 3.3: Spatial covariate of "distance to nearest blood vessel", as imaged in spot 59_A.

N_d dummy points per cell type are distributed across the image and interactions between dummy points and real points are calculated, according to Equations 3.4 and 3.5. This way of formulating the problem allows one to use fast, optimized inferential procedures developed for logistic regression. Specifically, at each point (dummy or data), the interactions with each cell type at that point are calculated, as well as the associations with spatial covariates. Then, a logistic regression is fitted in which the interactions are the independent variables and the response variable is binary, coded 1 for data points and 0 for dummy points. Increasing N_d will decrease the approximation error of this inferential technique up to a point, and an estimate of the proportion of the standard error of each coefficient that is due to this approximation can additionally be calculated. For more information, the interested reader is referred to (A. Baddeley, Coeurjolly, et al., 2014; Flint et al., 2022).

Using 2 cores on a laptop, fitting of the models on these datasets took between 30 seconds - 2 minutes per model, while estimating the standard errors for each parameter (a separate step) took 3-10 minutes per model. Memory costs for the fitting of these models are also not extravagant, typically only taking 1-2 GB.

3.2.3.2 Interpretation of Model Parameters and Predictions

After fitting the models, we extracted the fitted interaction parameters α_{m_i,m_j},γ_{m_i,m_j} and β_{1,m_i} (corresponding to the effect that distance to nearest blood vessel had on the distribution of each cell type m_i), along with the respective confidence interval for each parameter. We then plotted these parameters for one image as heatmaps, in order to demonstrate how to answer Questions 1 and 3. We also plotted the exponential potential functions scaled by each interaction parameter, in order to demonstrate how to answer Question 1a. We identified the cell type with the highest magnitude interaction parameter with tumor cells in each image, to answer Question 1b.

Next, we showed how one can answer Question 2 by identifying the cell type that has the largest magnitude contribution to the linear predictor of the conditional intensity of tumor cells at each point in space.

Lastly, we extracted the interaction parameters from all fitted models and plotted heatmaps of the difference in median interaction between the two patient groups, allowing us to answer Question 4.

3.2.3.3 Heterogeneity in Model Parameter Estimates Within Patient Groups and Across Patients

In order to answer Question 5, we calculated the median absolute deviation (MAD) of each interaction parameter at two levels: pooled within each patient group, as well as across the whole patient cohort, and plotted the MAD estimates as heatmaps. This allows us to identify interactions that are highly variable and likely to be more context-dependent, as well as those that remain more consistent and homogeneous.

3.2.3.4 Predictive Modeling of Patient Outcomes

We employed multivariate Cox proportional hazards models to estimate hazard ratios for interaction effects derived from the fitted MGPP models, in order to answer Question 6. Briefly, the interaction parameters for a given pair of cell types and spatial scale were extracted and used as a predictor of patient survival in a Cox model, while adjusting for the age and sex of each patient, as in Equation 3.6.

$$\begin{aligned}\lambda(t|\alpha_{ijk}) = \lambda_0(t)\exp(&\beta_{0,jk} + \beta_{1,jk}\alpha_{ijk} \\ &+ \beta_{2,jk}Age_i + \beta_{3,jk}Sex_i)\end{aligned} \tag{3.6}$$

Here, α_{ijk} indicates the α_{jk} parameter from the i-th image. Models were fit for the γ_{jk} parameters in a similar fashion. Pairs of types that appeared in less than 10 images were excluded from this analysis. The resulting hazard ratios serve as potential prognostic biomarkers, indicating the prognostic significance of the identified spatial interactions within the tumor microenvironment. Furthermore, to account for the fact that there are multiple interaction estimates per patient (because there are multiple images, and thus multiple model fits, per patient), we fit these Cox models using a robust sandwich variance estimator for each fitted parameter that appropriately accounts for the correlation between interaction estimates within patients (Spiekerman and Lin, 1998).

3.2.3.5 Evaluation of Model Fit

Finally, in answering Question 7, we assessed the fit of the MGPP models by employing the Area Under the Curve (AUC) metric to quantify the predictive accuracy that the model had for each cell type within the tumor microenvironment. Briefly, the AUC is a metric that determines how well the predicted conditional intensity is able to partition the spatial window into regions of higher and lower concentrations of points (A. Baddeley and Turner, 2005). By examining the AUC for each cell type, we are able to see how well the model

Table 3.2: Coefficients β_{1,m_j} as estimated from spot 59_A.

type	estimate	CI
B cells	-1.4e-04	(-6.2e-03, 5.9e-03)
CD163+ macros	-1.5e-03	(-2.9e-03, -6.1e-05)
CD8+ T cells	9.9e-04	(-1.8e-03, 3.8e-03)
granulocytes	-2.2e-03	(-3.4e-03, -1.1e-03)
memory CD4+ T	-1.9e-03	(-4.7e-03, 9.7e-04)
plasma cells	-1.0e-03	(-2.9e-03, 9.3e-04)
stroma	-6.9e-03	(-9.6e-03, -4.3e-03)
Tregs	3.4e-03	(-5.2e-03, 1.2e-02)
tumor cells	-1.9e-03	(-5.2e-03, 1.3e-03)

predicts the spatial distribution of each cell type. We then calculated the median AUC per cell type, both within each patient group and across patient group, which helped identify cell types that could be reliably predicted by the MGPP models. Furthermore, we calculated the correlation between the AUC values and the abundance of each cell type, to clarify whether cell type abundance improved predictability of that cell type.

3.3 Results from Example Analysis

3.3.1 Interpretation of Model Parameters and Predictions

We extracted the β_{1,m_j} parameters from each of the fitted models, and examples of the fitted parameters from one image are show in Table 3.2, allowing us to answer Question 3. The β parameters from this image demonstrate that the intensity of stroma, granulocytes and CD163+ macrophages are negatively associated with distance from vasculature - that is, as the distance from vasculature increases, the less likely it is that stromal cells, granulocytes or CD163+ macrophages will be encountered.

We next extracted the α_{m_i,m_j} and γ_{m_i,m_j} (subsequently α and γ) parameters from each of the fitted models. Examples of the fitted parameters from one image are given in Figure 3.4. The α parameters from this image demonstrate relatively consistent within-type positive interaction, or clustering, while between-type interactions at this scale tend to be negative, or repulsive. Of course, a significant outlier to this trend is the interaction between B cells and many other cell types, in particular with memory CD4+ T cells. The γ parameters, on the other hand, showcase a quite different story. Here, we again see that within-type interactions are generally positive (with the notable outlier of Tregs), though now tumor

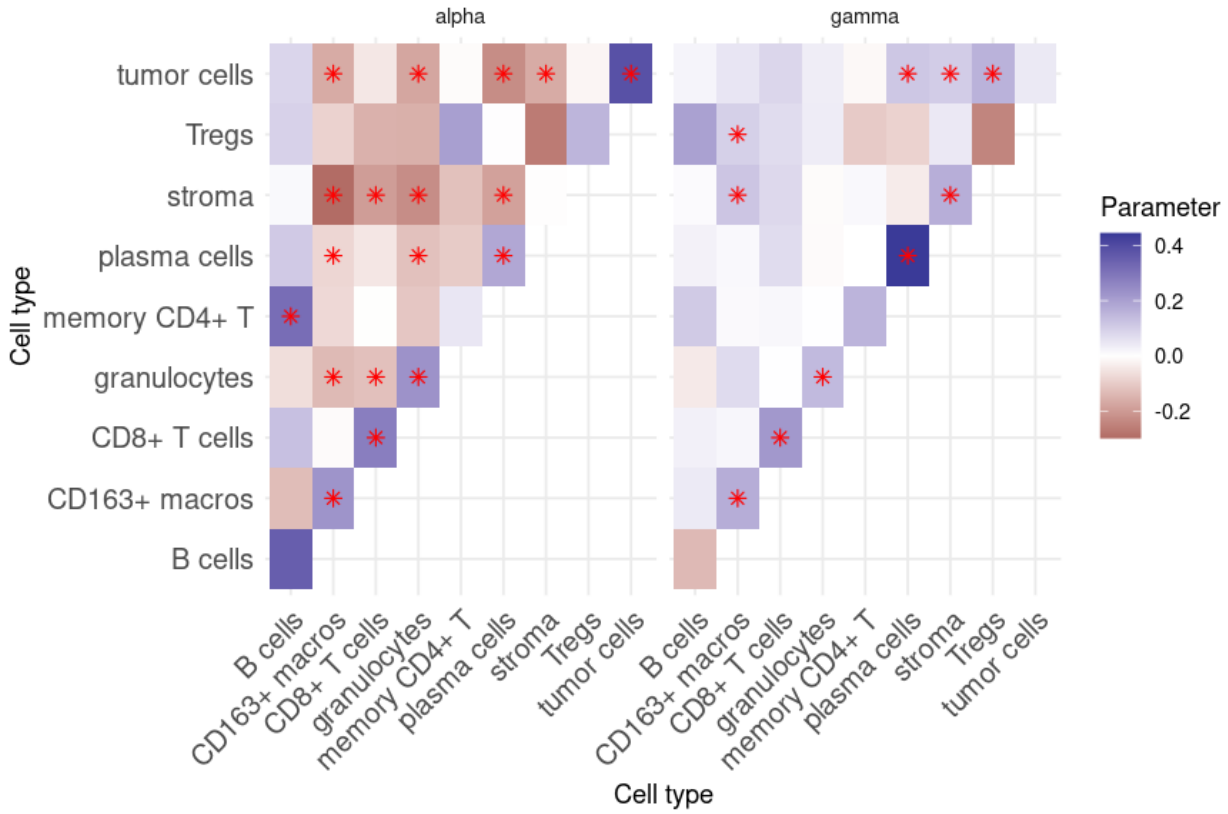


Figure 3.4: Interaction parameters α_{m_i,m_j} and γ_{m_i,m_j} as estimated from spot 59_A. Stars indicate that the interaction parameter was found to be significant within that image at a significance level of 0.05.

cells and CD163+ macrophages (or M2 macrophages) are generally positively associated with other cell types.

Heatmaps of the interaction parameters, as in Figure 3.4, allow one to quickly answer questions such as Question 1. From these heatmaps, we can also quickly answer questions like Question 1b: here, we can see that tumor cells seem to be quite strongly positively correlated with each other at short scales, while quite negatively correlated with plasma cells, also at short scales. Interestingly, at larger scales, tumor cells are no longer correlated with each other, and are now positively correlated with plasma cells!

One is likely also to want to compare how these interaction parameters contrast between various patient groups, in order to answer questions like Question 4. Figure 3.5 demonstrate the difference in the median interaction parameters α and γ from each patient group (averaging across all images within a group). Strikingly, it appears that both the α and γ parameters for smooth muscle generally have the strongest differences in median interaction parameter, with patients from the CLR group generally exhibiting greater attraction

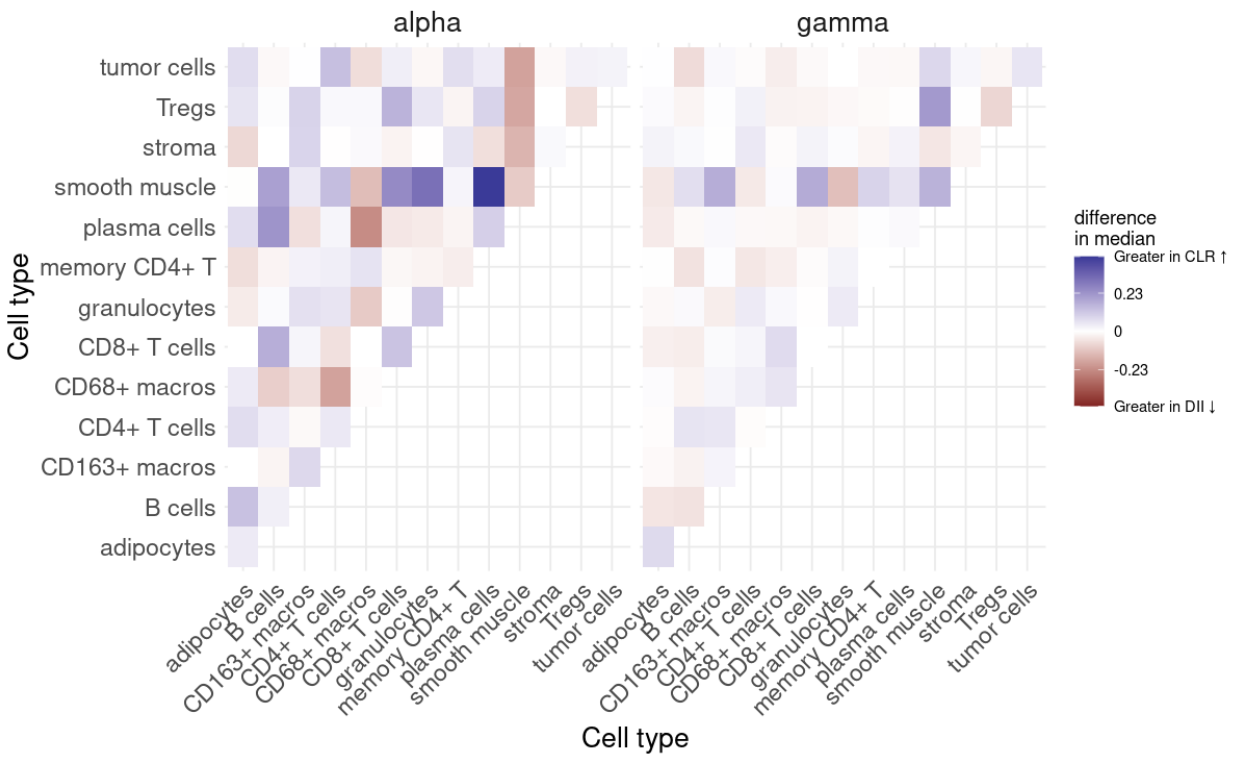


Figure 3.5: Difference in the median of the short-range interaction parameters α_{m_i,m_j} and γ_{m_i,m_j} between patient groups.

between smooth muscle and Tregs, CD8+ T cells and CD163+ macrophages, while smooth muscle and granulocytes exhibit greater repulsion in CLR patients.

Since the interaction parameters α and γ scale the potential functions used during the model fit (here, the exponential function), in order to answer questions liked Question 1a, it is also useful to view the pairwise potential functions themselves after they have been scaled by the appropriate parameters, as well as the combined effect of the short-range and medium-range scaled potential functions. An example of this can be seen in Figure 3.6, where the short-range potential (scaled by the appropriate α), medium-range potential (scaled by γ), and combined (sum of scaled short- and medium-range potentials) can be seen. The combined (or overall) potential thus gives a more complete picture between two cell types across spatial scales. Here we can see that B cells and memory CD4+ T cells are quite mutually attractive at shorter scales, but much less so at longer spatial scales. Such potentials can also be plotted simultaneously between all cell types of interest, as in Figure 3.7.

From both of these figures, it is also apparent that there are many significant interactions occurring between these cell types, at multiple scales. These interactions indicate the spatial

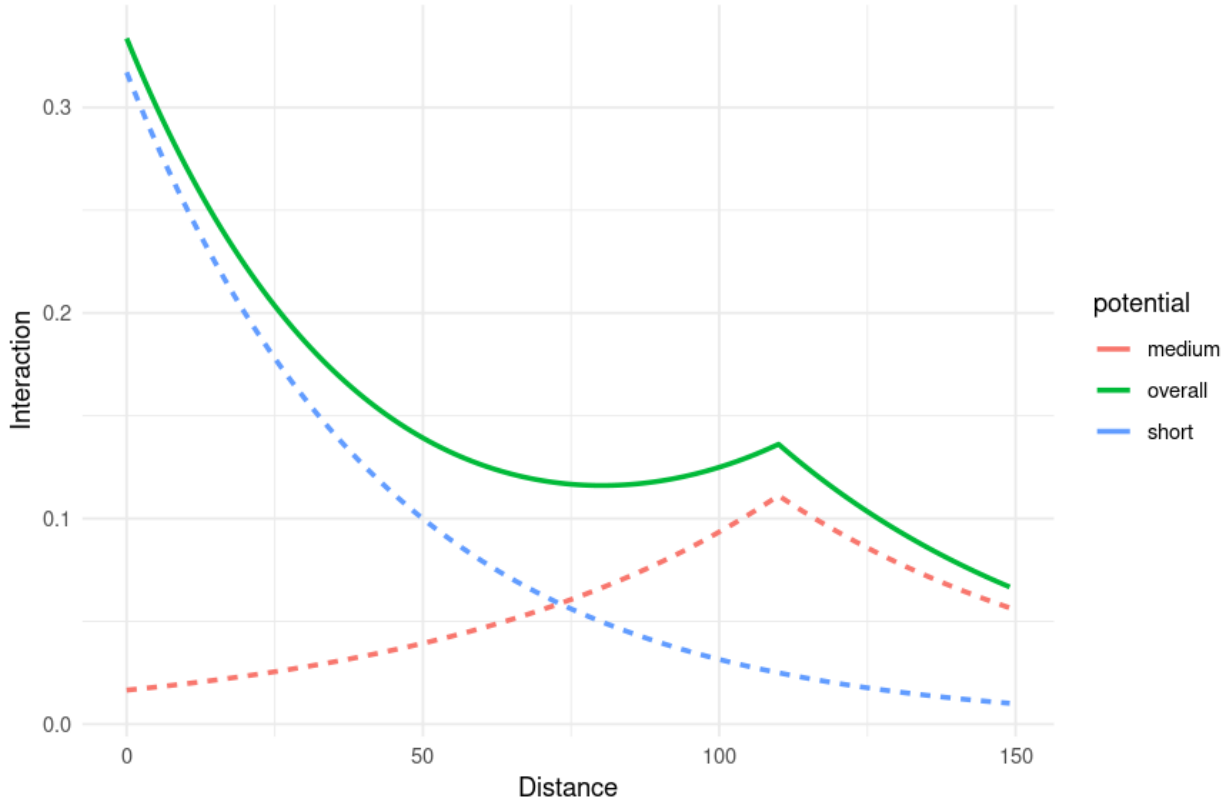


Figure 3.6: The short-range, medium-range and combined potentials between B cells and memory CD4+ T cells. From model fit on spot 59_A.

covariance between pairs of cell types that is unique to that pair. These quantities thus estimate the degree to which one cell type uniquely predicts the concentration of another, after accounting for the contributions to the prediction from all other cell types. From this information, we are thus able to assess which cell type most predicts the intensity of another cell type at each point in space, as in Figure 3.8. Such information allows one to answer questions like Question 2. Here we can see that different cell types are correlated more strongly with tumor cells at different points in the image. For example, granulocytes are the best predictors of tumor cells on the right side of the image, whereas CD163+ macrophages and stromal cells are the best predictors in other regions of the image.

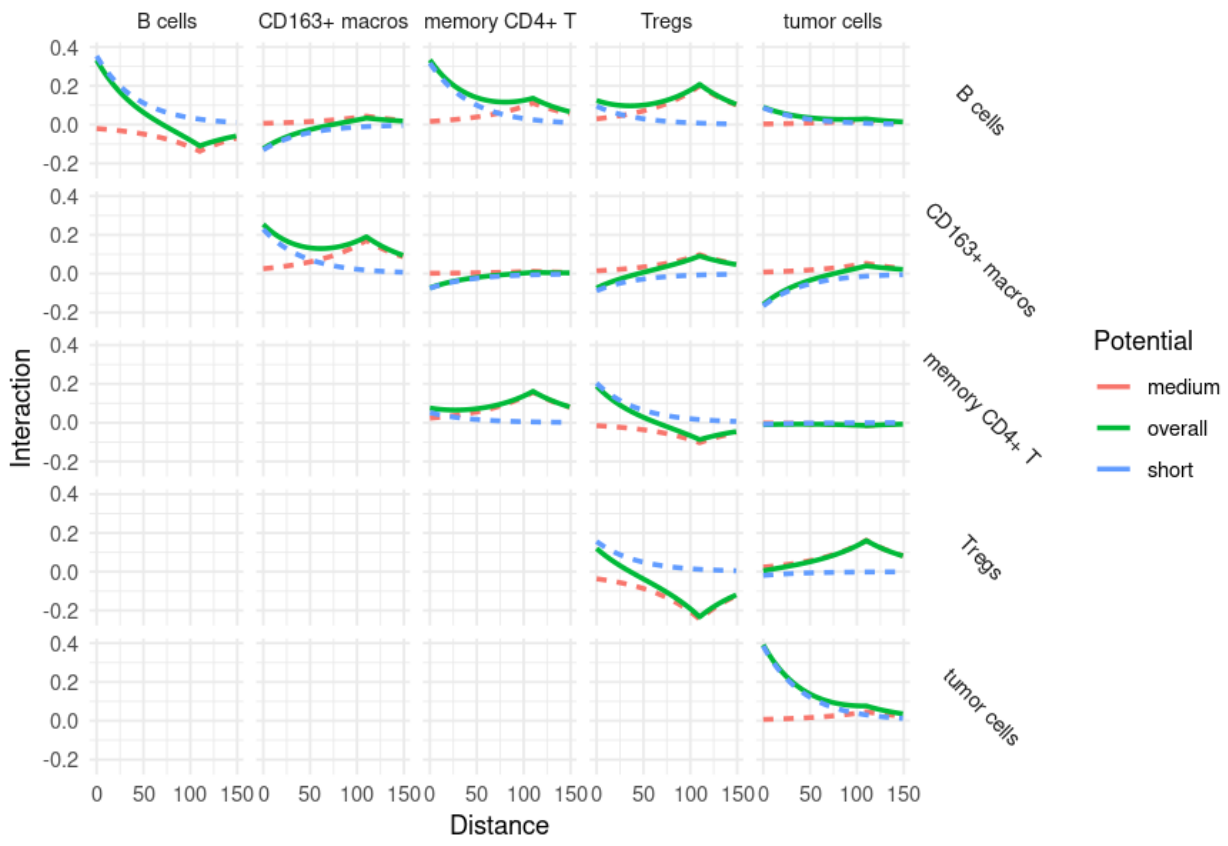


Figure 3.7: The short-range, medium-range and combined potentials between a subset of cell types. From model fit on spot 59_A.

3.3.2 Heterogeneity in Model Parameter Estimates Within Patient Groups and Across Patients

Estimates of the median absolute deviation of the short-range interaction parameters α can be seen in Figures 3.9 and 3.10, allowing us to answer Question 5. In Figure 3.9, we can see that, across the entire patient cohort, within-type interactions for several of the immune cell types (in particular, B cells, CD163+ macros, CD4+ T cells, CD68+ macros and CD8+ T cells) are the most heterogeneous between images, while the interaction between smooth muscle and CD4+ T cells appears to be the most consistent across images.

In Figure 3.10, it appears that DII patients generally have more homogeneous interactions between images. However, interactions involving smooth muscle in CLR patients appear to be the most homogeneous among CLR patients, sometimes more so than their DII counterpart (for example, the interaction between smooth muscle and stroma). Interestingly, interactions involving smooth muscle also seem most different between the two patient groups, as detailed in the previous section.

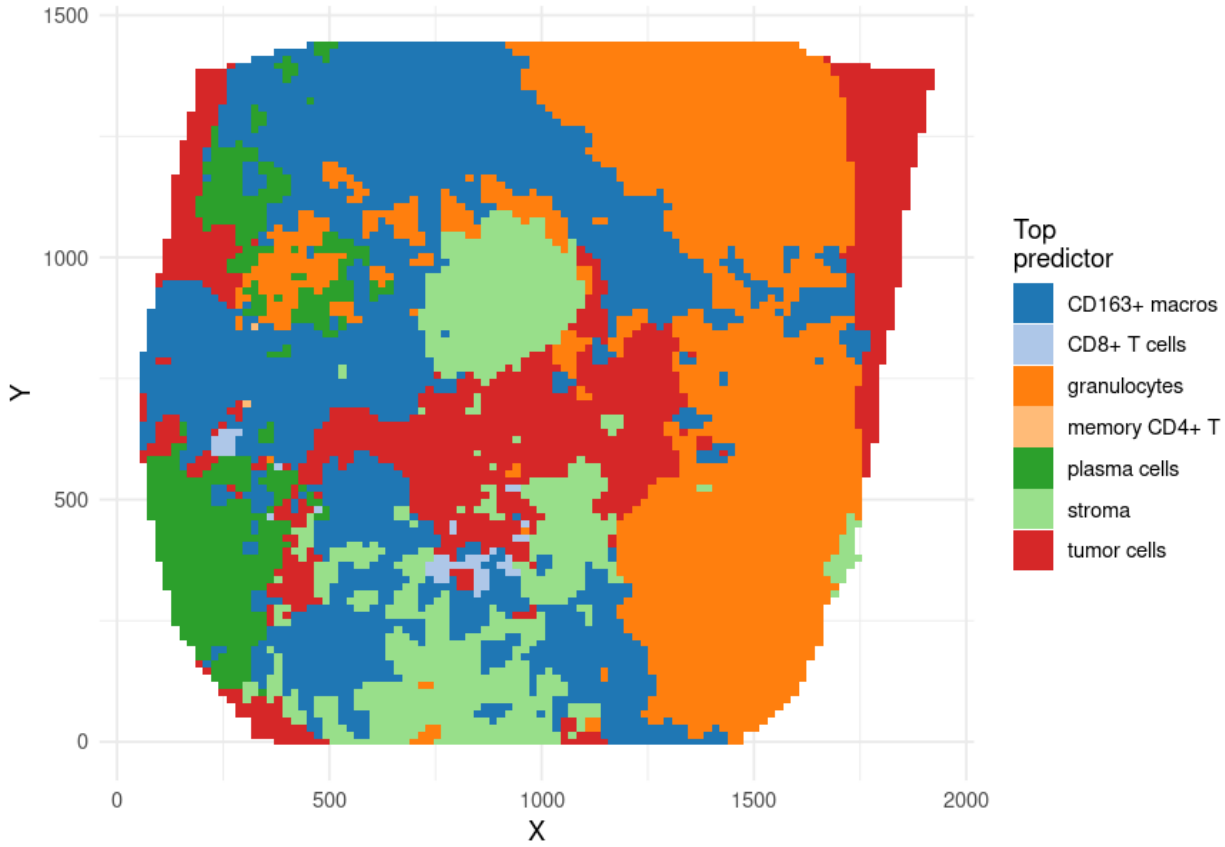


Figure 3.8: The top contributor to the predicted intensity of tumor cells. From model fit on spot 59_A.

3.3.3 Predictive Modeling of Patient Outcomes

Next, we turn our attention to answering Question 6. Estimated hazard ratios for both short- and medium-range interactions that were found to be significant after FDR adjustment for multiple hypothesis testing are shown in Figure 3.11. Conspicuously, increased short-range interactions between granulocytes and stroma seemed to be correlated with better patient outcomes, as well as granulocytes and M1 macrophages (CD68+ macros). Furthermore, the short-range interaction between M2 macrophages (CD163+ macrophages) and M1 macrophages seems to be indicative of poor patient outcome.

At longer spatial scales (γ parameters from Figure 3.11), the Treg and M1 macrophage interaction seems to be highly correlated with poor patient outcome, while the plasma cell - B cell and CD4+ - CD4+ interactions were associated with better patient outcomes.

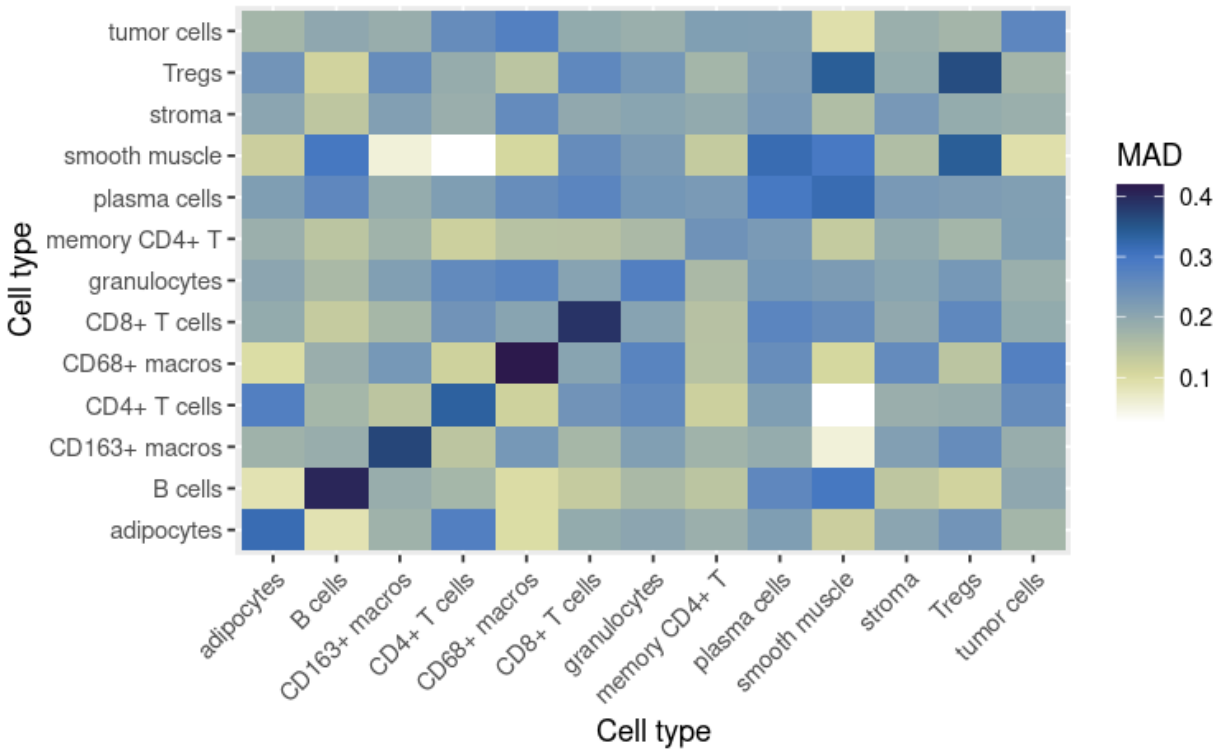


Figure 3.9: Estimates of the median absolute deviation (MAD) for α parameter estimates, across all patients.

3.3.4 Evaluation of Model Fit

Lastly, we focus on the issue of model trust, in an effort to answer Question 7. An example of the predicted conditional intensity for plasma cells can be seen in Figure 3.12. The AUC (0.835) for this cell type seems relatively high, indicating that the model seems to predict the intensity of plasma cells quite well. Additionally, the observed positions of the plasma cells seem to match higher values of the predicted conditional intensity, mirroring the information given by the AUC.

Table 3.3 shows the median AUC per cell type across all images, as well as the median AUC within each patient group. From this table, we can see that the cell types tend to be similarly predictable on average. Furthermore, there does not seem to be a significant difference in predictability between the patient groups, except, notably, for CD4+ T cells. At first, it may seem that this could be simply because of the much larger number of CD4+ T cells in CLR patients than in DII patients. However, as noted in Table 3.4, B cells are represented almost 4 times more in CLR patients than DII patients, a larger order difference than found in CD4+ T cells, while maintaining median AUCs that are approximately the same.

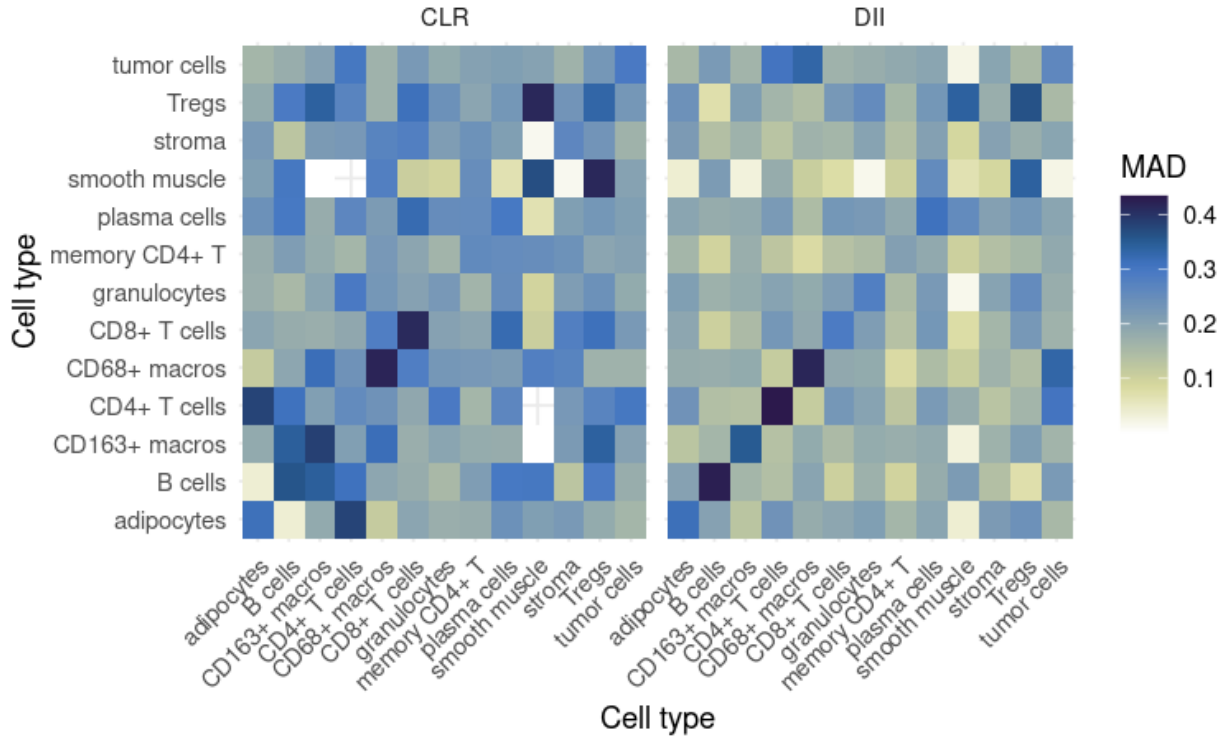


Figure 3.10: Estimates of the median absolute deviation (MAD) for α parameter estimates, in both patient groups.

To make the aforementioned precise, we measured the correlation between the AUC and abundance of each cell type. Table 3.5 shows the correlation between cell type and abundance, as quantified using AUCs from all model fits. All correlations can be seen to be negative, some with a quite large magnitude. This indicates that simply having more cells of a given type in an image does not make that cell type necessarily easier to predict using MGPP models. This likely indicates the importance of incorporating higher-order information on cell-type interactions, beyond overall cell-type abundance, in order to correctly predict distributions of cells.

Overall, from Table 3.3, we can see that the models predict the spatial distribution of cells quite well, indicating that we can likely trust the results from our models.

3.3.5 Comparison with Previous Analysis

Since the CRC dataset used here comes from a previously published study (Schürch et al., 2020), here we outline a few key points in comparison to the results from that study.

1. There is evidence of greater attraction between tumor cells and some subsets of immune

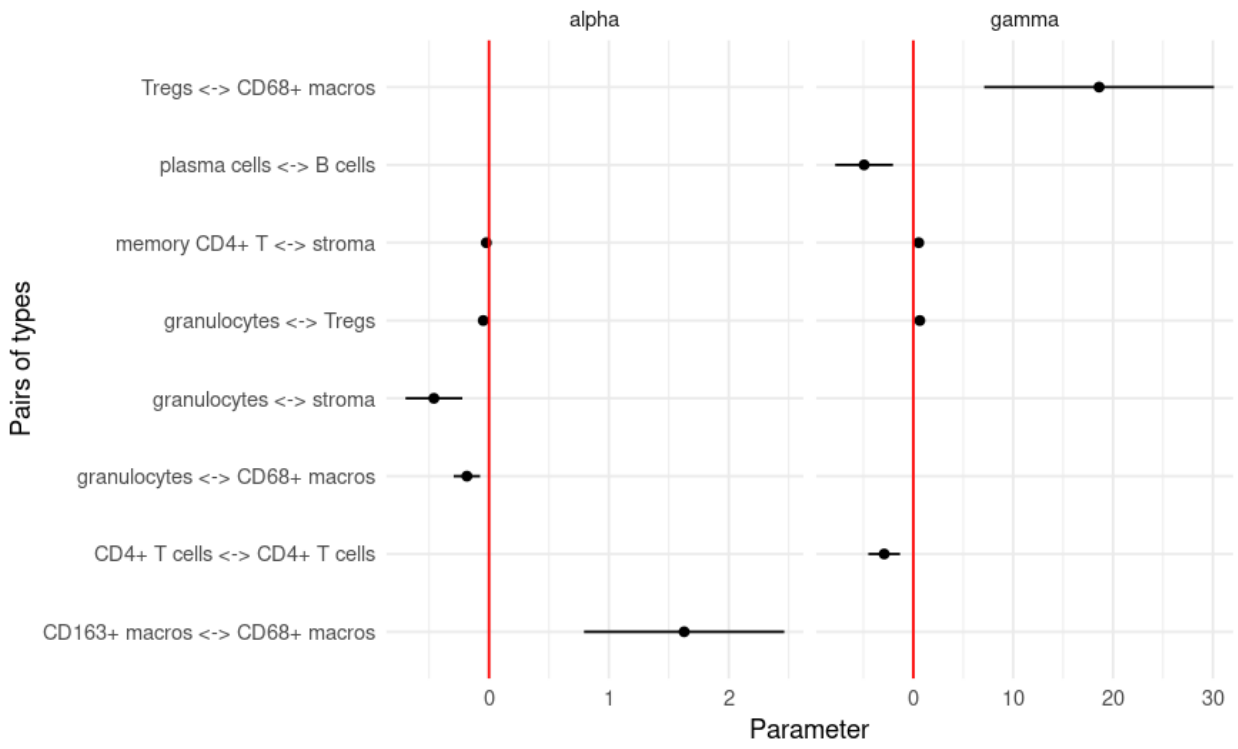


Figure 3.11: Log-hazard ratios estimated from short- and medium-range interaction parameters that were found to be significant after FDR adjustment.

cells in DII patients - for example, B cells and M1 macrophages - than CLR patients - similarly, (Schürch et al., 2020) identified the presence of a mixed tumor/immune compartment in DII patients, as opposed to CLR patients, in whom the tumor and immune components were more segregated. Interestingly, though, at short ranges, CD4+ T cells and tumor cells are found to be more attractive in CLR patients (Figure 3.5).

2. Schurch et al found much greater contact between a T-cell enriched neighborhood and a macrophage-enriched neighborhood in DII patients. Similarly, we found that M1 macrophages were more attractive with CD4+ T cells and plasma cells in DII patients at short ranges (Figure 3.5).

3. The authors further identified that the functional state of a granulocyte-enriched cellular neighborhood was indicative of patient survival in DII patients. Similarly, we also found that granulocyte associations were frequently implicated as the most prognostic biomarkers across patient groups derived from our models (Figure 3.11).

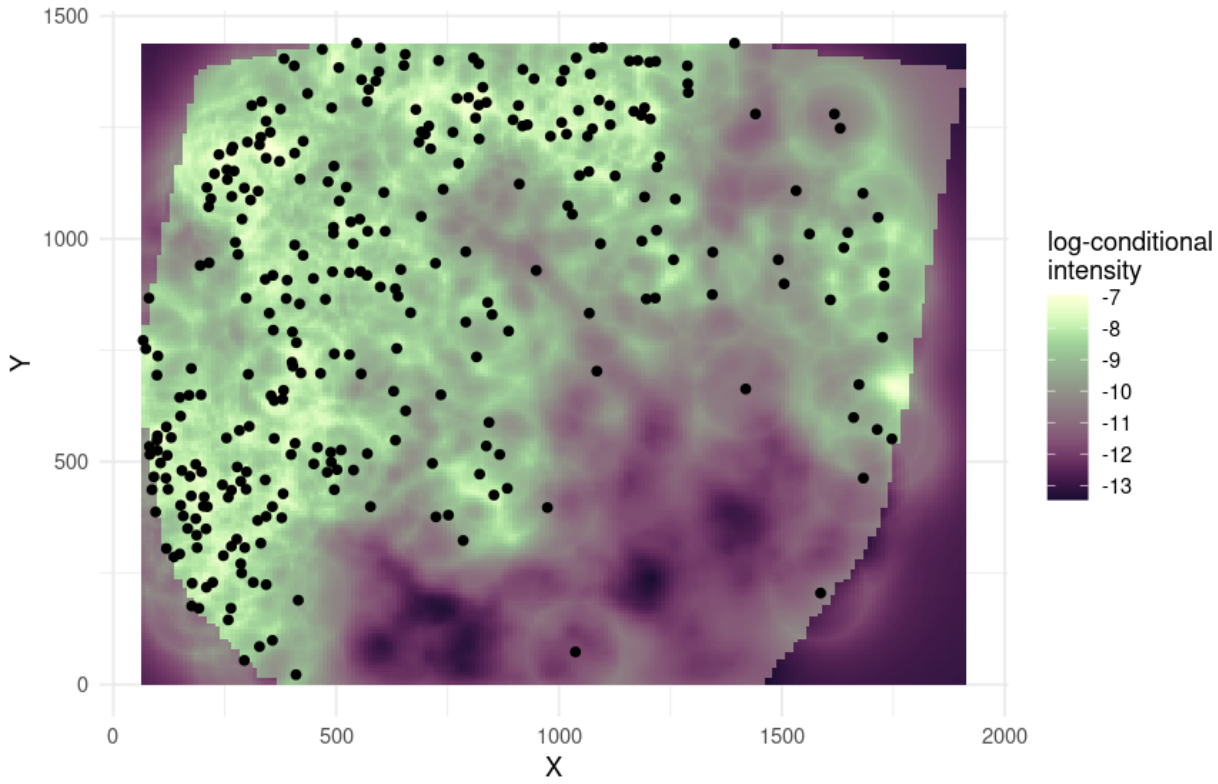


Figure 3.12: The predicted log-conditional intensity of plasma cells from the model fit for spot 59_A. The AUC for this cell type was 0.835, indicating high predictive ability for plasma cells.

3.4 Discussion

In this paper, we apply a novel approach for analyzing ecological interactions within the tumor microenvironment (TME). Specifically, we employ a multitype Gibbs point process model (MGPP) as a joint species distribution model, tailored to the complexities of the TME. This modeling framework was previously introduced as a statistical analysis framework in forest ecology (Flint et al., 2022). Our aim in this work is to highlight the significance of this modeling framework and its potential implications for addressing key questions in cancer biology.

Our primary contribution lies in the application of a specific MGPP model to characterize the data-generating process governing cell-cell and cell-environment interactions in the TME. Given the intricate nature of the TME, where diverse cell types interact spatially, the MGPP framework provides a powerful tool for capturing and quantifying these interactions. What sets this modeling framework apart from many others that currently exist is its ability to

Table 3.3: Median AUC for each cell type, across model fits from all images within each group and across the whole patient cohort

	AUC		
type	CLR	DII	All
B cells	0.83	0.82	0.82
CD163+ macros	0.72	0.69	0.70
CD4+ T cells	0.81	0.73	0.77
CD68+ macros	0.80	0.76	0.77
CD8+ T cells	0.77	0.74	0.75
NK cells	NA	0.72	0.72
Tregs	0.80	0.77	0.78
adipocytes	0.80	0.78	0.80
generic immune	0.76	0.76	0.76
granulocytes	0.77	0.73	0.75
memory CD4+ T	0.75	0.73	0.74
plasma cells	0.79	0.80	0.79
smooth muscle	0.77	0.75	0.76
stroma	0.72	0.70	0.70
tumor cells	0.72	0.76	0.75

model relationships between cell types simultaneously, while also being able to quantify the uncertainty in these associations as well as produce predictions. Researchers currently lack such a unified statistical framework capable of jointly modeling interactions between cell types and spatial variables across multiple spatial scales. With a specific MGPP model, we showcase how to bridge this gap, enabling a diverse array of inquiries.

Our emphasis extends beyond model development to model validation and diagnostics. We highlight the utility of our fitted models in generating predictions in order to examine goodness-of-fit and demonstrate the area under the curve (AUC) as a concise summary of model fit. Additionally, we stress the availability of further model diagnostics based on residuals.

Beyond model fitting, our study demonstrates the utility of extracting estimated parameters from single-image models and summarizing them across patient cohorts. This approach can serve as a valuable tool for identifying potential prognostic biomarkers, furthering our understanding of cancer biology and personalized medicine.

While hyperparameter selection is a critical aspect of model development, we do not delve deeply into this topic here. Interested readers can find comprehensive discussions on the selection of hyperparameters, including the scale and shape of potential functions, the

Table 3.4: Abundance of each cell type in each patient group.

	abundance	
type	CLR	DII
adipocytes	8.5	4.0
B cells	14.0	17.0
CD163+ macros	172.5	316.5
CD4+ T cells	5.5	2.0
CD68+ macros	9.5	20.0
CD8+ T cells	62.0	109.0
generic immune	12.0	9.0
granulocytes	27.0	123.5
memory CD4+ T	53.0	80.0
NK cells	1.5	1.0
plasma cells	28.0	22.0
smooth muscle	77.0	124.0
stroma	104.0	157.5
Tregs	8.0	21.0
tumor cells	102.5	222.0
vasculature	61.5	75.0

saturation parameter and the number of dummy points, in (Flint et al., 2022).

Lastly, MGPP models offer a unique capability which we also did not address in this paper: the generation of simulations while conditioning on various aspects of the data. This feature has the potential to serve two primary purposes—validating model outputs by comparing them to expected results and simulating new cell types based on different configurations. Such simulations can shed light on the spatial organization of cell types within the TME.

In conclusion, our investigation into cell-type interactions in the tumor microenvironment using multitype Gibbs point process models presents a valuable tool for researchers interested in the complex ecology of the TME. This framework empowers researchers to explore complex questions, validate models, and gain deeper insights into the dynamics of the TME, with potential implications for cancer research and therapy.

Table 3.5: Correlation between AUC score and abundance for each cell type.

type	correlation
B cells	-0.18
CD163+ macros	-0.39
CD4+ T cells	-0.30
CD68+ macros	-0.41
CD8+ T cells	-0.49
NK cells	-0.99
Tregs	-0.35
adipocytes	-0.42
generic immune	-0.59
granulocytes	-0.54
memory CD4+ T	-0.29
plasma cells	-0.41
smooth muscle	-0.58
stroma	-0.49
tumor cells	-0.16

CHAPTER 4

The SHADE Model for Asymmetric Multilevel Spatial Interactions

4.1 Introduction

Spatial dependencies between cell types play a central role in immune dynamics, tumor behavior, and tissue organization, motivating statistical models that can capture such interactions (Y. Yuan, 2016; Maley et al., 2017). The TME is a spatially structured system where cell arrangements are closely linked to disease progression and treatment response. Notably, these spatial interactions are frequently asymmetric: the spatial pattern of one cell type may be associated with that of another without necessarily implying reciprocity. For example, anti-tumor immune cells often cluster near tumor cells as part of immune surveillance, while tumor positioning is largely determined by vasculature and tissue architecture rather than by the presence of immune cells (Binnewies et al., 2018; Gilkes, Semenza, and Wirtz, 2014; Bindea et al., 2013). Quantifying such asymmetric associations is important for understanding tissue-level biological processes.

Recent advances in multiplexed imaging technologies, including mIF and other high-resolution spatial profiling methods, have made it possible to quantify cell type spatial distributions and interactions within the TME at single-cell resolution (Sheng et al., 2023). However, conventional spatial statistics, such as Ripley's K-function and pair correlation functions, primarily describe co-location patterns and do not explicitly model conditional relationships. Although Gibbs point process models offer a more flexible probabilistic framework (Moller and Rasmus Plenge Waagepetersen, 2003; A. Baddeley, Rubak, and Turner, 2015), standard formulations assume symmetric interactions, thereby overlooking the possibility that one cell type may be associated with a directional spatial arrangement of another. While hierarchical Gibbs models have been proposed to accommodate directionality (HöUgmander and SäUrkkä, 1999; Grabarnik and Särkkä, 2009), these methods often depend on parametric interaction functions that impose restrictive assumptions.

Observations based on the G-cross function suggest that spatial interactions between cell types are inherently asymmetric, with the inferred dependence structure varying by the choice of reference (Tsang et al., 2024). These findings motivate statistical models that directly account for directional spatial effects specified by biological hypotheses.

In this work, we develop a flexible statistical framework for modeling asymmetric spatial associations in tissue microenvironments, making it well-suited for the analysis of multiplexed imaging datasets. Our approach extends hierarchical Gibbs models by incorporating *spatial interaction curves (SICs)*, which flexibly quantify how the presence of one cell type relates to the expected density of another across different spatial scales. This is achieved by modeling the *conditional intensity function*, which describes the expected density of a target cell type at any spatial location given the observed pattern of a reference cell type. Unlike traditional step-function interaction terms, we model SICs via a flexible spline basis, enabling smooth, data-driven representations of spatial associations without imposing arbitrary interaction radii or functional forms—while maintaining interpretability through an explicit model of how local cell configurations contribute to overall spatial organization.

Beyond asymmetry, our method accounts for multilevel variation across images, patients, and cohorts. Biological spatial patterns vary at multiple levels, yet existing spatial models typically analyze images independently, potentially overlooking important heterogeneity across patients. We integrate the estimation of SICs into a multilevel Bayesian framework, which enables partial pooling across biological scales while preserving biologically meaningful variation.

To ensure computational efficiency, we approximate conditional intensity estimation using logistic regression with quadrature-based dummy points rather than direct Poisson likelihood estimation (A. Baddeley, Coeurjolly, et al., 2014). This transformation avoids computational challenges such as instability in fine-grained quadrature approximations and biases inherent in conventional spatial regression techniques.

We refer to this overall modeling framework as SHADE (Spatial Hierarchical Asymmetry via Directional Estimation). SHADE combines asymmetric spatial association modeling with hierarchical Bayesian inference to capture complex spatial organization across biological scales. An overview of the SHADE workflow is provided in Figure 4.1.

Our approach builds on existing spatial point process models while addressing several key limitations in the analysis of tissue microenvironments. First, we introduce a conditional intensity modeling framework that explicitly captures directional spatial associations between cell types, enabling the study of asymmetric interactions. Second, we develop a nonparametric representation of spatial interactions through spatial interaction curves, which provide a flexible and interpretable summary of how cell types influence one another as a function

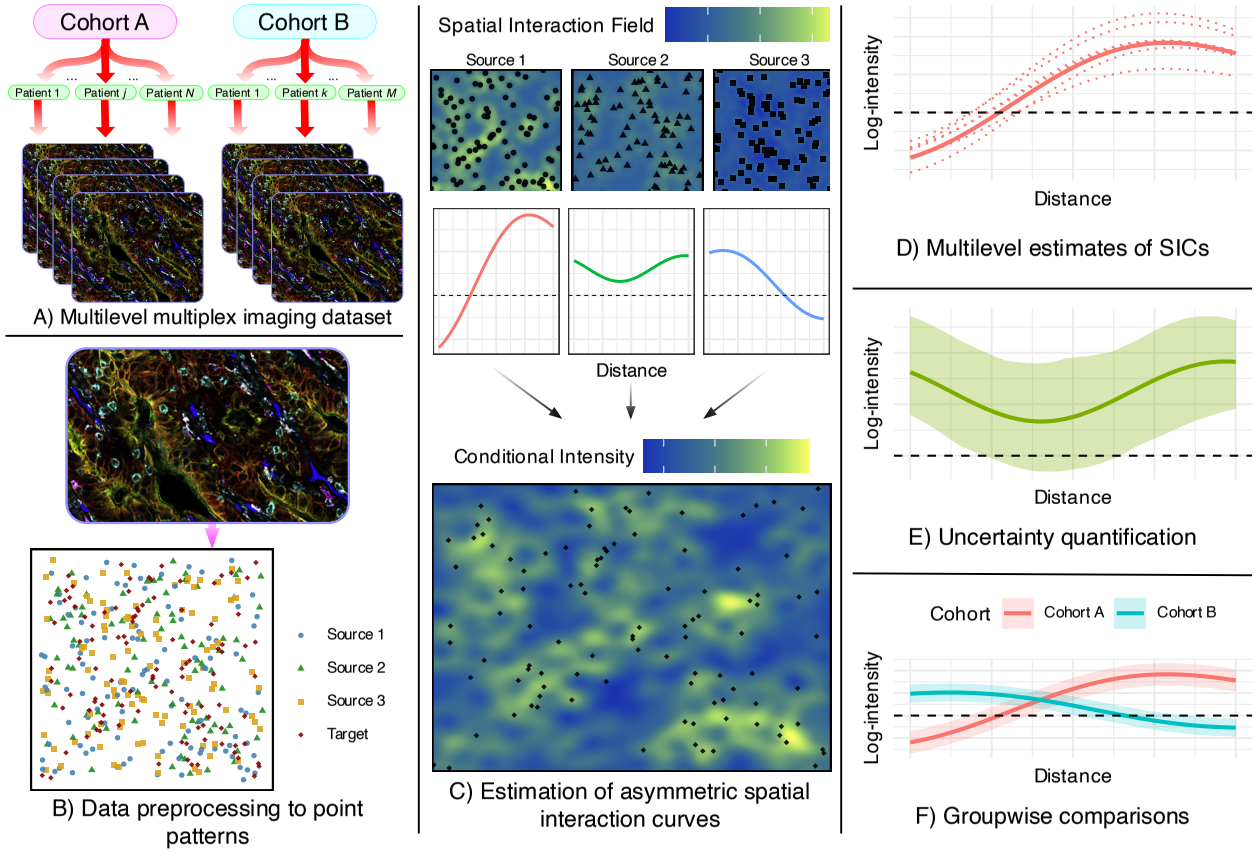


Figure 4.1: Summary of the SHADE framework. **A)** Multiplexed imaging data is structured hierarchically across cohorts, patients, and images. Multiplex images were adapted from Schurch et al. (2020) and are used under a Creative Commons CC BY 4.0 license. **B)** Images are processed into spatial point patterns with cell type annotations. **C)** SHADE estimates Spatial Interaction Curves (SICs) that capture directional associations between cell types across spatial scales. **D)** SICs are estimated at cohort, patient, and image levels, enabling multilevel analysis of spatial heterogeneity. **E)** Posterior distributions provide uncertainty quantification. **F)** SICs can be compared across cohorts to assess differences in spatial organization.

of distance. Third, we incorporate a multilevel Bayesian structure that models variation in spatial interactions across images, patients, and cohorts, allowing for partial pooling and uncertainty quantification at each level. Finally, we adopt a logistic regression approximation to the Poisson likelihood, which enables scalable and numerically stable inference without the need for fine spatial discretization.

4.2 Methods

4.2.1 Hierarchical Modeling of Conditional Spatial Point Processes

We model the spatial distribution of a target cell type B given the presence of one or more conditioning cell types $A_1, A_2, \ldots, A_K$, using a hierarchical framework based on conditional spatial point processes. Our approach extends hierarchical Gibbs models (HöUgmander and SäUrkkä, 1999; Grabarnik and Särkkä, 2009) and is designed to flexibly estimate asymmetric spatial association patterns while accounting for multilevel variation across images, patients, and cohorts.

Let $X_{A_k}, X_B \subset W$ denote the observed spatial point patterns of cell types A_k and B, respectively, within a two-dimensional tissue region $W \subset \mathbb{R}^2$. Formally, $X_{A_k} = \{x_i^{(k)} \in W \mid i = 1, \ldots, N_{A_k}\}$, $\quad X_B = \{y_j \in W \mid j = 1, \ldots, N_B\}$, where $x_i^{(k)} = (x_{i1}^{(k)}, x_{i2}^{(k)}) \in \mathbb{R}^2$ denotes the two-dimensional spatial coordinate of the i-th cell of type A_k, and $y_j = (y_{j1}, y_{j2}) \in \mathbb{R}^2$ denotes the coordinate of the j-th cell of type B. The quantities N_{A_k} and N_B indicate the total number of observed cells of type A_k and B, respectively. The observation window W corresponds to the region of tissue captured in the image, typically a rectangular subset of the plane defined by the image dimensions.

In practice, X_{A_k} and X_B are obtained from image segmentation and cell type classification pipelines applied to high-resolution tissue images, such as those generated by multiplexed imaging platforms.

We model the spatial distribution of X_B (the *target* cell type, e.g., immune cells) conditional on $X_{A_1}, \ldots, X_{A_K}$ (the *source* cell types, e.g., tumor cells and vasculature) by assuming that $X_B \mid X_{A_1}, \ldots, X_{A_K}$ follows an inhomogeneous Poisson point process (A. Baddeley, Rubak, and Turner, 2015), the likelihood of which we write as:

$$L(X_B \mid X_{A_1}, \ldots, X_{A_K}) = \left[\prod_{v \in X_B} \lambda(v \mid X_{A_1}, \ldots, X_{A_K}) \right] \times \exp\left(- \int_W \lambda(v \mid X_{A_1}, \ldots, X_{A_K})\, dv \right), \tag{4.1}$$

where the product runs over all observed locations of type B, while the integral accounts for the total expected intensity over the observation window W. The point process for X_B is entirely characterized by its *conditional intensity function* $\lambda(v \mid X_{A_1}, \ldots, X_{A_K})$, which defines the expected local density of type B cells at any location $v \in W$, conditioned on the spatial configurations of the conditioning cell types, that is, $\lambda(v \mid X_{A_1}, \ldots, X_{A_K}) = \lim_{|dv|\to 0} \frac{1}{|dv|} E[N_B(dv) \mid X_{A_1}, \ldots, X_{A_K}]$.

We model $\lambda(\cdot)$ as depending log-linearly on the observed patterns of $A_1, \ldots, A_K$ cells:

$$\log \lambda(v \mid X_{A_1}, \ldots, X_{A_K}) = \boldsymbol{z}^\top(v)\boldsymbol{\beta} + \sum_{k=1}^{K} \boldsymbol{q}_{A_k}^\top(v)\boldsymbol{\psi}_{A_k} \tag{4.2}$$

In (4.2), the term $\boldsymbol{z}(v) \in \mathbb{R}^J$ is a vector of covariates at location v (including an intercept), with corresponding coefficients $\boldsymbol{\beta} \in \mathbb{R}^J$. The vector $\boldsymbol{q}_{A_k}(v) \in \mathbb{R}^P$ encodes spatial interaction features between type A_k and type B; its p-th element is defined as:

$$[\boldsymbol{q}_{A_k}(v)]_p = \sum_{x \in X_{A_k}} \phi_p\left(\operatorname{dist}(v, x)\right), \tag{4.3}$$

where $\phi_p(\cdot)$ is a basis function (e.g., a B-spline or Gaussian kernel) that modulates the influence of a type A_k cell at a given distance from v. Finally, $\boldsymbol{\psi}_{A_k} \in \mathbb{R}^P$ are the corresponding coefficients. This formulation flexibly captures how proximity to different cell types influences the expected density of type B cells across spatial scales. Our use of smooth interaction features in (4.3) reflects biologically realistic, distance-dependent associations. In doing so, we generalize traditional Gibbs process formulations by avoiding rigid parametric forms such as fixed interaction radii (Grabarnik and Särkkä, 2009), which fail to capture the continuous and often subtle variations observed in biological spatial interactions (A. Baddeley and Turner, 2005).

While the model specified in (4.2) provides a flexible representation of spatial interactions through basis expansions, directly interpreting the estimated coefficients $\boldsymbol{\psi}_{A_k}$ can be difficult. In particular, biological interest often centers on how the strength and direction of association between a conditioning cell type A_k and a target cell type B vary as a function of distance. To facilitate interpretable biological insights, we next define a derived quantity, the *spatial interaction curve* (SIC), which summarizes the estimated effect of proximity to A_k cells on the expected density of B cells across spatial scales.

4.2.1.1 Spatial Interaction Curve

The SIC summarizes the asymmetric spatial association between a conditioning cell type A_k and a target cell type B as a function of distance s:

$$\mathrm{SIC}_{A_k \to B}(s) = \sum_{p=1}^{P} \psi_{A_k}^{(p)} \phi_p(s), \tag{4.4}$$

This curve represents the expected contribution of type A_k cells to the log-intensity of type B cells as a function of distance s from a type A_k cell.

As illustrated in Figure 4.2, a positive SIC value at distance s indicates that the presence of a type A_k cell is associated with an increased expected density of type B cells at that radius, consistent with spatial attraction or clustering. Conversely, a negative SIC value suggests spatial repulsion or avoidance at that distance. This function provides an interpretable, distance-dependent summary of directional spatial association between cell types.

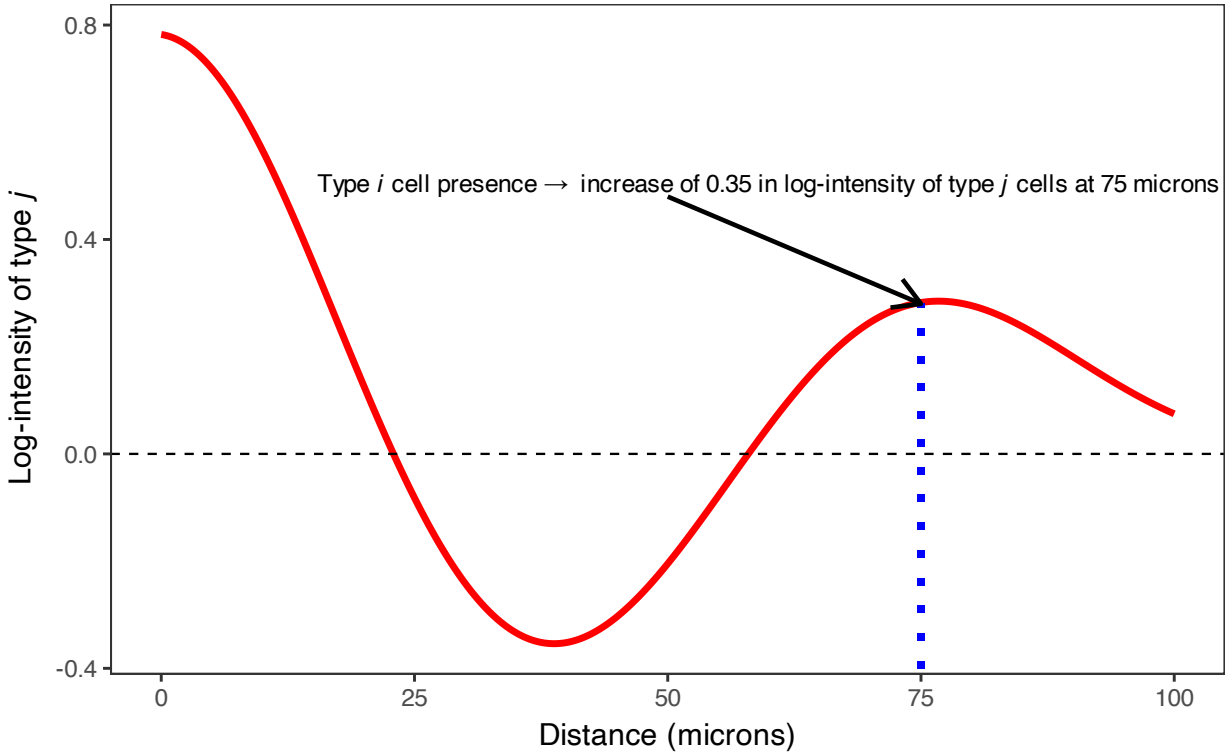


Figure 4.2: An example spatial interaction curve showing the effect of a source cell type A_k on a target cell type B. At each distance s, the curve value represents the change in log-intensity of type B associated with the presence of a type A_k cell at distance s.

Figure 4.3 shows two example SICs representative of those estimated from real data. Type 2 cells (cyan) exhibit a strong positive association at short distances, consistent with

clustering behavior observed in immune infiltration around tumor cells. In contrast, Type 1 cells (red) show negative association at close range, suggestive of exclusion zones, as might be induced by physical barriers or competitive interactions in the tissue microenvironment.

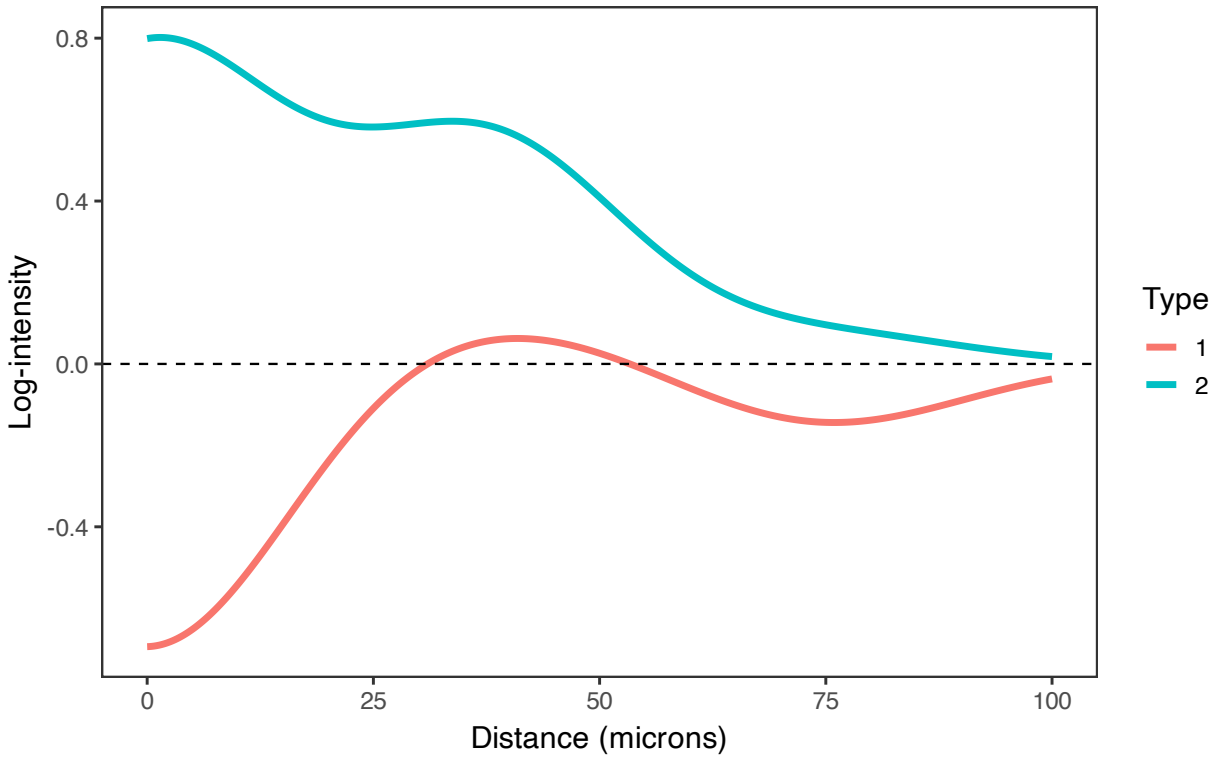


Figure 4.3: Estimated SICs for two source cell types. Type 2 shows a strong positive association with the target population at short distances, while Type 1 exhibits negative association at short range. These contrasting patterns highlight the flexibility of the SIC framework for capturing biologically meaningful spatial structure.

Figure 4.4 demonstrates how these spatial interactions manifest in tissue space. Panels 1 and 2 show the spatial predictors $\boldsymbol{q}_{A_1}(v)$ and $\boldsymbol{q}_{A_2}(v)$ generated by the individual SICs from Figure 4.3—each reflecting the localized contribution to the log-intensity of type B due to each source. Each panel is analogous to a summand from the sum on the right hand side of (4.2). Panel 3 shows the additive combination of these effects, illustrating how multiple source cell types jointly shape a heterogeneous spatial intensity landscape. The target cells are observed to cluster in high-intensity regions and avoid areas of low expected density.

A key advantage of our framework is that the spatial interaction terms $\boldsymbol{\psi}_{A_k}$ can take both positive and negative values, allowing for flexible modeling of attraction and repulsion. This mirrors the flexibility of hierarchical Gibbs models while overcoming constraints in symmetric pairwise interaction models, which typically restrict interaction terms to be non-positive and

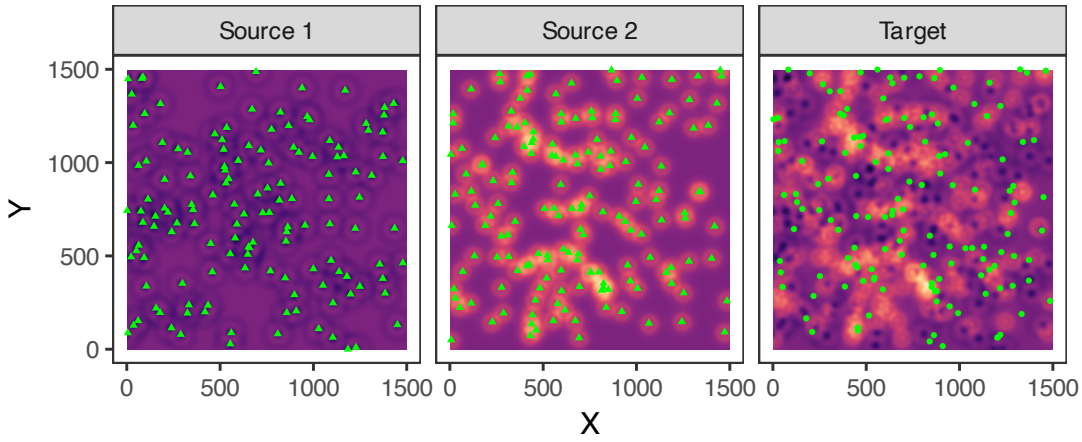


Figure 4.4: Spatial interaction fields implied by SICs from two source cell types (Panels 1 and 2) and their additive combination (Panel 3). The combined field governs the spatial distribution of target cells, which are more likely to occur in regions of elevated log-intensity.

therefore cannot directly model spatial attraction.

4.2.1.2 Characterizing Predictive Asymmetry

Our modeling framework captures *predictive* spatial associations, not biological causation. A strong $A \to B$ SIC indicates that the presence of A is statistically predictive of local B density, but it does not imply that A exerts a causal biological effect on B. Apparent directional associations may arise from confounding factors such as differences in the spatial distribution of A (e.g., sparser, more clustered, or more localized patterns), which can make A a stronger statistical predictor without implying direct biological influence.

To mitigate such confounding, the model incorporates spatial covariates or trend surfaces (e.g., distance to the tumor margin) to adjust for first-order intensity effects. This adjustment ensures that estimated directional SICs capture residual spatial association beyond shared background structure.

4.2.2 Multilevel Bayesian Model

To capture biological variability across individuals and sampling levels, we impose a hierarchical prior structure on the spatial interaction coefficients. For each conditioning cell type A_k, the interaction coefficients are indexed hierarchically across three levels: cohort (ψ), patient (γ), and image (δ). Specifically, for each basis function p, we define:

$$
\begin{aligned}
\psi_{A_k}^{(g,p)} &\sim \mathcal{N}(0, \sigma^2_{\text{cohort}}), \\
\gamma_{A_k}^{(n,p)} &\sim \mathcal{N}(\psi_{A_k}^{(g(n),p)}, \sigma^2_{\text{patient}}), \\
\delta_{A_k}^{(m,p)} &\sim \mathcal{N}(\gamma_{A_k}^{(n(m),p)}, \sigma^2_{\text{image}}),
\end{aligned}
\tag{4.5}
$$

where $g(n)$ maps patient n to its cohort, and $n(m)$ maps image m to its corresponding patient. Each image-level coefficient $\delta_{A_k}^{(m,p)}$ governs the localized effect of source cell type A_k on target cell type B at distance scale p, within image m.

This hierarchical formulation enables partial pooling across the dataset structure, improving estimation stability while preserving biologically meaningful heterogeneity in spatial interactions. By modeling variation at the cohort, patient, and image levels, the framework supports inference on both shared and context-specific spatial association patterns.

Hyperpriors for the variance components σ^2_{cohort}, $\sigma^2_{\text{patient}}$, and σ^2_{image} are detailed in the Supplement.

4.2.3 Logistic Regression Approximation for Computational Efficiency

Direct estimation of the Poisson likelihood in spatial point process models typically requires numerical integration over a fine spatial grid, which becomes computationally expensive and unstable in high-resolution images. To address this, we follow the logistic regression approximation introduced by A. Baddeley, Coeurjolly, et al. (2014), which avoids spatial gridding by introducing a set of dummy points D sampled from a homogeneous Poisson process with known intensity λ_{dummy}. We define the combined set of points $Y = X_B \cup D$, and assign binary labels:

$$
I(v) = \begin{cases} 1, & v \in X_B \\ 0, & v \in D \end{cases}
$$

The conditional probability that a point v is a true (observed) point of type B is given by:

$$
P(I(v) = 1) = \frac{\lambda(v)}{\lambda(v) + \lambda_{\text{dummy}}},
$$

leading to the following logistic regression model:

$$
\log \frac{P(I(v) = 1)}{P(I(v) = 0)} = \boldsymbol{z}^\top(v)\boldsymbol{\beta} + \sum_{k=1}^{K} \boldsymbol{q}_{A_k}^\top(v)\boldsymbol{\delta}_{A_k}^{(m)} - \log \lambda_{\text{dummy}},
\tag{4.6}
$$

This approximation circumvents key computational challenges of direct Poisson modeling. In Poisson models, fine spatial discretization leads to large numbers of empty pixels, often resulting in singular design matrices and unstable inference (A. Baddeley, Rubak, and Turner, 2015). The Hauck-Donner effect (Hauck and Donner, 1977) can further distort uncertainty estimates. By reframing the problem as a binary classification task over observed and dummy points, the logistic approximation enables scalable, stable inference of spatial interaction effects at the image level.

4.2.4 Summary of Estimation Procedure

Model fitting is implemented in `Stan` using Hamiltonian Monte Carlo (HMC) via `cmdstanr` (Stan Development Team, 2024; Gabry et al., 2024).

To approximate the likelihood, we first generate a set of dummy points D from a homogeneous Poisson process with intensity λ_{dummy} over the observation window W. For each location $v \in X_B \cup D$, we then compute spatial covariates $\boldsymbol{z}(v)$ and interaction features $\boldsymbol{q}_{A_k}(v)$, where each interaction feature encodes basis-function-weighted distances to cells of type A_k.

Using these constructed features, we fit a multilevel Bayesian logistic regression model based on the approximation in Equation (4.6), with spatial interaction coefficients $\boldsymbol{\delta}_{A_k}^{(m,p)}$ modeled hierarchically according to the priors in Equation (4.5). Finally, we extract posterior draws of the interaction coefficients and reconstruct the spatial interaction curves using Equation (4.4).

We provide an open-source implementation of the SHADE model at: github.com/jeliason/SHADE. All code used to generate the figures and analyses presented in this manuscript is available at: github.com/jeliason/shade_paper_code.

4.3 Simulation Study

To evaluate the proposed model, we conducted a simulation study that generates synthetic spatial point patterns incorporating both asymmetric interactions and multilevel structure. Spatial patterns are simulated within a bounded domain $W = [0, S] \times [0, S]$, where points of source types are generated using a homogeneous Poisson process with baseline intensity $\lambda_{\text{base}} = \frac{N_{\text{points}}}{S^2}$. Points of target types are generated with spatial dependencies modeled through the previously described spatial interaction curves, which are functions of the distances from the source cell types. To introduce hierarchical structure, the interaction coefficients governing the SICs are generated according to (4.5).

The final point pattern is transformed into a logistic regression dataset by introducing dummy points distributed according to a Poisson distribution with intensity λ_{dummy} and calculating spatial interactions as in (4.3). The full simulation procedure is described in detail in the Supplement.

4.3.1 Number of cells of each type and number of dummy points

We first investigate how inference quality is affected as the number of cells of each type N_t as well as the number of dummy points used for quadrature changes. We simulated 5 realizations each of every combination of N_t and the ratio of dummy points to real points R_d, where $N_t \in \{20, 80, 150, 300\}$ and $R_d \in \{0.5, 1, 2, 5, 10\}$, according to our simulation procedure described in the Supplement. We fixed the number of types $N_T = 3$, with the last cell type being the focal cell type that is target on the other cell types. We set the number of patient groups to 1, the number of patients per group to 40 and the number of images per patient to 2. The rest of the simulation parameters were kept at their defaults, as described in the Supplement.

We found that the average RMSE for the image-level coefficients $\delta^{(m,p)}_{t_1 \to t_2}$ decreased as the ratio of dummy points increased, regardless of the number of observed points per cell type. This suggests that using a finer quadrature grid improves the inference quality for image-level spatial interaction estimates (Figure A.1). However, we observed a different trend at higher levels of the hierarchy. In particular, the inference quality for patient-level coefficients $\gamma^{(n,p)}_{t_1 \to t_2}$ and cohort-level coefficients $\psi^{(g,p)}_{t_1 \to t_2}$ worsened beyond a dummy point ratio of approximately 2–5, especially in regimes with very low or very high numbers of observed cells. This indicates that excessively increasing the dummy point ratio may have diminishing or even adverse effects on inference for higher-level parameters, though the magnitude of these effects is relatively small in absolute terms (note the log-scale on the y-axis in Figure A.1).

We also examined how inference quality varied across spatial scales for each level of the hierarchical model (Figure 4.5). For image-level coefficients $\delta^{(m,p)}_{t_1 \to t_2}$, the average RMSE generally decreased with increasing dummy point ratio, particularly at larger spatial scales. In contrast, the inference quality for patient-level coefficients $\gamma^{(n,p)}_{t_1 \to t_2}$ remained relatively stable across dummy point ratios and spatial scales. For cohort-level coefficients $\psi^{(g,p)}_{t_1 \to t_2}$, inference was worst for short-range coefficients and showed increased RMSE at moderate dummy point ratios (approximately 2–5).

A representative example of an estimated cohort-level SIC is shown in Figure A.2, comparing the inferred SIC to the ground truth across various parameter configurations. Inference quality for this SIC was poor when the number of observed cells per type was low. The best

performance—reflected in both reduced bias and narrower credible intervals—occurred when the number of points per type was moderate (80–300) and a wide range of dummy points was used, consistent with the trends observed in the average RMSE plots (Figure A.1).

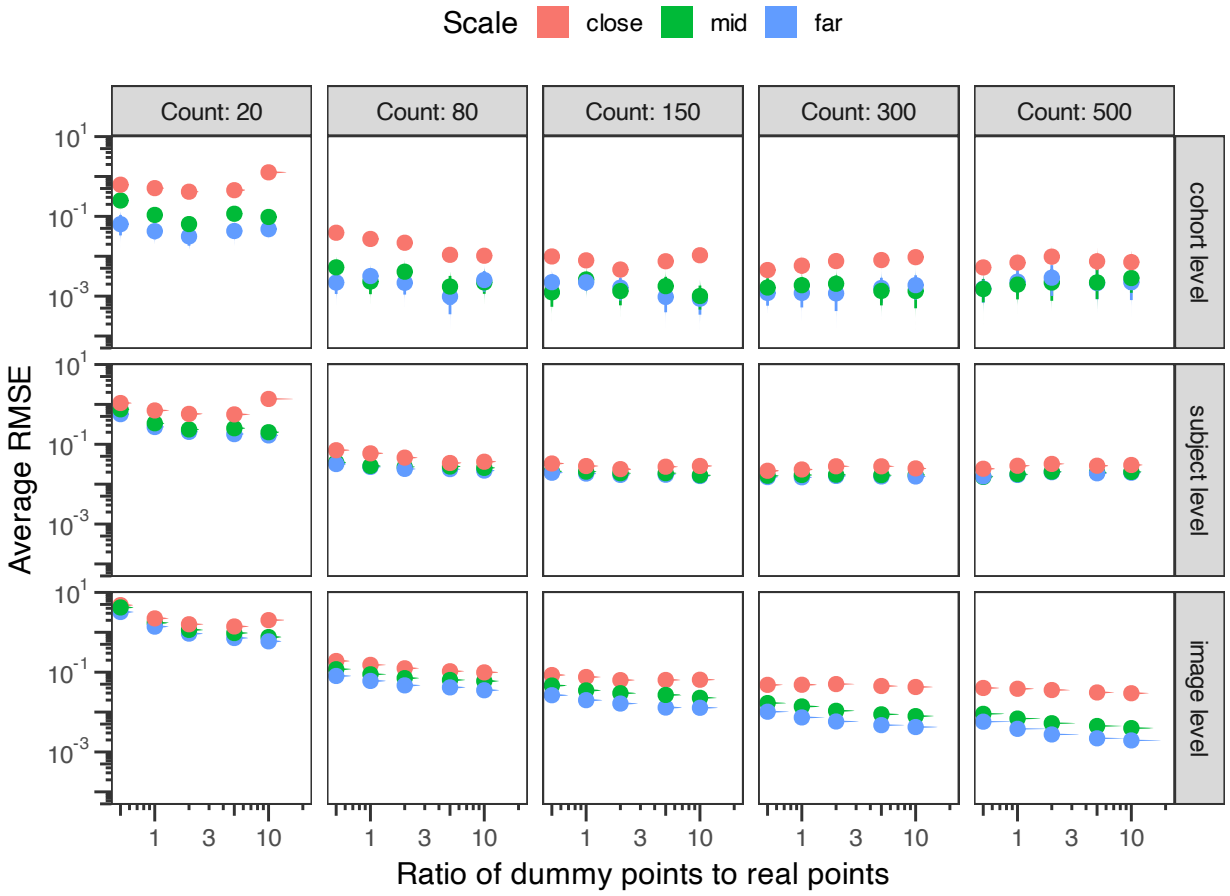


Figure 4.5: RMSE for spatial interaction coefficients across spatial scales, shown separately for each hierarchical level: $\delta^{(m,p)}_{t_1 \to t_2}$, $\gamma^{(n,p)}_{t_1 \to t_2}$, and $\psi^{(g,p)}_{t_1 \to t_2}$. Results are shown for varying cell counts and dummy point ratios.

4.3.2 Number of patients and images per patient

Next, we examined the sensitivity of inference quality to the total number of patients as well as the number of images per patient. Here, we set the number of points per type to 150 and the ratio of dummy points to actual points to 2, while keeping the rest of the default parameters the same, as outlined in the Supplement. We then simulated 15 realizations each of every combination of the number of images per patient $N_{\text{images per patient}}$ and the total number of patients N_{patients}, where $N_{\text{images per patient}} \in \{1, 2, 4\}$ and $N_{\text{patients}} \in \{10, 20, 40\}$,

according to our simulation procedure described in the Supplement.

We found some interesting trends in the average RMSE as the number of patients per patient group increased. For cohorts that only had one image per patient, average RMSE decreased as the number of patients increased (Figure A.3). However, for patients that had 2 images, the average RMSE was highest for the highest number of patients, which we found to be a counterintuitive finding. Furthermore, for patients with 4 images, having a cohort with 40 patients was associated with the lowest RMSE, as expected, though the second-lowest was, unexpectedly, having a cohort with 10 patients, rather than 20.

For $\delta_{t_1 \to t_2}^{(m,p)}$ coefficients, longer-range coefficients seemed to have the lowest average RMSE (Figure 4.6), while close-range coefficients had the worst quality. RMSE, however, was very close to constant across the number of patients, though it did decrease slightly as the number of images increased. RMSE for $\gamma_{t_1 \to t_2}^{(n,p)}$ coefficients decreased as the number of images increased, though, as we noticed earlier, the RMSE in this case has a nonlinear relationship with the number of patients. The RMSE of $\psi_{t_1 \to t_2}^{(g,p)}$ coefficients exhibited the same counterintuitive nonlinear association with number of patients, though in a more pronounced way.

Finally, we demonstrate an example of an estimated global SIC from this simulation study (Figure A.4). Here, we can see that the variance of the SIC estimate decreases both as the number of patients increases and the the number of images increases, which matches our expectations. However, we can also see that there is a slight amount of persistent bias at close range for simulations in which the number of patients equals 40 - this may be indicative of too much shrinkage during estimation and may be the reason for the nonlinear association of RMSE with number of patients.

4.3.3 Hierarchical vs non-hierarchical modeling

In our last simulation study, we wanted to examine the effect that accounting vs not accounting for the hierarchical structure of the dataset would have on inference quality. In this study, we simulated from 2 patient groups which had 20 patients each. Each patient had 4 images. We set the number of points of each type in each image to 150, with the ratio of dummy points to real points as 2. The rest of the simulation parameters were set to their defaults, as described in the Supplement. We simulated 100 separate realizations from these parameter settings, and fit two models to each realization: the full model that we consider here that takes into account the hierarchical structure of the data, and a model that fits image-level SICs individually, without any shrinkage towards individual or global SICs. We evaluate the inference quality of these two models on simulations via the average RMSE of

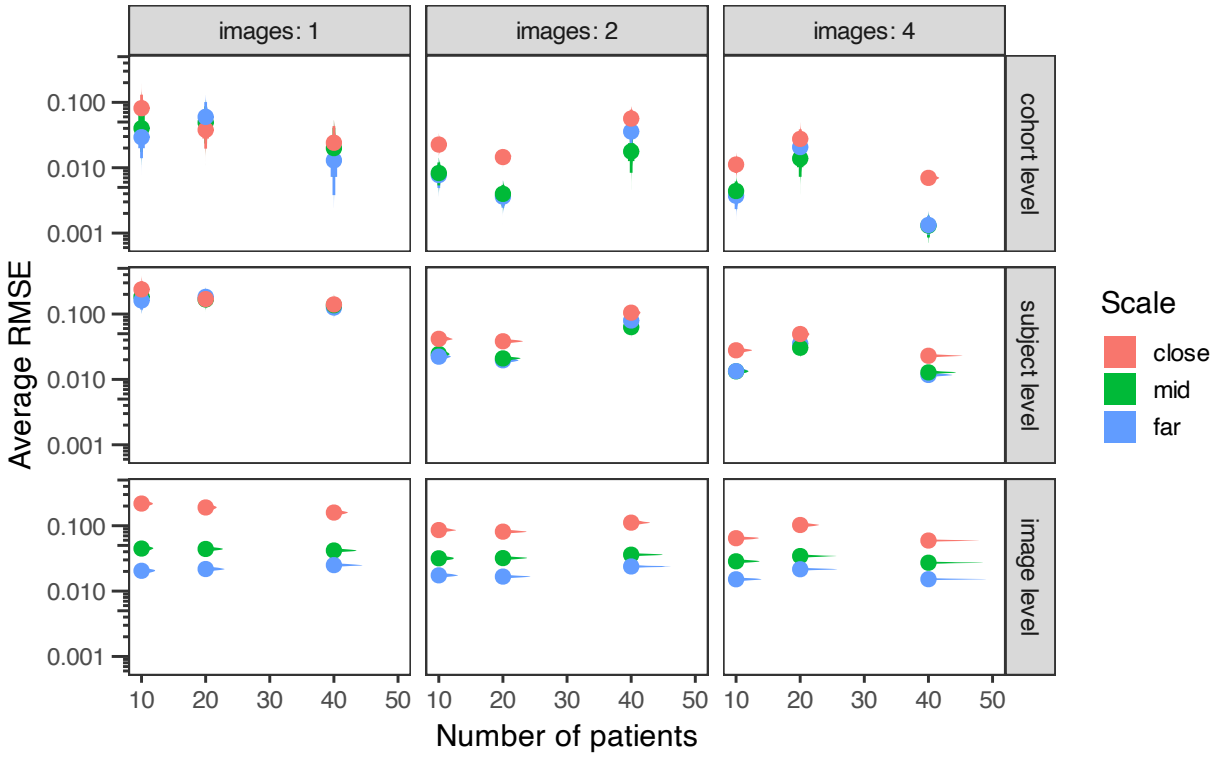


Figure 4.6: Average RMSE across spatial scales for spatial interaction coefficients at the image ($\delta^{(m,p)}_{t_1 \to t_2}$), patient ($\gamma^{(n,p)}_{t_1 \to t_2}$), and cohort ($\psi^{(g,p)}_{t_1 \to t_2}$) levels, across varying numbers of patients and images per patient.

the $\delta^{(m,p)}_{t_1 \to t_2}$ coefficients.

We found that on average, and across all spatial scales, the average RMSE of $\delta^{(m,p)}_{t_1 \to t_2}$ coefficients was lower for those estimated by the full hierarchical model than the model that fit images individually, most significantly so for closer-range coefficients (Table 4.1).

Hierarchical	**Scale: All**	**Scale: Small**	**Scale: Medium**	**Scale: Large**
False	0.142 (0.097, 0.196)	0.355 (0.230, 0.509)	0.022 (0.016, 0.031)	0.048 (0.034, 0.066)
True	0.039 (0.027, 0.043)	0.072 (0.043, 0.084)	0.017 (0.013, 0.020)	0.028 (0.021, 0.036)

Table 4.1: Average RMSE of $\delta^{(m,p)}_{t_1 \to t_2}$ coefficients in aggregate and across spatial scales, from both a model that fully accounts for the hierarchical structure as well as a model that does not account for hierarchy.

Figure 4.7 shows several examples of image-level SICs estimated with both models from this simulation study. In all cases, we can see that the SIC is better estimated when hierarchical structure is accounted for, in terms of both the bias and and variance of the estimate.

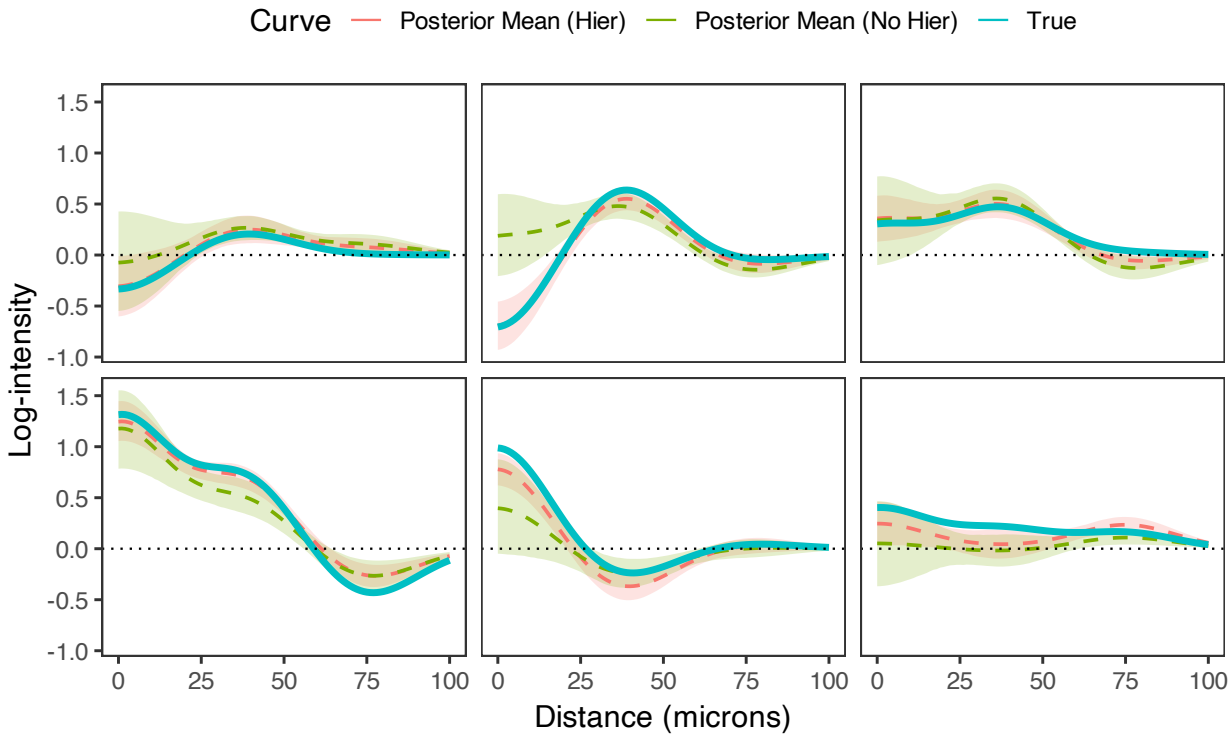


Figure 4.7: Examples of estimated image-level SICs from hierachical modeling simulations, demonstrating better estimation of SICs when accounting for multilevel structure.

4.4 Results: Multiscale inference of directional spatial interactions in CRC

4.4.1 Description of colorectal cancer dataset

We applied our hierarchical spatial interaction model to a publicly available colorectal cancer (CRC) dataset of multiplexed tumor tissue images from 35 patients. These images capture the spatial organization of malignant, immune, and stromal cell types, with high-dimensional protein expression measured across 56 markers. Patients were stratified into two groups based on immune architecture at the tumor invasive front: those with organized tertiary lymphoid structures (TLSs), termed Crohn's-like reaction (CLR), and those with diffuse inflammatory infiltration (DII) lacking TLSs. CLR tumors are typically more immune-infiltrated ("hot"), while DII tumors show immune exclusion ("cold") (Schürch et al., 2020). Each patient contributed four images, each from a separate biopsy, yielding a multilevel structure: images nested within patients (4 per patient), patients nested within groups (17–18 per group), and two patient groups in total (CLR and DII).

The dataset includes single-cell annotations for ˜200,000 cells across 16 cell types. For analysis, we focused on the 8 most abundant types. We refined two labels: "stroma" cells were reclassified as hybrid epithelial-mesenchymal (E/M) cells based on co-expression of cytokeratin and vimentin (Kuburich et al., 2024), and "smooth muscle" cells were relabeled as cancer-associated fibroblasts (CAFs) due to expression of α-SMA and vimentin (Cao et al., 2025). Additional cell types included CD163+ macrophages (TAMs), CD8+ T cells, granulocytes, memory CD4+ T cells, tumor cells, and vasculature—key players in the tumor immune microenvironment.

We focused our analysis on a fixed set of directional interactions motivated by prior biological knowledge of cell type function and spatial influence in the tumor microenvironment. $CD8^+$ T cells, memory $CD4^+$ T cells, and granulocytes were treated as target populations, as these immune cell types are typically recruited or excluded in response to local environmental cues. As source populations, we selected vasculature, tumor cells, CAFs, TAMs, and hybrid E/M cells due to their known roles in structuring tissue architecture, secreting chemotactic or immunosuppressive factors, and organizing spatial niches.

Vasculature acts as an entry point for immune cells and shapes infiltration patterns; tumor cells and CAFs contribute to immune exclusion through ECM remodeling and cytokine signaling; TAMs modulate immune suppression and local inflammation; and hybrid E/M cells may influence spatial dynamics through motility and stromal interactions. Target populations were selected based on their functional relevance to anti-tumor immunity and their expected sensitivity to local microenvironmental structure. For each target cell type, we jointly estimated SICs with respect to all source types, yielding interpretable, distance-resolved summaries of unique spatial association.

To implement this modeling approach, we generated the necessary inputs for each target–specific model. First, we created a quadrature for each image consisting of 1,000 dummy points for each cell type. We then used a set of 3 radial basis functions for the basis set ϕ_p, as shown in Figure A.5, to create the distance-based interaction features $\boldsymbol{q}_{A_k}(v)$. After data preparation, we fit our model using variational inference and sampled 1,000 posterior draws after model fitting.

4.4.2 How are specific cell types spatially organized around others in the tumor microenvironment?

A fundamental question in tumor microenvironment analysis is whether certain cell types tend to cluster near or avoid others, and how these patterns depend on spatial scale. These spatial relationships reflect underlying mechanisms of communication, attraction, repulsion,

or physical constraints, and can offer insight into processes such as immune surveillance, tumor evasion, and niche formation. The directional SICs estimated by SHADE allow us to detect cell-type-specific clustering and exclusion behaviors that reflect the underlying structure of immune, stromal, and tumor cell organization.

We extracted all sets of image-, patient- and cohort-level parameters ($\delta^{(m,p)}_{t_1\to t_2}$,$\gamma^{(n,p)}_{t_1\to t_2}$ and $\psi^{(g,p)}_{t_1\to t_2}$) and calculated the SICs as described in (4.4). Figure 4.8 shows an example SIC for CTLs as the target cell type, highlighting their interaction with CAFs as one of several jointly modeled source cell types. For example, at approximately 75 microns, CTLs in CLR patients appear to cluster around CAFs (log-intensity > 0), whereas in DII patients, a relative depletion of CTLs is observed near CAFs. Patient-level SICs are also shown to illustrate within-group heterogeneity (Figure A.6). Across all cell type pairs, we observed consistently negative associations at short spatial ranges (< 25 microns), likely reflecting physical crowding and space constraints that limit close proximity between cells. At intermediate distances (25–75 microns), many pairs exhibited a distinct positive bump in their spatial interaction curves, suggestive of clustering behavior; notably, the magnitude of this effect often differed between patient groups. In contrast, associations at longer ranges (> 75 microns) were comparatively sparse, indicating limited coordinated spatial organization at these distances.

4.4.3 How do spatial interaction patterns vary across patients and tissue sections?

Spatial interactions may differ substantially across individual tumors and even between tissue sections from the same patient. This variation may reflect biologically meaningful heterogeneity in tumor architecture, immune activity, or microenvironmental organization, and is important for understanding the extent to which spatial structure is conserved or patient-specific. SHADE's hierarchical structure enables the estimation of SICs at the image, patient, and cohort levels, allowing us to evaluate both intra-patient and inter-patient variability. This analysis helps identify which spatial patterns are consistent across patients and which reflect individual-specific tumor architecture or microenvironmental features.

To assess variability in spatial interaction structure across samples, we summarize between-patient and between-image heterogeneity using median absolute deviation (MAD)-based estimates. These measures are defined in the Supplement. Figure 4.9 shows heatmaps of the resulting variability estimates for each source–target cell type pair.

We observed substantial variability in spatial interaction structure across both patients and images (Figure 4.9). Notably, granulocytes clustering around TAMs and vasculature ex-

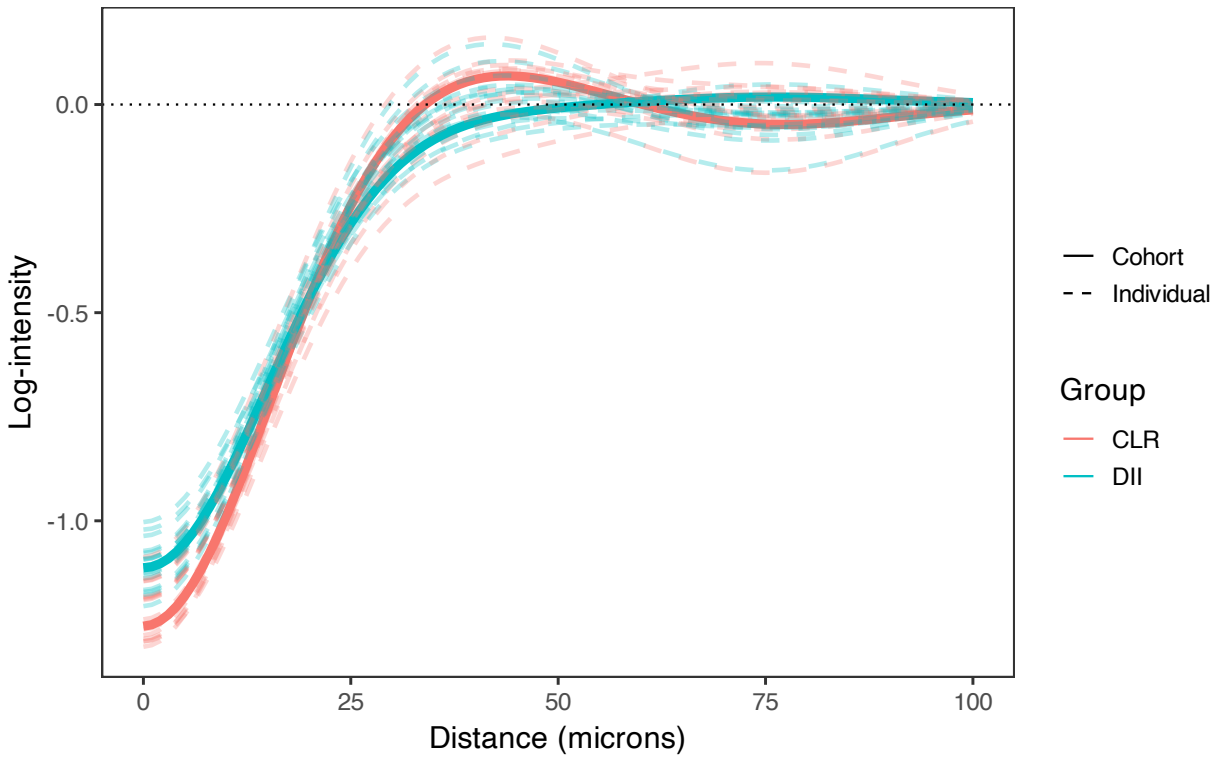


Figure 4.8: SIC showing the directional association of CAFs (source) with CTLs (target), stratified by patient group. Solid lines indicate cohort-level estimates for CLR and DII groups, while dashed lines represent patient-level SICs, highlighting within-group variability. The plot illustrates how CTLs tend to slightly cluster around CTLs in CLR patients at mid-to-long ranges (positive log-intensity), whereas a relative neutral to avoidant behavior is observed in DII patients—demonstrating a group-specific spatial pattern in CAF-CTL organization.

hibited high between-patient heterogeneity. This suggests that granulocyte recruitment and localization near TAMs and vessels may be modulated by patient-specific microenvironmental states, such as differences in angiogenic signaling or myeloid activation. Within patients, granulocyte interactions also showed high heterogeneity, but principally in association with vasculature and hybrid E/M cells. These image-level differences may reflect localized fluctuations in endothelial activation or epithelial plasticity, which may differentially attract or retain granulocytes within specific tissue regions. Examples of patient-level curves for two patients, along with their image-level curves, are shown in Figure 4.10.

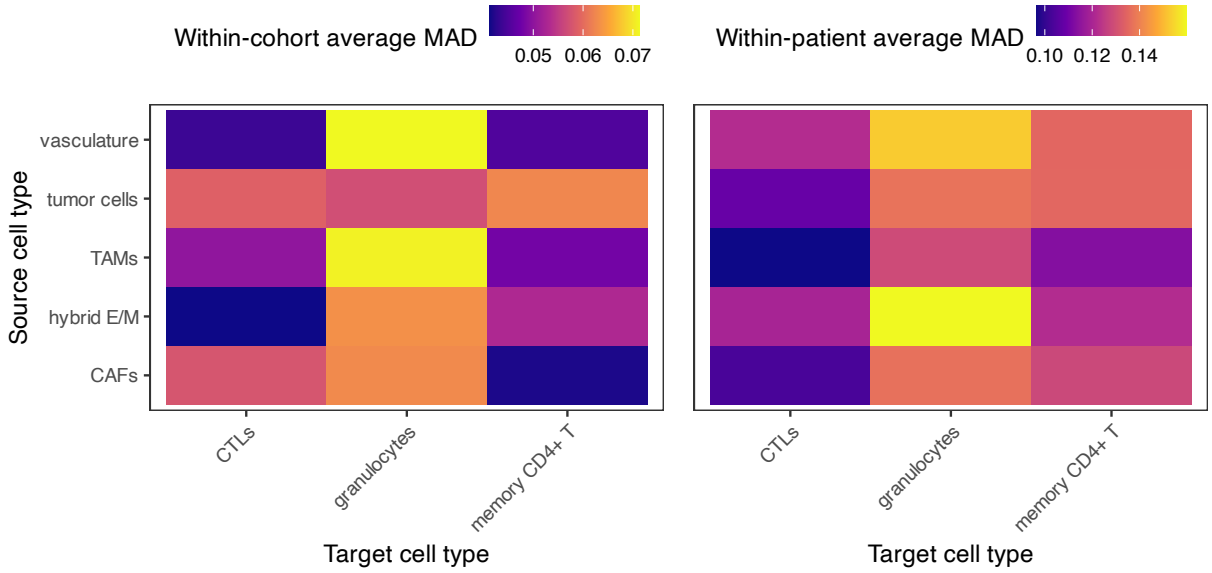


Figure 4.9: Heatmaps show the median absolute deviation (MAD) of SICs across spatial distances, summarizing two sources of variability: between patients (within cohort) and between images (within patient). Higher values indicate greater heterogeneity in spatial structure.

4.4.4 How do immune and stromal spatial interactions differ between hot and cold tumors?

Distinct tumor immune phenotypes—such as immune-infiltrated (CLR) versus immune-excluded (DII) tumors—are associated with differential immune activity and prognosis. These differences are often accompanied by changes in how immune and stromal cells are spatially arranged in relation to tumor cells and each other. We used SHADE to compare SICs across these patient groups and identify interactions that are enriched or depleted in each context. This allowed us to link differences in spatial structure to known functional differences in immune activity.

We observed several notable differences in spatial interaction curves (SICs) between CLR and DII patients (Figure 4.11). Below, we highlight a subset of interactions where the between-group difference in log-intensity exceeds 0.05 (dashed lines in Figure 4.11), focusing on plausible underlying biological mechanisms. First, CLR patients showed increased clustering of CTLs and memory CD4+ T cells around vasculature at short spatial ranges, suggesting enhanced immune surveillance and trafficking in these patients. The vasculature serves as a conduit for lymphocyte infiltration into the tumor microenvironment, and vascular normalization in immunologically active tumors has been linked to improved T cell

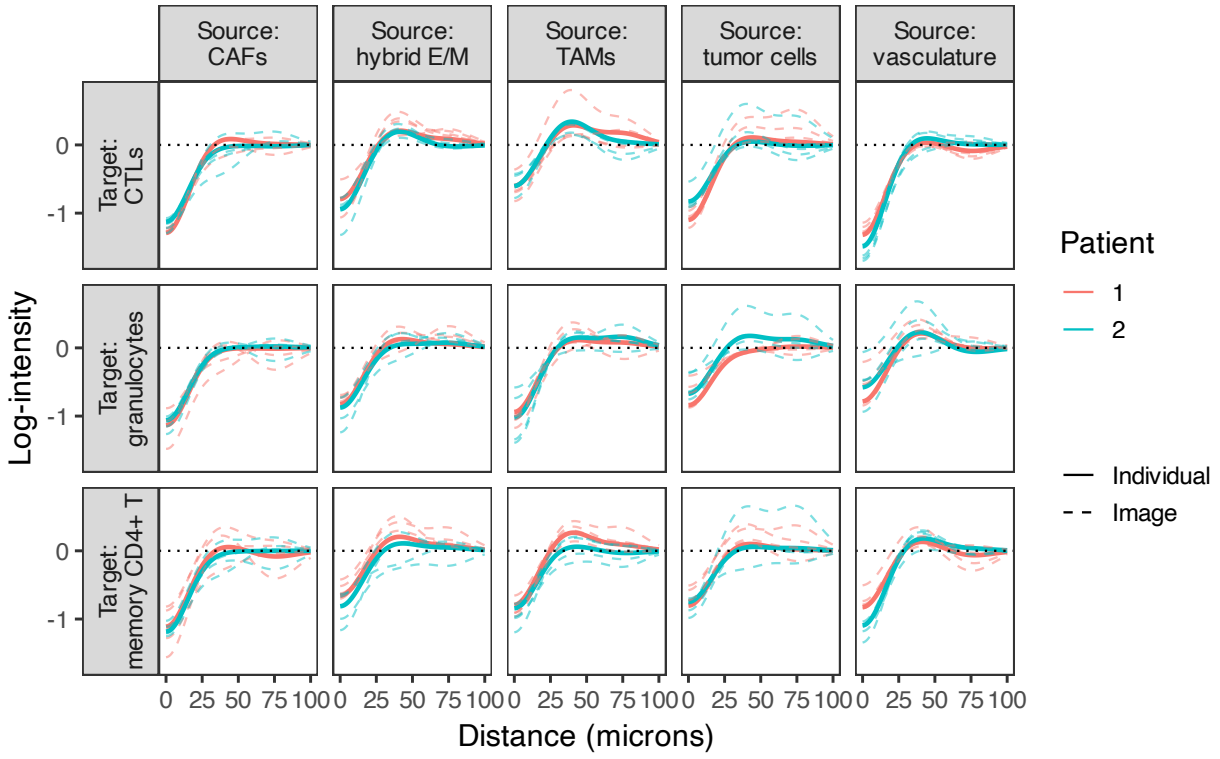


Figure 4.10: Examples of patient- and image-level SICs for cell type pairs with high within-patient MAD values. Each dotted line represents a SIC estimated from a single tissue section, while each solid line indicates a patient-level SIC.

recruitment and anti-tumor immunity (Lanitis, Irving, and Coukos, 2015; Hendry et al., 2016). This may reflect a more organized immune response in CLR patients, consistent with better clinical outcomes. In contrast, DII patients exhibited greater clustering of CTLs around tumor cells at all spatial ranges. While this might suggest immune recognition, it may alternatively indicate ineffective or exhausted T cell responses that fail to clear tumor cells. Persistent CTL-tumor colocalization without effective cytolysis has been associated with immune dysfunction and tumor immune escape (Dolina et al., 2021; Raskov et al., 2021).

We also observed distinct patterns in CTL clustering around CAFs: DII patients displayed increased clustering at short and long spatial ranges, whereas CLR patients showed elevated clustering at medium ranges. CAFs are known to mediate immune suppression via direct inhibition of CTLs and by establishing physical and cytokine-mediated barriers (L. Jenkins et al., 2022; Freeman and Mielgo, 2020). The long-range clustering in DII patients may reflect exclusionary barriers or CAF-mediated trapping of CTLs, whereas the medium-range interactions in CLR patients could represent transient engagement or immune cell penetration

into CAF-dense zones under less suppressive conditions. Lastly, CLR patients demonstrated increased clustering of memory CD4+ T cells at short distances from both CAFs and hybrid E/M cells. This may suggest spatial niches where helper T cells coordinate localized immune responses, potentially promoting anti-tumor immunity or influencing the differentiation state of tumor cells. CD4+ T cells have been implicated in the modulation of EMT processes, and their proximity to hybrid E/M cells may reflect attempts to restrain phenotypic plasticity and metastatic potential (Xie et al., 2025; Milosevic and Östman, 2024).

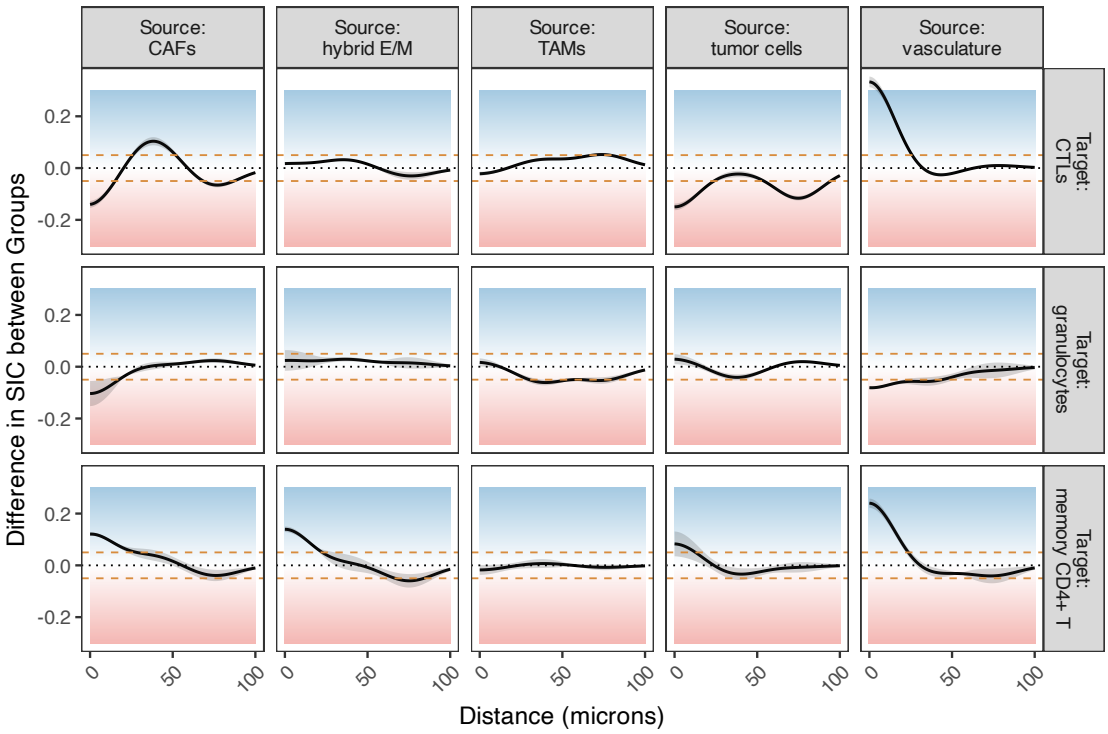


Figure 4.11: The difference in SICs between the two patient groups. Curves above 0 (in the blue area) indicate that the log-intensity is greater in CLR patients than DII patients; curves in the red area show the reverse.

4.4.5 How well can spatial interactions predict the organization of individual cell types across tumor regions?

In addition to estimating SICs, SHADE generates predictions of the spatial distribution of each cell type, conditional on the presence of all others. Understanding how well these interactions predict observed spatial patterns can help identify which cells are tightly governed by spatial associations and which may be spatially organized by other factors. These predictions provide insight into how well spatial interactions explain observed cellular ar-

rangements, and which cell types are more or less predictable. By evaluating predictive accuracy across images and patient groups, we gain further understanding of the robustness of spatial organization and how it may vary with tumor subtype.

For example, in image 47_B from the dataset, we can produce the following predictions of each target cell type, conditional on the source cell types (Figure 4.12). We can see that the target cell types are quite well predicted, especially CTLs. Likely due to their relative paucity, granulocytes are not quite as well predicted in this image (regions of localization are not as distinct).

We calculated AUCs per cell type and per group (Figure A.7). All three target cell types were better predicted in CLR patients than in DII patients, though there was a relatively large amount of variability in image-level AUCs.

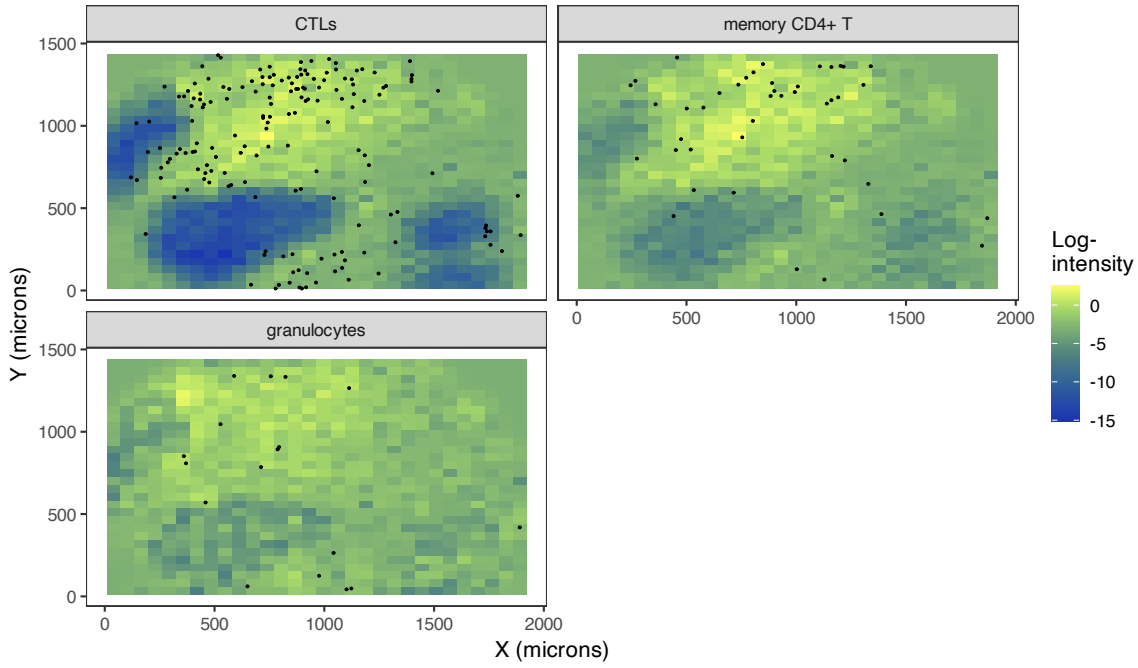


Figure 4.12: Predicted spatial distributions for target cell types in a representative image (47_B), based on the estimated conditional intensity functions.

4.5 Discussion

In this work, we introduced SHADE, a Bayesian hierarchical model for quantifying asymmetric spatial interactions in multiplexed imaging data. By modeling the conditional intensity function and estimating spatial interaction curves, our approach enables the identification

of directional associations between cell types while accounting for multilevel variation across images, patients, and cohorts.

Through simulation studies, we demonstrated that SHADE effectively captures asymmetric spatial associations while leveraging hierarchical structure to improve estimation. Including multilevel priors improves inference accuracy, especially for short-range interactions, by reducing bias and variance in estimated coefficients. We also found that increasing the number of dummy points improves inference up to a point, but can introduce diminishing returns or mild instability in some settings.

SHADE revealed distinct and biologically meaningful patterns of cellular organization in colorectal tumors, highlighting how immune and stromal interactions differ between immune-infiltrated (CLR) and immune-excluded (DII) phenotypes. By estimating directional spatial interaction curves across multiple spatial scales and model hierarchies, we identified key differences in the local microenvironment that may underlie divergent immune states. Notably, CLR tumors exhibited increased short-range clustering of cytotoxic and memory T cells around vasculature, consistent with enhanced immune infiltration via normalized vessels and the formation of perivascular immune niches. In contrast, CTLs in DII tumors showed elevated spatial association with tumor cells across all distances—a pattern that may reflect either failed immune surveillance or dysfunctional retention of exhausted T cells.

Our analysis also uncovered cell-type-specific interaction patterns with stromal components. CTLs in DII patients showed greater clustering near cancer-associated fibroblasts at both short and long spatial ranges, whereas in CLR patients, CTL–CAF interactions peaked at intermediate distances. Given CAFs' established role in immune suppression and physical exclusion, these patterns may reflect differential stromal constraint across tumor types. Similarly, CLR patients exhibited increased short-range clustering of memory CD4+ T cells around both CAFs and hybrid epithelial–mesenchymal cells. These interactions suggest a potentially immunoregulatory or anti-tumor role for helper T cells in CLR tumors, possibly including modulation of EMT plasticity or the stabilization of immune-supportive niches.

Beyond cohort-level differences, we also found substantial heterogeneity in spatial organization across patients and across tissue sections from the same patient. This variability was particularly pronounced in granulocyte interactions with TAMs, vasculature, and hybrid E/M cells, suggesting patient-specific or localized microenvironmental factors influencing spatial localization. Finally, SHADE's ability to predict target cell distributions from source cell types varied by cell type and tumor subtype: immune cell organization was more predictable in CLR tumors, indicating more structured immune environments compared to the more disorganized spatial context of DII tumors. Collectively, these results emphasize the need for multiscale spatial modeling to capture both consistent interaction motifs and the

biologically meaningful variability that characterizes tumor immune landscapes.

A key strength of SHADE is its flexibility in modeling directional interactions across spatial scales, avoiding the restrictive symmetry assumptions of traditional spatial models. Moreover, the use of logistic regression for conditional intensity estimation allows for scalable and stable inference, sidestepping common numerical pitfalls of direct Poisson modeling.

Nonetheless, SHADE has limitations. First, while SICs capture directional spatial association, they remain correlational and cannot determine causality or infer mechanisms of interaction. Second, the SICs reflect spatial dependence rather than molecular signaling pathways, which may limit biological interpretability in some settings. Third, although SHADE supports biologically motivated conditioning structures, exploratory analyses may benefit from modeling both directions of association when directionality is uncertain.

Future extensions of SHADE could incorporate functional covariates—such as marker intensity, proliferation, or exhaustion scores—into the SIC framework, enabling joint analysis of spatial structure and functional state. This would open the door to modeling not just where cells are, but how they behave spatially in context.

Overall, our results highlight the importance of capturing asymmetric and hierarchical structure in spatial models, and position SHADE as a powerful tool for dissecting the complex architecture of tissue microenvironments in cancer and beyond.

CHAPTER 5

The BIPPS Model for Shared Spatial Dependencies Across Patients

5.1 Introduction

Characterizing the spatial structure of the TME requires quantifying both within-type spatial dependence and cross-type interactions among cells. Multivariate geostatistical models achieve this through marginal and cross-correlation functions, which describe how spatial correlation decays with distance (Banerjee, Carlin, and Gelfand, 2014; Genton and Kleiber, 2015). These functions capture spatial organization within and between cell types, offering a comprehensive and interpretable framework for modeling the TME.

However, extending these models to multi-subject imaging data presents key challenges. Existing spatial models are computationally intensive and often restricted to single-subject analyses, limiting generalizability. Some multi-subject approaches rely on univariate spatial models (Xu et al., 2009; Sanyal and Ferreira, 2012) or non-spatial factor models (De Vito et al., 2019; Vito et al., 2021; Wang and Lock, 2024; Lock, Park, and Hoadley, 2022), while others use black-box machine learning methods that lack interpretability and principled uncertainty quantification (Hu et al., 2021; McKinley et al., 2022; Z. Wu et al., 2022). Tools focused on image processing (Schapiro, Sokolov, et al., 2022; Warchol et al., 2023) similarly fall short of providing statistical inference.

To address these gaps, we introduce Bayesian Inference for Point Processes in Space (BIPPS), a hierarchical Bayesian model for multi-subject multiplexed imaging data. We model each image as a log-Gaussian Cox process (LGCP, Diggle et al. 2013), using a multivariate Gaussian process to capture both spatial and inter-cell-type dependence. Crucially, BIPPS shares spatial parameters across subjects, enabling partial pooling while retaining subject-level heterogeneity. The model yields interpretable estimates of all marginal and cross-correlation functions, making it especially valuable for cohort-level spatial inference in studies with limited per-subject samples.

We apply BIPPS to two multiplexed imaging datasets. The first is a pancreatic cancer dataset from the University of Michigan Pancreatic Cancer Center (UM-PCC) (S. N. Krishnan, Mohammed, et al., 2022; S. N. Krishnan, Ji, et al., 2024), while the second is a public colorectal cancer (CRC) dataset with 35 patients (Schürch et al., 2020). Our framework enables rigorous analysis of disease-specific spatial structure, shedding light on tumor-immune interactions and their relevance to cancer progression and treatment.

Section 5.2 describes the datasets and preprocessing. Section 5.3 presents our model. Section 5.4 demonstrates performance on synthetic data. Sections 5.5 and 5.6 present applications to the UM-PCC and CRC datasets, respectively. We conclude with a summary and discussion. The Supplement provides extended methods, simulation diagnostics, sensitivity analyses, and additional results. An `R` software package for implementing our methods is available at github.com/mkln/bipps, with code to reproduce our analyses at github.com/jeliason/bipps_paper_code.

5.2 Many-subject multiplexed imaging datasets

Multiplexed imaging datasets from multiple subjects enable the spatial characterization of tissue organization at single-cell resolution. Despite differences in disease context, these datasets share a common structure: each image consists of spatially localized cells annotated with phenotypic labels derived from standardized staining and imaging pipelines. The primary challenge in analyzing such datasets lies in modeling spatial dependencies across subjects while accounting for inter-patient variability in cell composition and tissue architecture. In Figure 5.1, we show the workflow for this analysis, beginning with the collection of biopsy samples from different patient cohorts, followed by multiplex imaging and cell phenotyping. The resulting spatial point patterns are then transformed into multivariate binned count data, which serves as the structured input for our method.

5.2.1 Pancreatic cancer dataset

Images from the pancreatic cancer dataset were obtained from patients who had presented at the University of Michigan Pancreatic Cancer Clinic and subsequently underwent surgical resection for various pancreatic pathologies. The study population consisted of patients diagnosed with six distinct pancreatic diseases, encompassing both malignant and non-malignant conditions. Our primary focus was on two specific patient groups: intraductal papillary mucinous neoplasm (IPMN, 70 patients) and pancreatic ductal adenocarcinoma (PDAC, 71 patients).

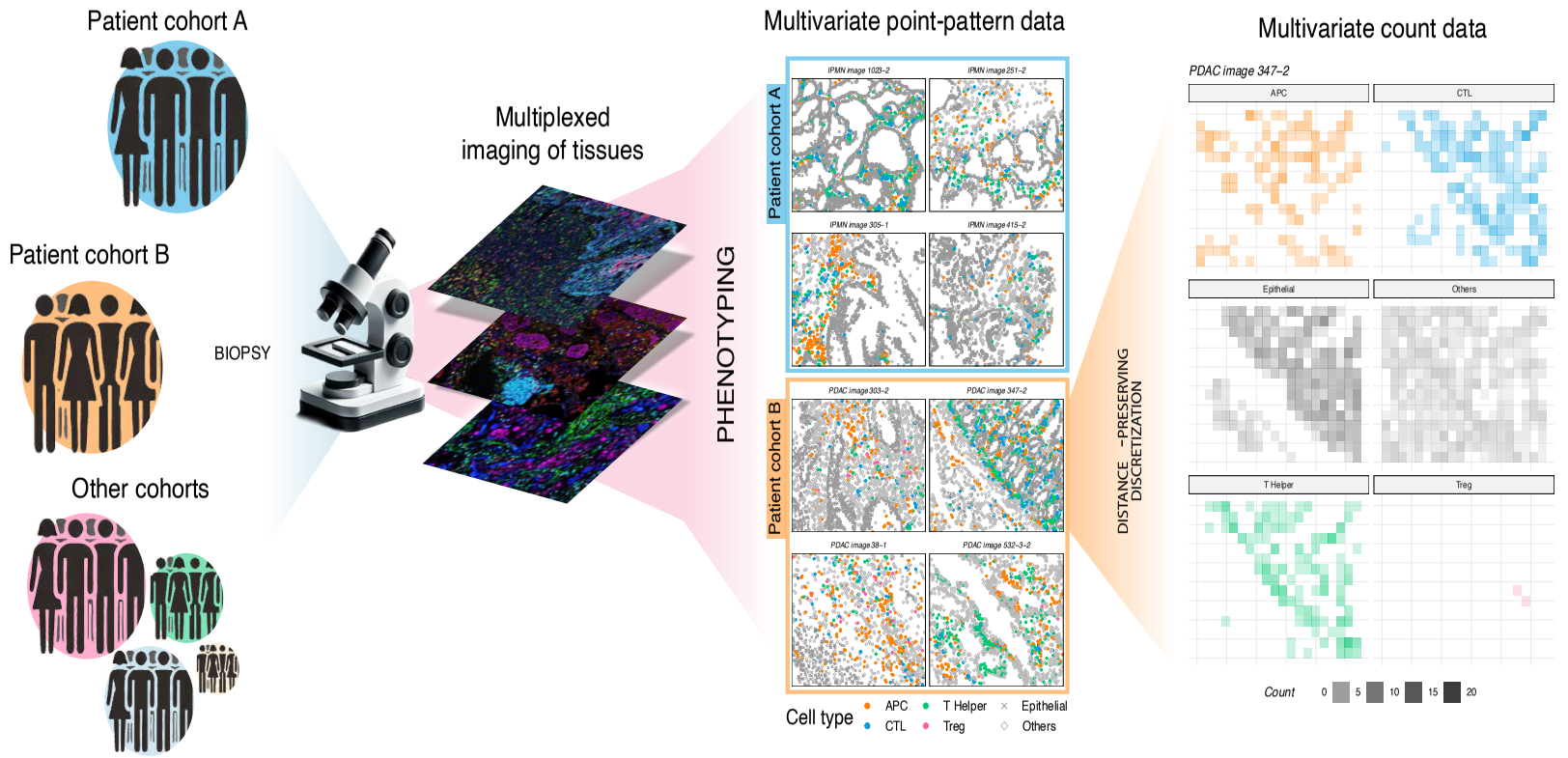


Figure 5.1: Overview of the imaging and statistical analysis workflow. Biopsy samples are first collected from patients across different cohorts. These biopsies are then processed and imaged using multiplex imaging techniques, allowing for high-dimensional spatial characterization of the tumor microenvironment. Following imaging, cell phenotyping is performed to classify cells based on marker expression, generating spatial point patterns for each sample. These point patterns are subsequently converted into multivariate binned count data, which serves as the structured input for the statistical method presented in this study. This figure was developed primarily by Michele Peruzzi.

Tissue images were obtained via multiplexed immunofluorescence (mIF) staining, which was performed following protocols outlined in Lazarus et al., 2018. Processing and subsequent phenotyping were carried out using the methods described by S. N. Krishnan, Mohammed, et al., 2022. Phenotypic classification was performed using user-defined thresholds and standardized with inForm Tissue Analysis software (Akoya, 2021) across the six original disease groups. Each cell was classified into one of six phenotypes: antigen-presenting cell (APC), cytotoxic T lymphocyte (CTL), epithelial (Epithelial), helper T cell (T Helper), regulatory T cell (Treg), or others. Figure 5.2 illustrates representative point patterns from the pancreatic cancer dataset, highlighting spatial distributions of these phenotypes in selected IPMN and PDAC images. The bottom row further demonstrates the transformation of an IPMN point pattern (sample 909-1) into a binned count representation, which aggregates local cell densities to serve as input for subsequent statistical analysis by our method.

5.2.2 Colorectal cancer dataset

The second dataset we consider consists of multiplexed images of tumor tissue samples sourced from a publicly available colorectal cancer repository (Schürch et al., 2020). These images depict various cell types within the tumor microenvironment, including malignant epithelial cells, various immune subtypes and other stromal cells. The PhenoCycler imaging technology (Goltsev et al., 2018; Black et al., 2021) was used to generate this dataset, which consists of multiplexed imaging data from 35 CRC patients, each contributing four biopsies. A single high-dimensional image was acquired per biopsy, capturing the expression of 56 distinct proteins across tumor tissue sections. Patients were categorized based on tumor histopathology into two groups: those with visible tertiary lymphoid structures (TLSs) at the invasive front (CLR, Crohn's-like reaction) and those with diffuse inflammatory infiltration (DII) without detectable TLSs. CLR patients generally exhibit an immune-hot, well-mixed tumor microenvironment and a more favorable prognosis, whereas DII patients have an immune-exclusive, cold TME (Schürch et al., 2020).

The original study annotated 16 cell types across nearly 200,000 cells, but we focus on the 10 most common. We also relabeled "stroma" cells as "hybrid E/M" due to co-expression of Vimentin and cytokeratin, indicating both epithelial and mesenchymal traits (Kuburich et al., 2024). Similarly, "smooth muscle" cells were reclassified as cancer-associated fibroblasts (CAFs) based on α-SMA and Vimentin expression (Cao et al., 2025). We excluded the six least frequent cell types from the CRC analysis because their rarity limits their contribution to spatial modeling. Since these cell types provide minimal spatial information, we focused on the more prevalent populations. In future work, we may extend the model to incorporate non-

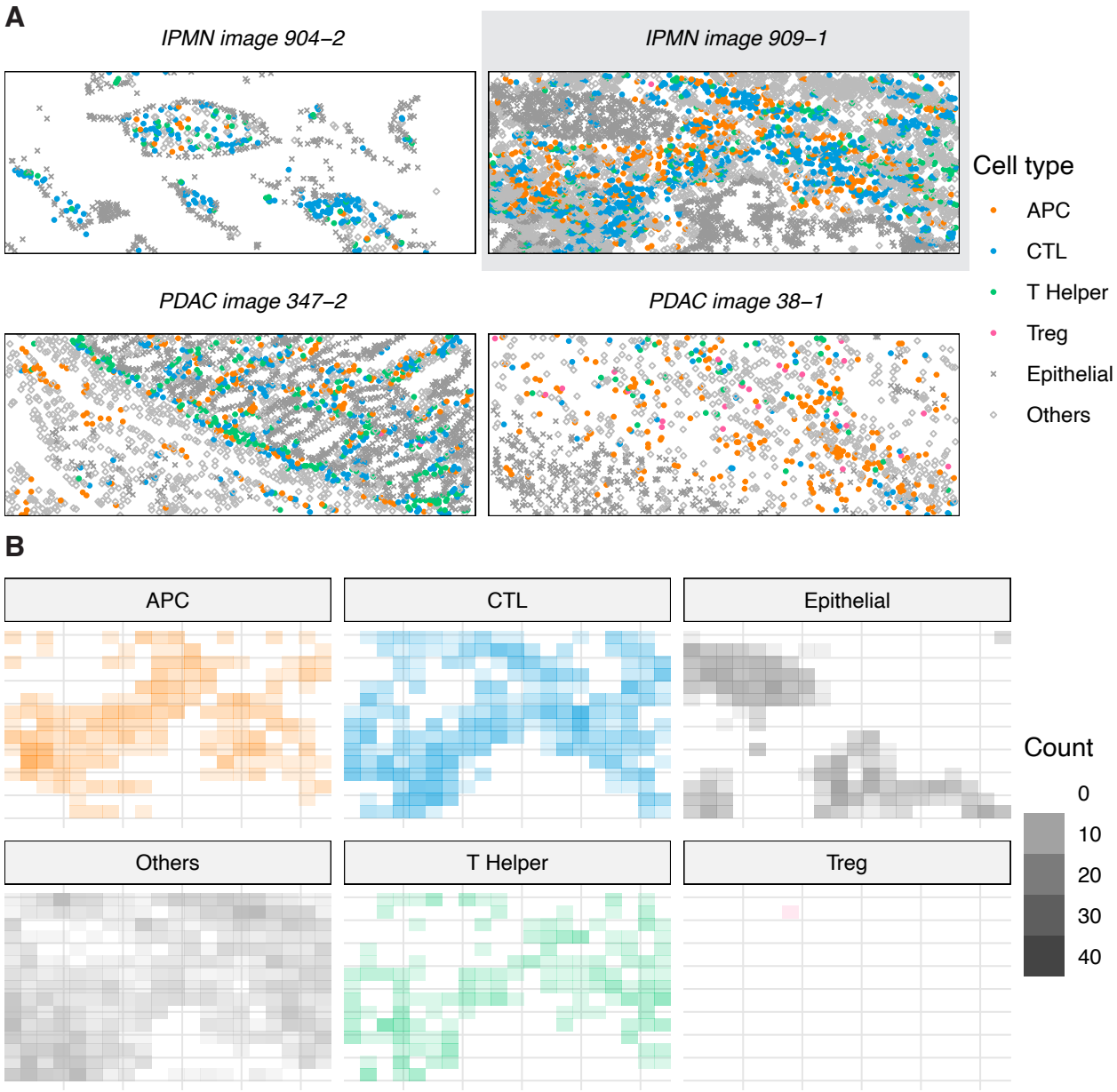


Figure 5.2: Point patterns derived from selected images in the pancreatic cancer dataset, including both IPMN and PDAC cases. **A)** Two images from each cancer type. **B)** IPMN image 909-1 is transformed into the binned count representation. The visualized areas in each image are approximately 1300 μm (x-axis) by 1000 μm (y-axis).

spatial components to account for these rare cell types. Figure 5.3 presents representative point patterns from the colorectal cancer dataset, illustrating spatial distributions in selected CLR and DII cases, with the bottom row showing the binned count representation for sample 59-A.

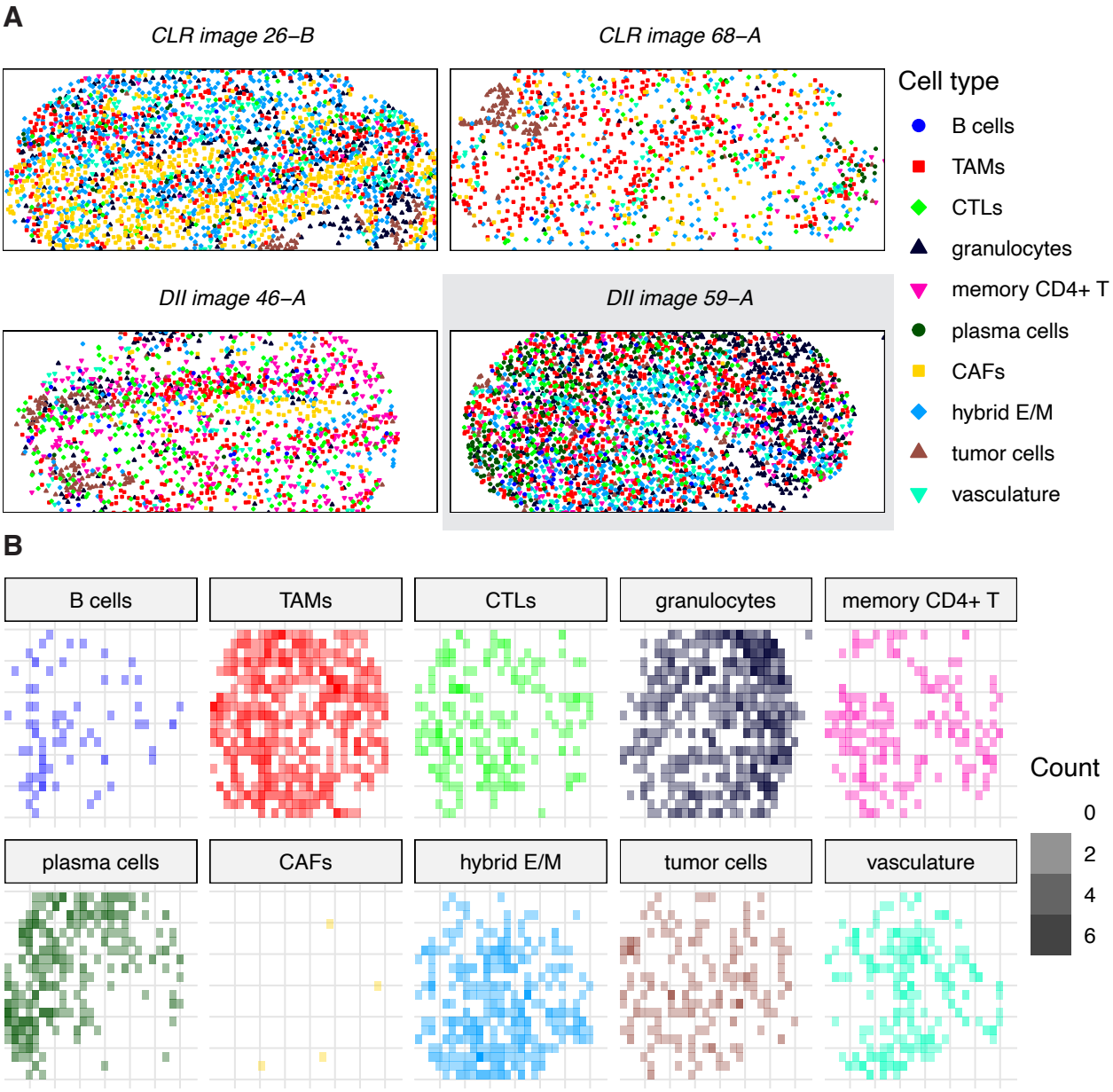


Figure 5.3: Point patterns derived from selected images in the colorectal cancer dataset, including both CLR and DII cases. **A)** Two images from each cancer type. **B)** DII image 59-A is transformed into the binned count representation. The visualized areas in each image are approximately 1800 μm (x-axis) by 1400 μm (y-axis).

5.3 Multi subject Bayesian latent spatial factor model

Let $\mathcal{D}_i \subset \Re^2$ represent the ith multiplexed image in the data, with $i = 1, \ldots, N$[1]. Moving forward, we assume that each subject is associated to a single image to simplify exposition

[1]This section was adapted from a manuscript coauthored with Michele Peruzzi and Arvind Rao. The original draft of this section was written primarily by Michele Peruzzi.

and without loss of generality. Then, let $\mathcal{Y}_{ij} = \{Y_{ij}(\boldsymbol{l}) : \boldsymbol{l} \in \mathcal{D}\}$ be a subject-specific inhomogenous Poisson process related to the point-pattern of cells of type j, where $j = 1, \ldots, q$. We then introduce $\mathcal{W} = \{\boldsymbol{w}(\boldsymbol{l}) : \boldsymbol{l} \in \mathcal{D}\}$ as a zero-mean multivariate GP and denote as $w_{ij}(\boldsymbol{l})$ the realization at $\boldsymbol{l}$ of its jth margin for subject i. $\mathcal{W}$ is completely defined by the user-specified cross-covariance matrix function indexed by parameter vector $\boldsymbol{\theta}$ and defined as $\boldsymbol{C}(\cdot, \cdot; \boldsymbol{\theta}) : \mathcal{D} \times \mathcal{D} \to \boldsymbol{M}_{q\times q}$, where $\boldsymbol{M}_{q\times q}$ is the set of $q \times q$ real-valued matrices, and such that $\boldsymbol{C}(\boldsymbol{l}, \boldsymbol{l}'; \boldsymbol{\theta}) = \boldsymbol{C}(\boldsymbol{l}', \boldsymbol{l}; \boldsymbol{\theta})^\top$ and $\sum_{i=1}^n \sum_{j=1}^n \boldsymbol{x}_i^\top \boldsymbol{C}(\boldsymbol{l}_i, \boldsymbol{l}_j; \boldsymbol{\theta}) \boldsymbol{x}_j > 0$ for any integer n, any finite set $\{\boldsymbol{l}_1, \boldsymbol{l}_2, \ldots, \boldsymbol{l}_n\}$, and for all $\boldsymbol{x}_i, \boldsymbol{x}_j \in \Re^q \setminus \{\boldsymbol{0}\}$ (Genton and Kleiber, 2015). Therefore, by construction, the (r, s)th entry of $\boldsymbol{C}(\boldsymbol{l}, \boldsymbol{l}')$ evaluates to $C_{rs}(\boldsymbol{l}, \boldsymbol{l}'; \boldsymbol{\theta}) = \text{cov}(w_r(\boldsymbol{l}), w_s(\boldsymbol{l}'))$. Under an assumption of stationarity, we have $\boldsymbol{C}(\boldsymbol{l}, \boldsymbol{l}'; \boldsymbol{\theta}) = \boldsymbol{C}(\boldsymbol{h}; \boldsymbol{\theta})$, where $\boldsymbol{h} = \|\boldsymbol{l} - \boldsymbol{l}'\|$. We then complete the LGCP by assuming that, conditional on $\mathcal{W}$, the intensity of $\mathcal{Y}_{ij}$ at location $\boldsymbol{l}$ is $\Lambda_{ij}(\boldsymbol{l}) = \exp\{\boldsymbol{x}_i^\top(\boldsymbol{l})\boldsymbol{\beta}_j + w_{ij}(\boldsymbol{l})\}$. We choose a spatial factor model as our cross-covariance matrix function: $C_{rs}(\boldsymbol{l}, \boldsymbol{l}') = \sum_{j=1}^k \lambda_{rj}\lambda_{sj}\rho(\boldsymbol{l}, \boldsymbol{l}'; \varphi_j)$, where λ_{rj} is the (r, j) element of the $q \times k$ matrix $\boldsymbol{A}$ of factor loadings, $\rho(\boldsymbol{l}, \boldsymbol{l}'; \varphi_j) = \exp\{-\varphi_j\|\boldsymbol{l} - \boldsymbol{l}'\|\}$, and therefore $\boldsymbol{\theta} = \{\boldsymbol{A}, \varphi_1, \ldots, \varphi_k\}$.

Our model specification has two main features: first, our conditional independence assumption implies that $\mathcal{Y}_{ij}(\boldsymbol{l})$ incur in spatial and cross-cell-type dependence by means of marginalization of $\mathcal{W}$. Our primary inferential targets are the marginal correlation functions $\rho_{rr}(\boldsymbol{h}; \boldsymbol{\theta}) = \frac{C_{rr}(\boldsymbol{h};\boldsymbol{\theta})}{C_{rr}(\boldsymbol{0};\boldsymbol{\theta})}$ and the cross-correlation functions $\rho_{rs}(\boldsymbol{h}; \boldsymbol{\theta}) = \frac{C_{rs}(\boldsymbol{h};\boldsymbol{\theta})}{\sqrt{C_{rr}(\boldsymbol{0};\boldsymbol{\theta})}\sqrt{C_{ss}(\boldsymbol{0};\boldsymbol{\theta})}}$, where $r \neq s$. These functions (which total $q(q+1)/2$ in number) give an easily interpretable characterization of the joint spatial pattern of the log-intensity process among all cell types. The second main feature of our model is that we assume $\boldsymbol{\theta}$ is common across *all* subjects in the same group. Because we group subjects by disease and $\boldsymbol{\theta}$ determines the shape of all correlation functions $\rho_{rs}(\cdot; \boldsymbol{\theta})$ for $r, s = 1, \ldots, q$, our model pools information from all subjects to characterize the disease-specific joint point pattern of cell types.

Because computations involving LGCPs are highly complex even in univariate single-subject settings, we follow Diggle et al., 2013 and discretize the image using a computational grid of dimension $n_x \times n_y$, for a total of n_{px} pixels. Suppose $\boldsymbol{l}$ and $\boldsymbol{l}'$ are two locations in any of the images in the data. The preprocessing pipeline we described in Section 5.3.1 ensures that $\boldsymbol{h} = \|\boldsymbol{l} - \boldsymbol{l}'\|$ is the same for all images and therefore ensure that each image correctly contributes to learning the common cross-covariance matrix function $\boldsymbol{C}(\cdot; \boldsymbol{\theta})$. We then refer to $y_{ij}(\boldsymbol{l})$ as the number of cells of type j at location $\boldsymbol{l}$ in the ith image. The LGCP model

on the discretized images translates to

$$\begin{aligned} y_{ij}(\boldsymbol{l}) \mid \lambda_{ij}(\boldsymbol{l}) &\sim \text{Poisson}(\lambda_{ij}(\boldsymbol{l})), & \boldsymbol{\lambda}_i(\boldsymbol{l}) &= \left[\lambda_{i1}(\boldsymbol{l}) \cdots \lambda_{iq}(\boldsymbol{l})\right]^\top \\ \log \boldsymbol{\lambda}_i(\boldsymbol{l}) &= \boldsymbol{\alpha}_i + \boldsymbol{X}_i(\boldsymbol{l})\boldsymbol{B} + \boldsymbol{w}_i(\boldsymbol{l}), & \boldsymbol{w}_i(\boldsymbol{l}) &= \boldsymbol{A}\boldsymbol{v}_i(\boldsymbol{l}), \\ \boldsymbol{v}_i(\boldsymbol{l}) &= \left[v_{i,1}(\boldsymbol{l}), \ldots, v_{i,k}(\boldsymbol{l})\right], ^\top & v_{i,j}(\cdot) &\sim GP(\boldsymbol{0}, \rho(\cdot,\cdot;\varphi_j)), \end{aligned} \tag{5.1}$$

where we assign zero-mean Gaussian priors to $\boldsymbol{\alpha}_i$, each column of $\boldsymbol{B}$, and each row of $\boldsymbol{A}$. Finally, we assign the prior $\varphi_j \sim U(0.1, 10), j = 1, \ldots, k$ to allow a wide range of possible spatial dependence. In (5.1), $\boldsymbol{\alpha}_i$ is a cell-type-specific offset for the overall intensity, whereas the term $\boldsymbol{w}_i(\boldsymbol{l}) = \boldsymbol{A}\boldsymbol{v}_i(\boldsymbol{l})$ represents a latent spatial factor model defined above. Because the latent spatial effects $\boldsymbol{w}_i(\boldsymbol{l})$ are subject specific, our model allows variability in the observed images without requiring that they align–indeed, that would be impossible because each image corresponds to a different tissue. Nevertheless, our model borrows information across all images to estimate all $q(q+1)/2$ cross-covariance functions. Finally, we assume the effect of covariates $\boldsymbol{X}_i(\boldsymbol{l})$ to be common across images; although relaxing this assumption is straightforward, it is beyond the scope of this article.

GPs are notoriously computationally inefficient when the number of spatial locations is large. Although one possible solution is to discretize the raw data on a coarser computational grid (i.e., lower n_{px}), reducing grid size negatively impacts the ability to infer LGCP paramters (Diggle et al., 2013). Instead, we replace the GP with a meshed GP (MGP, Peruzzi, Banerjee, and Finley, 2022). An MGP is scalable to large datasets because it restricts spatial dependence using a user-specified sparse directed acyclic graph (DAG) which operates on a partition of the spatial domain. Intuitively, each domain subregion is independent of non-neighboring regions, given its neighbors. Scalable GPs based on sparse DAGs have become extremely popular in geostatistical settings due to their numerous desirable properties; notably, they lead to algorithms that scale linearly with the number of observed spatial locations, (Vecchia, 1988; Datta et al., 2016; Katzfuss and Guinness, 2021), with additional computational savings on imaging data (Peruzzi, Banerjee, and Finley, 2022), and with theoretical guarantees regarding their ability to approximate unrestricted GPs (Zhu et al., 2024). Therefore, any such DAG-based GP can scalably approximate an unrestricted GP, with inferential performance being indistinguishable in practice. We fit MGPs to our datasets via Masrkov-chain Monte Carlo methods for multivariate count data, extending those outlined in Peruzzi and Dunson (2024) to the multi-subject setting. More details on the MCMC algorithm are in the Supplement.

5.3.1 Data preprocessing workflow

To ensure compatibility with our method and its software implementation, we first standardized the spatial coordinates of each image. Specifically, we first shifted each image so that its bottom-left corner aligns with the origin $(0, 0)$. Next, we rescaled the coordinates to normalize image dimensions across the dataset. Let $l_{x,i}$ and $l_{y,i}$ represent the x-axis and y-axis lengths (in microns) of the ith image. We computed the maximum axis length across all images as $l^* = \max\{l_{x,i}, l_{y,i}\}$. Each image was then rescaled so that its spatial domain is given by $\mathcal{D}_i = \left[0, \frac{l_{x,i}}{l^*}\right] \times \left[0, \frac{l_{y,i}}{l^*}\right] \subset [0, 1]^2$. This transformation ensures that all images share a common unit of measurement while preserving relative distances within each image. This rescaling step simplifies the specification of priors for spatial range parameters while maintaining the integrity of spatial relationships.

It is customary to fit LGCPs by first discretizing the raw point-pattern data into a high resolution pixel image where each pixel is a vector of cell-type counts (Diggle et al., 2013). Therefore, we constructed an $n_x \times n_y$ grid over each image and counted the number of cells of each type within each grid pixel. We denote the resulting count data as $\boldsymbol{y}_i(\boldsymbol{l})$, a vector of dimension q, with its jth element representing the number of cells of type j at location $\boldsymbol{l}$ in the ith image. Examples of the processed images after these transformations are shown in Figure 5.2B and Figure 5.3B.

5.4 Simulation studies

We performed two simulation studies in which we examined the impact of hyperparameters such as the number of latent factors k and the number of grid pixels n_{px} on inference quality. In particular, we studied the effect of misspecification of k on the convergence of MCMC sampling, the recovery of cross-correlation curves between cell types, and the use of information criteria to appropriately choose k. In the second study, we consider the case where we use a coarser grid of pixels than is assumed in the generative model, again focusing on how MCMC sampling and cross-correlation curve recovery are affected. In both of these simulation studies, we quantified the degree of cross-correlation curve recovery via the median absolute deviation (MAD) between the observed and fitted cross-correlation curves.

5.4.1 Recovering cross-correlations when the number of latent factors is misspecified

In this set of simulations, we set the true, generative k (what we call k^*) as $k^* = 3$ and then examined how the recovery of parameters is affected as we fit the model with $k \in \{2, 3, 6\}$.

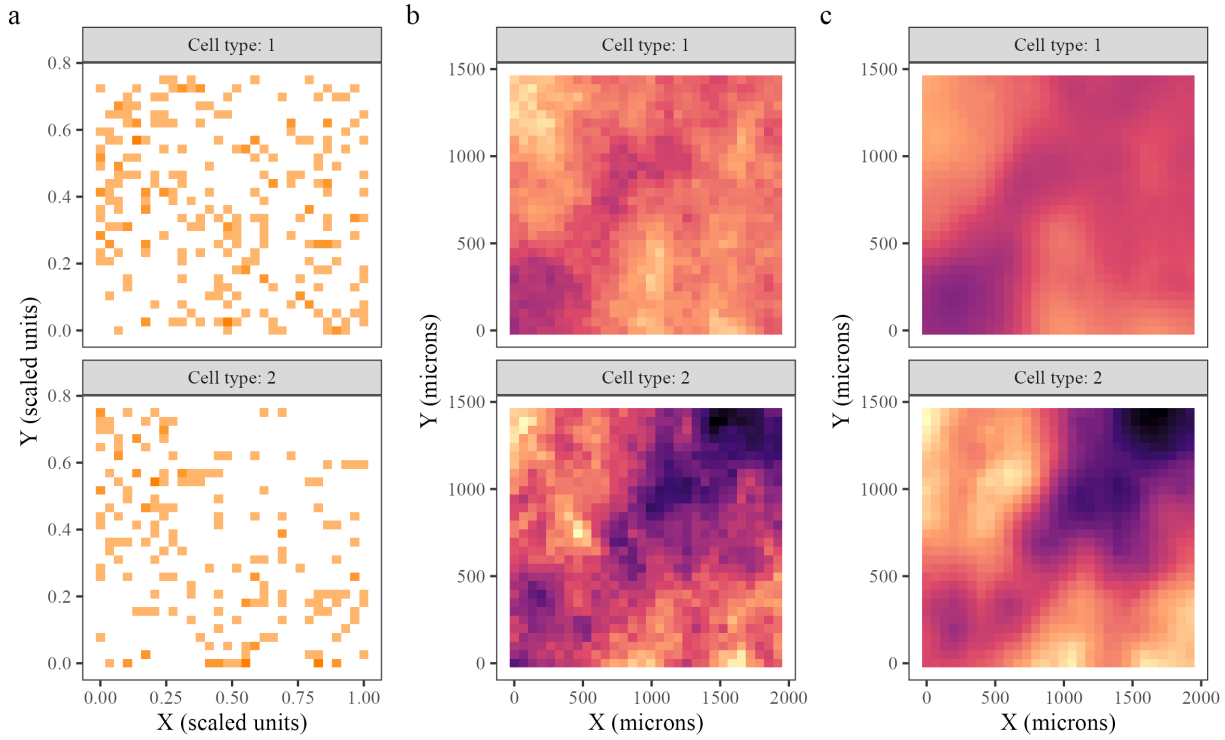


Figure 5.4: Example simulation from study on misspecification of k. a) Simulated data for a single image. b) Underlying latent factors of simulated data. c) Recovered latent factors.

For each k, we simulated 10 datasets of $N = 40$ images each, where each image is a 30×30 pixel grid of 10 cell types, with pixel coordinates for all images from the rectangular domain $[0, 1] \times [0, 0.75]$. We employed an exponential covariance function for the spatial latent factors $\boldsymbol{v}_i(\boldsymbol{l})$, in which the spatial decay φ_j of each factor $v_{i,j}(\boldsymbol{l})$ was sampled uniformly from $[1, 3]$. Lastly, we used an intensity offset of $\alpha_{ij} = -2$ for each $i = 1, \ldots, N$ and $j = 1, \ldots q$. We then simulated from our model–refer to the Supplement for additional details. After generation of ground-truth simulations, we fit our model to these simulations and sampled 5000 MCMC samples after a 5000-sample burn-in period. An example of simulated data can be seen in Figure 5.4a, while an example of the underlying latent factors $\boldsymbol{w}_i(\boldsymbol{l})$ can be seen in Figure 5.4b.

Recovery of the cross-correlation curves was best when $k = k^* = 3$ (Figure 5.5). If k was misspecified, we found that $k > k^*$ yielded cross-correlation curves that were very similar to the true curves (Figure 5.5). Underspecified k had the worst recovery at closer spatial distances. Convergence diagnostics, such as effective sample size (ESS) (C. J. Geyer, 2011) and Gelman's $\hat{R}$ (Vehtari et al., 2021), however, showed that smaller k demonstrated better MCMC convergence, even when $k < k^*$. In general, MCMC convergence of cross-correlation

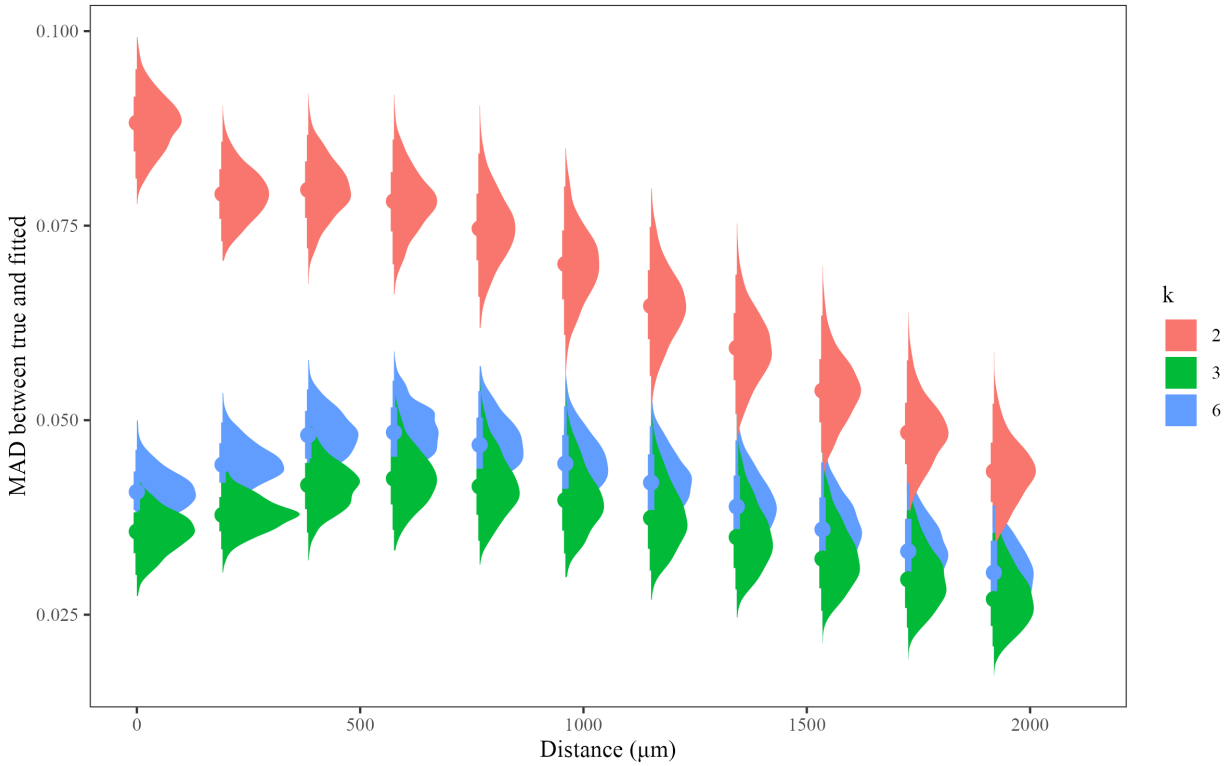


Figure 5.5: MAD of cross-correlation curves at various distances between observed and predicted when fitting with various k. Here the generative model has $k^* = 3$.

estimates was the best for zero-distance cross-correlations. Lastly, we applied the Watanabe-Akaike information criterion (WAIC) (Watanabe, 2010) to models fit under various k and found, again, that underspecified k leads to the worst predicted out-of-sample predictive performance, while fitting with over-specified k has a WAIC that is very similar to that of models fit with k^*. Based on these results, in general we recommend that users perform model fitting at various k and then use the model at a value of k for which the WAIC is no longer improving. Doing so will ensure the best out-of-sample predictive performance, as well as high-fidelity parameter recovery. The Supplement includes additional details and Figures related to our simulation study.

5.4.2 The effect of grid size on cross-correlation recovery

In the next set of simulations, we set $k^* = 4$ and then examined how the recovery of parameters is affected as we fit the model with number of pixels on each dimension $n_x \in 12, 24, 48$. We also set $n_y = n_x$, so that each image is $n_x \times n_x$. For each n_x, we simulated 10 datasets of $N = 40$ images each, where each image was a 48×48 pixel grid of 10 cell types, with pixel coordinates for all images from the rectangular domain $[0, 1] \times [0, 0.75]$. We then coarsened each pixel grid to an $n_x \times n_x$ grid by aggregating neighboring pixels together. We used an intensity offset of $\alpha_{ij} = -3$ for each $i = 1, \ldots, N$ and $j = 1, \ldots q$. We kept the rest of the parameters the same as from the last simulation–specifically, we used spatial latent factors $\boldsymbol{v}_i(\boldsymbol{l})$ with exponential covariance and sampled each spatial decay parameter φ_j from $U(1, 3)$. We then simulated from our model–additional details are in the Supplement. After generation of ground-truth simulations, we fit our model to these simulations and sampled 1000 MCMC samples thinned every 5 draws after a 5000-draw burn-in period. An example of simulated data at various resolutions can be seen in Figure 5.6.

We found that the MAD between observed and fitted was much closer between simulations with different n_x than in the previous simulation suite, and seems to follow a similar trend over spatial distance: increasing up to about 1000 μm and then decreasing (Figure 5.7). Models fit with the largest number of pixels $n_x = 48$ had nominally the best cross-correlation curve recovery. In the Supplement, however, we show that the convergence diagnostics for models fit on data with higher numbers of pixels were significantly worse, and that cross-correlation curves fit on different n_x seem to be visually indistinguishable from one another. Finally, the use of an information criterion was not useful to us in this situation, as the datasets we would be comparing are of different sizes, but we can recommend that users wanting to determine the appropriate number of pixels should run a sensitivity analysis at various pixel count and determine the coarsest number of pixels that they can use without

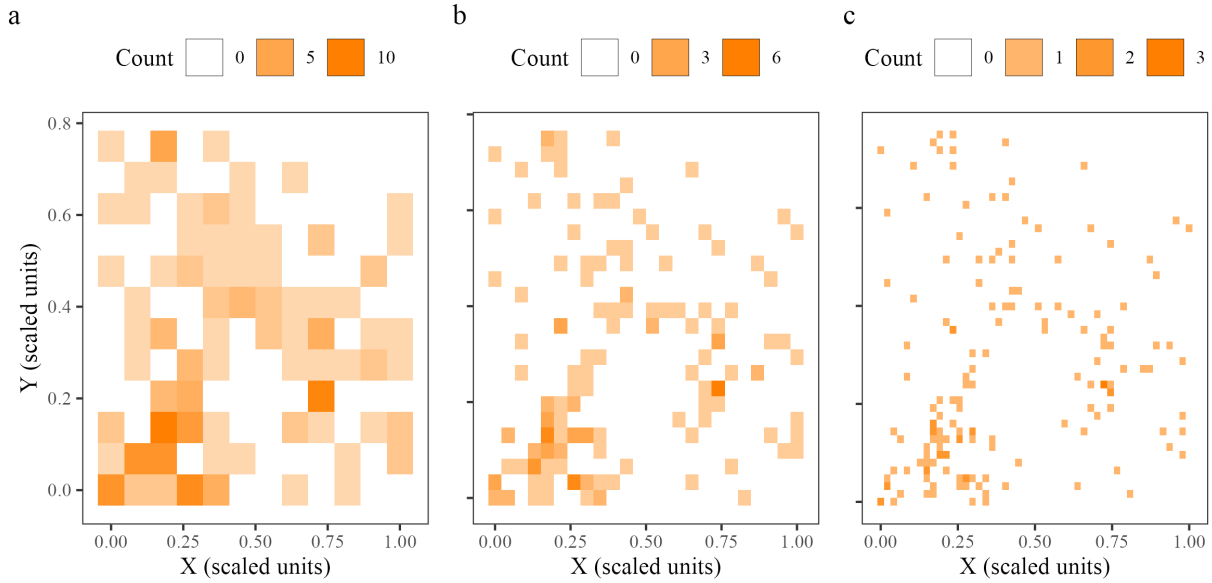


Figure 5.6: Simulated data for a single image with a) $n_x = 12$, b) $n_x = 24$ and c) $n_x = 48$.

affecting the cross-correlation curve recovery too much.

5.5 Application to pancreatic cancer data

We applied our method to the full set of 143 images in the PDAC cohort and 89 images in the IPMN cohort, focusing on key cell types such as regulatory T cells (Tregs), helper T cells, epithelial (tumor) cells, cytotoxic T lymphocytes (CTLs), antigen-presenting cells (APCs), and others. Each image was scaled and divided into a grid of pixels, each of which was approximately $70 \times 70 \mu m$, following our data preprocessing workflow (see Figure 5.2B for an example of a preprocessed image). After preprocessing, we fit our model separately to the PDAC and IPMN cohorts, testing $k \in \{2, 3, 4, 5\}$ factors. We run MCMC for 30,000 iterations, discarding the first 20,000 as burn-in and thinning 10:1 the remaining ones, retaining a total of 1,000 MCMC samples.

WAIC values for both patient groups were lowest for $k = 5$ (Table 5.1). However, results from the $\hat{R}$ diagnostic were mixed - no single model had universally better convergence of cross-correlation curves for all cell type pairs. In both groups, the $\hat{R}$ of the cross-correlation curves for most cell type pairs at $k = 5$ were quite good, indicating acceptable MCMC convergence. Thus, we chose $k = 5$ for both models. The Supplement includes additional details on MCMC diagnostics.

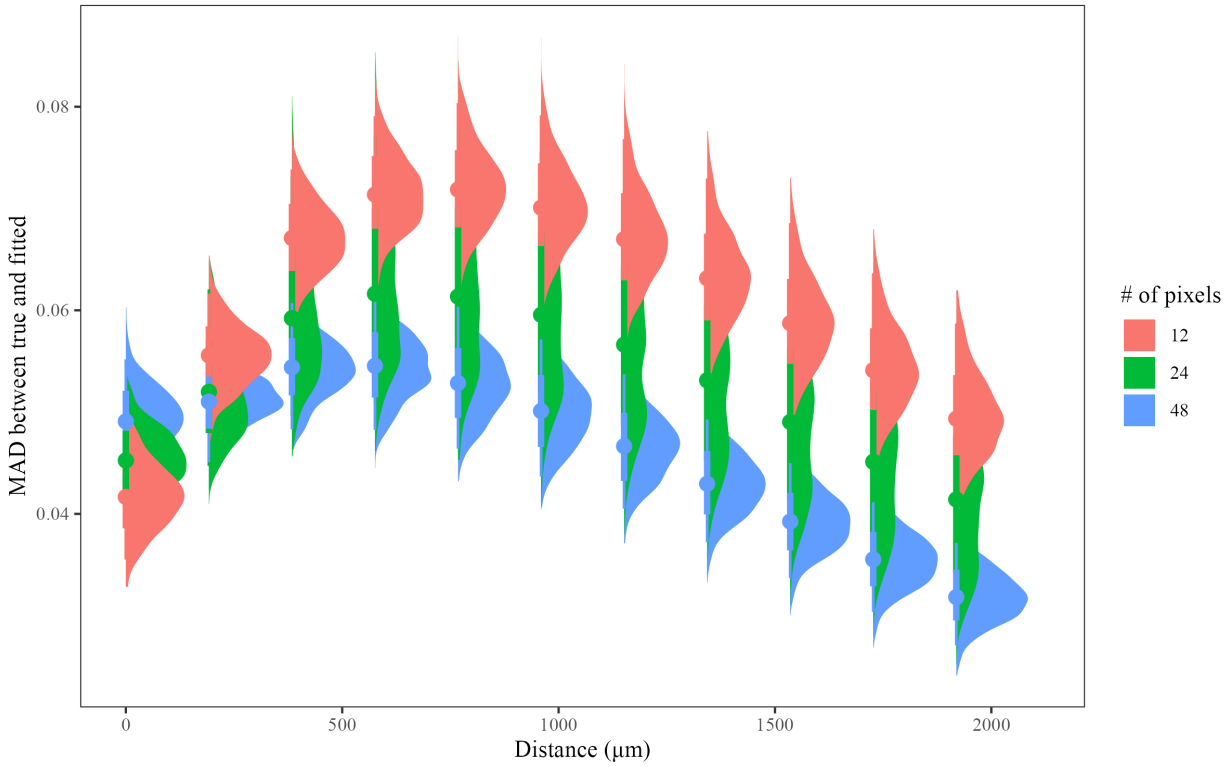


Figure 5.7: MAD of cross-correlation curves at various distances between observed and predicted when fitting with various numbers of pixels.

Table 5.1: The WAIC of the fitted model for various k, for both patient groups in the pancreatic cancer dataset.

	WAIC Scores	
k	**WAIC in PDAC group**	**WAIC in IPMN group**
2	3197	2933
3	3090	2941
4	2958	2894
5	2940	2843

5.5.1 Spatial Organization of the Pancreatic Cancer TME Across Patient Groups

The spatial cross-correlation analysis reveals distinct cellular organization patterns between PDAC and IPMN tumors, highlighting key differences in immune cell spatial distribution (Figure 5.8). In general, PDAC tumors exhibit stronger spatial associations among key immune cell populations, whereas IPMN tumors display weaker correlations, indicating a less structured immune landscape.

Regulatory T Cell Expansion and Clustering in PDAC A striking finding is the significantly higher spatial correlation of Tregs in PDAC tumors compared to IPMN tumors. While Tregs show strong self-association in both groups at close distances, their correlation persists across greater spatial ranges in PDAC, suggesting extensive clustering. This spatial pattern may indicate an immunosuppressive niche within PDAC tumors, where Tregs accumulate to dampen anti-tumor immune responses. Such clustering has been implicated in immune evasion and resistance to immunotherapy in pancreatic cancer (Mota Reyes et al., 2022).

Stronger T Cell Interactions in PDAC Tumors CTLs and helper T cells also display stronger spatial correlations in PDAC compared to IPMN. This pattern suggests that T cells in PDAC tumors are more spatially organized and less dispersed, potentially reflecting their recruitment into structured immune environments. The increased correlation between CTLs and THelpers further suggests coordinated immune activation in PDAC tumors, albeit within a potentially immunosuppressive microenvironment.

Additionally, PDAC tumors exhibit a higher correlation between CTLs and Tregs at moderate to far distances, reinforcing the hypothesis that PDAC fosters a structured but immune-suppressive T cell landscape, where regulatory mechanisms counteract effector re-

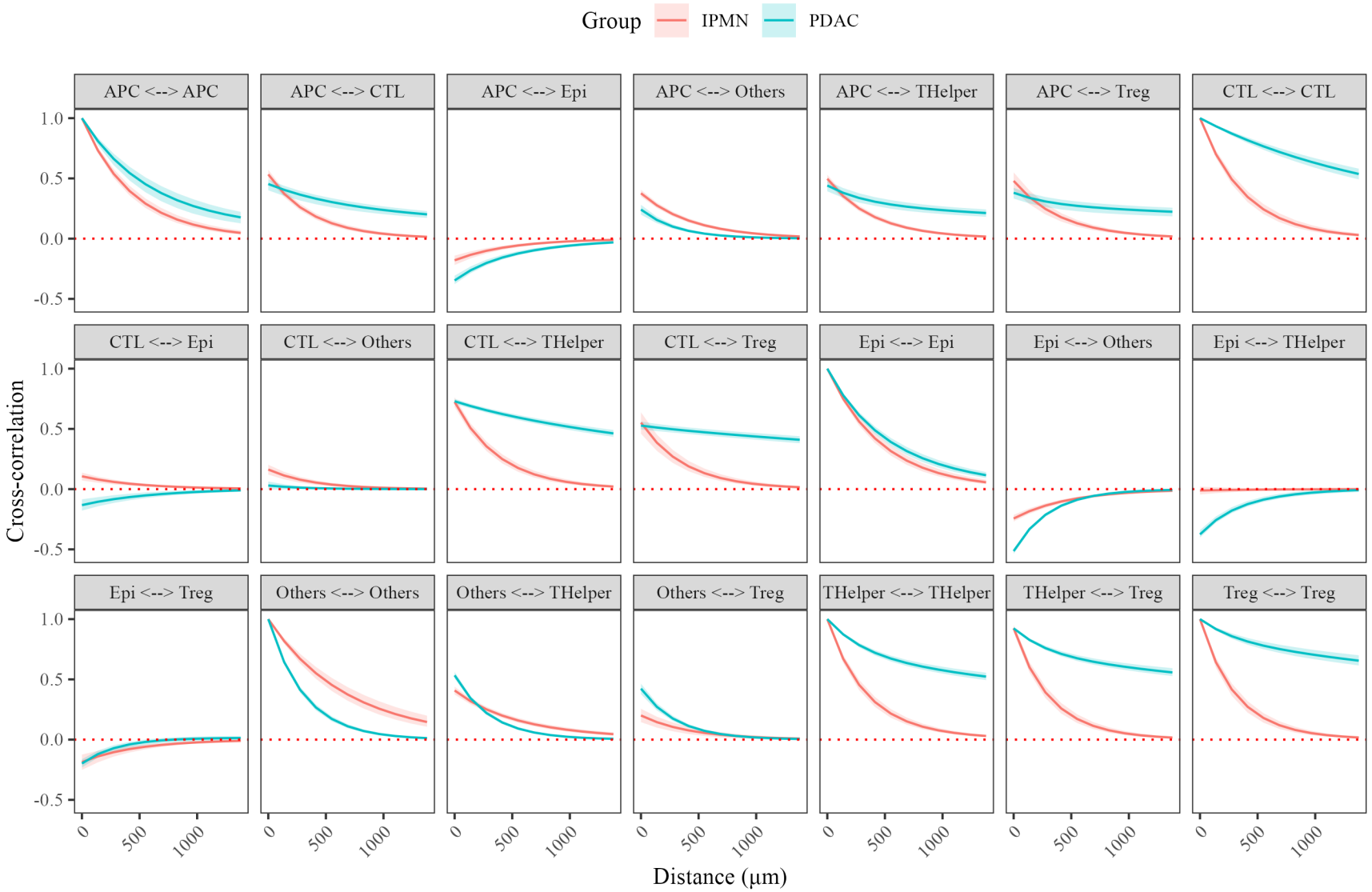


Figure 5.8: The estimated spatial cross-correlations between all pairs of cell-type combinations from patients in the IPMN group and the PDAC group. The colored line in each plot indicates the posterior expectation, while the shaded region represents the pointwise 95% credible interval.

sponses (Goulart et al., 2021).

Divergent Antigen-Presenting Cell Distribution Between Groups APCs show markedly different spatial correlations with T cells between PDAC and IPMN tumors. In PDAC, APCs exhibit stronger associations with CTLs, THelpers, and Tregs, suggesting that these tumors maintain closer interactions between antigen-presenting and lymphoid cells. This organization could support both immunosuppressive and immune-activating roles, depending on the APC phenotype (Poh and Ernst, 2021). The weaker APC-T cell associations in IPMN tumors indicate a less structured immune landscape, which may reflect differences in tumor immunogenicity or microenvironmental constraints on APC activity.

Conclusion Overall, PDAC tumors exhibit a more structured and spatially organized immune microenvironment compared to IPMN tumors. The increased clustering of Tregs, the stronger spatial correlations among T cells, and the heightened APC interactions in PDAC tumors suggest an immune landscape that is both more engaged and more tightly regulated. These findings align with previous observations that PDAC tumors often develop robust immunosuppressive mechanisms to evade immune-mediated tumor clearance. In contrast, the weaker immune correlations in IPMN tumors suggest a less compartmentalized and potentially less immunosuppressive immune landscape, which may have implications for differential treatment strategies targeting the immune microenvironment in pancreatic cancer. Interestingly, however, the spatial relationship between epithelial cells and the other cell types seemed generally much more similar - indicating that the spatial differences between these diseases were largely due to the the distributions of immune cells rather than those of precancerous and cancerous cells.

5.6 Application to CRC data

We removed four images with extensive contiguous missing data (37% to 53%), as the software implementation of our method relies on dividing images into blocks with sufficient non-missing data. We identified all cell types across both patient groups that had a median count in each image greater than 10 and dropped the other cell types from the data. Lastly, we applied our preprocessing workflow to each image by converting each image to a pixel grid with pixels of size $70 \times 70 \mu m$, as described previously (Figure 5.3B). After data preprocessing, we fit our model separately on the images from each cohort with $k \in \{2, 4, 6, 8, 10\}$ for a total of 30,000 MCMC samples, of which 20,000 discarded as burn-in and 1000 retained after 10:1 thinning of the remainder chain.

Table 5.2: The WAIC of the fitted model for various k, for both patient groups in the CRC dataset.

	WAIC Scores	
k	**WAIC in CLR group**	**WAIC in DII group**
2	3121	5945
4	2913	5799
6	2646	5565
8	2564	5471
10	2571	5449

We found that the models with the lowest WAIC were those with higher values of k (Table 5.2). However, such models fit with these higher values of k also demonstrated more problems with MCMC convergence, with $k = 4$ seeming to demonstrate the best MCMC convergence (in terms of $\hat{R}$). The Supplement includes additional details on MCMC fitting, with figures indicating that cross-correlation curves estimated with $k = 2$ are signficantly different than the others, while curves estimated with higher values of k look much more similar. We decided to compromise on these three aspects of model selection and choose $k = 6$ for our final model.

5.6.1 Spatial Organization of the CRC TME Across Patient Groups

The spatial cross-correlation analysis reveals distinct organizational patterns in the TME between CLR and DII patients, reflecting their respective immune-hot and immune-cold phenotypes (Figure 5.9). These differences highlight how immune and stromal cells are distributed differently across patient groups, with CLR tumors exhibiting a more spatially intermixed immune landscape and DII tumors displaying stronger compartmentalization, consistent with an immune-excluded state (Fu et al., 2021).

Immune Exclusion and Stromal Structuring in DII Tumors DII tumors exhibit significantly stronger spatial correlations between plasma cells and vasculature, suggesting that humoral immune responses in these tumors are spatially restricted to perivascular niches. This pattern indicates that plasma cells in DII tumors preferentially accumulate near blood vessels rather than distributing more broadly throughout the tumor microenvironment. Such confinement could limit their interactions with effector immune cells and tumor cells, reinforcing a segregated immune landscape (Nowosad, Marine, and Karras, 2023). Similarly,

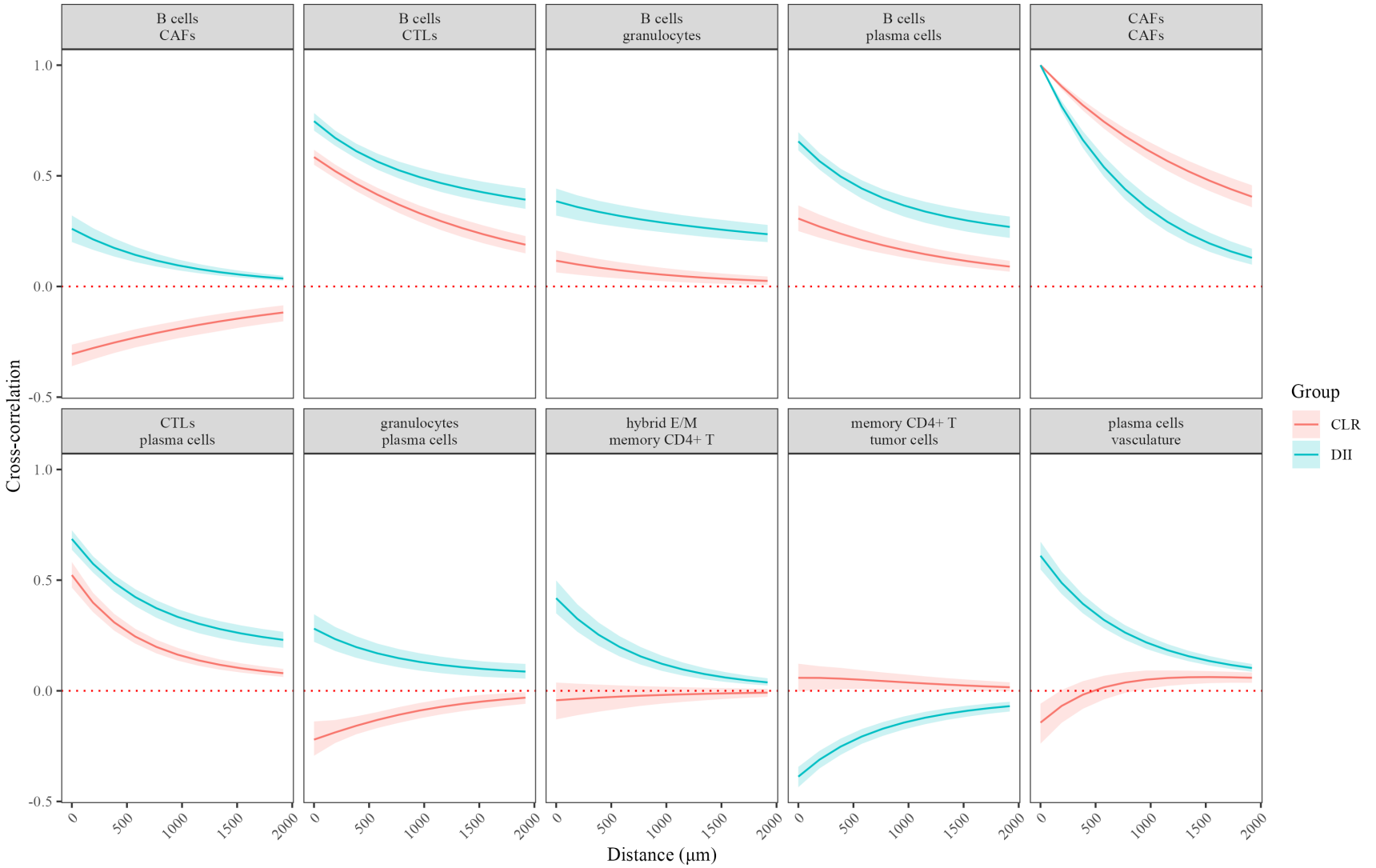


Figure 5.9: The estimated spatial cross-correlations between select pairs of cell-type combinations from patients in the CLR group and the DII group. The colored line in each plot indicates the posterior expectation, while the shaded region represents the pointwise 95% credible interval.

B cells show a much stronger association with cancer-associated fibroblasts in DII tumors, whereas in CLR tumors, this association is exclusionary. This suggests that CAFs in DII patients may be spatially structuring B cell responses, potentially forming barriers that restrict immune infiltration into tumor-rich regions (Mhaidly and Mechta-Grigoriou, 2021).

Granulocytes in DII tumors also exhibit a stronger correlation with plasma cells and B cells, which may indicate spatial compartmentalization of humoral immunity and myeloid-driven immune suppression. Given the known tumor-supportive roles of tumor-associated neutrophils (TANs), this spatial pattern could reflect an immune landscape where granulocytes are suppressing adaptive immune responses rather than promoting direct anti-tumor activity (L. Wu and X. H.-F. Zhang, 2020). Hybrid epithelial-mesenchymal cells in DII tumors also show significantly stronger correlations with memory CD4+ T cells, a pattern that may indicate localized immune suppression where hybrid E/M cells contribute to immune evasion by modifying their surrounding microenvironment (Fedele and Melisi, 2020).

Notably, cytotoxic T lymphocytes exhibit stronger spatial correlations with plasma cells and B cells in DII tumors compared to CLR tumors. This suggests that in the immune-excluded environment of DII tumors, CTLs are more confined to lymphoid aggregates or peritumoral regions rather than being evenly distributed across the tumor microenvironment. This restricted pattern may reflect an immune response that is less capable of directly engaging tumor cells, reinforcing the immune-excluded phenotype of DII tumors (Kinker et al., 2021; Wortman et al., 2021).

More Spatial Mixing and Less Exclusion in CLR Tumors In contrast to the spatial segregation observed in DII tumors, CLR tumors exhibit a more intermixed immune and stromal landscape, where spatial correlations between cell types are generally closer to zero. This suggests that CLR tumors lack the pronounced immune compartmentalization seen in DII tumors, enabling more uniform mixing of immune and stromal populations. For example, memory CD4+ T cells in CLR tumors exhibit weaker spatial correlation with vasculature and CAFs compared to DII tumors, suggesting that helper T cells are more evenly distributed throughout the tumor rather than being concentrated in specific stromal niches. This well-mixed distribution could enhance immune surveillance and support cytotoxic responses in CLR patients (A. T. Baker et al., 2021; M. Chen et al., 2024).

In contrast to most other stromal-immune associations, CAFs in CLR tumors exhibit a notably stronger self-association than in DII tumors, suggesting that fibroblasts in CLR patients are more spatially structured. This pattern may indicate that certain CAF populations in CLR tumors play an active role in reinforcing immune-permissive microenvironments, potentially supporting tertiary lymphoid structure (TLS) formation or immune cell recruitment

(M. Chen et al., 2024; Gunaydin, 2021; Guo et al., 2023). Rather than forming exclusively physical barriers, as they appear to in DII tumors, CAFs in CLR tumors may contribute to organized immune-stromal interactions that facilitate adaptive immune responses.

Interestingly, while CLR tumors generally do not exhibit significantly stronger immune cell interactions in the way that DII tumors do, they show less exclusionary spatial patterns between immune and tumor cells. For instance, memory CD4+ T cells exhibit a much less negative spatial correlation with both tumor cells and hybrid E/M cells in CLR tumors, suggesting that helper T cells are not being actively excluded from tumor-rich regions to the same extent as in DII tumors. Rather than forming segregated zones, the immune and tumor compartments in CLR tumors appear to be more diffusely distributed, supporting a microenvironment where immune cells may have better access to engage tumor cells directly.

Conclusion The spatial correlation patterns between immune and stromal components highlight the fundamental differences between CLR and DII tumors. While DII tumors exhibit strong compartmentalization, particularly through structured immune-stromal associations, CLR tumors show a more intermixed spatial organization where immune and tumor cells are less strictly segregated. This suggests that CLR tumors may support greater immune cell mobility and accessibility to tumor regions, potentially facilitating more effective anti-tumor immune responses.

5.7 Discussion

In this work, we developed a novel Bayesian hierarchical framework for jointly modeling spatial dependencies in multiplexed imaging data across multiple subjects. Our methods provide a principled statistical approach to pool information across patients and infer cell-type-specific spatial co-location patterns. The interpretable output from our model is a set of disease-specific spatial correlation functions that quantify how strongly different cell types associate spatially over varying distances. We applied our methods to UM-PCC and CRC data, revealing novel insights into immune infiltration patterns, immune suppression, TME organization, and the role of spatial organization in cancer progression and immune evasion.

Specifically, on pancreatic cancer data, our methods revealed a stark contrast in immune cell spatial patterns between IPMN and PDAC patients. In IPMN, a more well-mixed immune microenvironment suggests a relatively intact immune surveillance system. Conversely, in PDAC, immune exclusion was evident through the spatial segregation of T cells together away from tumor cells. This supports the hypothesis that tumor progression is associated with restructuring of the immune microenvironment to evade immune detection (Jamouss

et al., 2025). Such findings align with prior work suggesting that PDAC fosters an immunosuppressive niche that inhibits effective anti-tumor responses (Hernandez et al., 2022). Similarly, our analysis of CRC data highlighted significant differences between the CLR and DII patient groups. CLR patients exhibited a more integrated immune microenvironment, consistent with the presence of tertiary lymphoid structures and enhanced immune surveillance (Maoz, Dennis, and Greenson, 2019). In contrast, the observed interactions between CAFs and immune cells in DII patients further underscored the complex crosstalk shaping tumor immune landscapes, particularly in promoting immune suppression and tumor progression (Dang et al., 2023).

Future research will address several open challenges. First, our approach assumes a common spatial covariance structure across all subjects within a given disease group. While this assumption facilitates information sharing, it may not fully capture inter-patient heterogeneity. Future work could explore hierarchical extensions that allow partial pooling of spatial parameters, better accommodating variations in spatial dependencies across individuals. Methodological advancements in this direction might build on recent developments in Bayesian factor modeling with multi-study data (De Vito et al., 2019; Mauri, Anceschi, and Dunson, 2025). Second, our model assumes that the spatial patterns of all cell types can be represented via the same set of spatial factors, leading to inflexibility in modeling cell-type-specific spatial characteristics. Recent developments in more flexible cross-covariance function models (Peruzzi, 2024) may help resolve these limitations. Third, our model can be extended via integration of covariates, such as patient-specific information, directly into the covariance model. Covariate-dependent spatial correlation curves have potential implications for therapeutic targeting.

Finally, our approach is generally applicable to other spatially resolved omics technologies, such as spatial transcriptomics or mass cytometry imaging. These modalities provide complementary information about gene expression and protein abundance, enabling a multimodal characterization of the tumor microenvironment.

CHAPTER 6

Bayesian Survival Modeling with Functional Spatial Covariates

6.1 Introduction

Cancer is a leading cause of death worldwide, and patient outcomes are increasingly understood to be shaped by the tumor microenvironment (American Cancer Society, 2024; Visser and Joyce, 2023). Recent advances in multiplexed imaging have enabled high-resolution spatial profiling of cells within tumor tissues, offering new opportunities to study the spatial organization of immune and tumor cells (Dam, Baars, and Vercoulen, 2022). Spatial point pattern analysis has become an important tool for quantifying these patterns, with summary statistics such as the nearest-neighbor distribution function, or G-cross function, providing interpretable measures of spatial dependence between cell types (Eliason, S. Krishnan, et al., 2025; Wrobel, Harris, and Vandekar, 2023; Behanova, Klemm, and Wählby, 2022; Benimam et al., 2025). These functions can reveal how patterns of proximity and avoidance may reflect underlying biological interactions relevant to tumor progression and treatment response.

Despite their growing use, applications of spatial summary statistics in survival analysis often rely on ad-hoc or restrictive modeling strategies. A common approach involves selecting a small number of spatial radii at which to summarize cell-cell interactions, which may overlook important information contained in the full interaction curve (Tsang et al., 2024; Eliason, S. Krishnan, et al., 2025). In addition, such spatial covariates are typically treated as fixed, even though they are subject to sampling variability due to the finite number of cells observed, measurement noise, and local tissue heterogeneity. Moreover, spatial data are often collected at multiple sites within a tumor or across multiple biopsies per patient, giving rise to a hierarchical structure that is frequently ignored. These limitations can compromise the interpretability and robustness of downstream survival analyses.

We introduce a Bayesian multi-level model for survival analysis that incorporates spatially derived functional covariates while accounting for measurement uncertainty. Rather

than summarizing spatial information at a few pre-selected radii, the model includes the full spatial interaction curve as a covariate by estimating a spatially varying regression coefficient function, $\beta(s)$. This allows different spatial scales to contribute flexibly to the hazard. To account for uncertainty in the estimated spatial covariates, we model sampling error based on a resampling approach. Specifically, we use a bootstrap procedure in which points of one type are resampled while the other type is held fixed, generating variability in the estimated G-cross function. The standard deviation of these replicates at each location provides an estimate of finite-sample variability, which is used as a known error term in the model. Furthermore, the framework supports multiple measurements per subject by modeling latent functional covariates within a hierarchical structure. This allows information to be partially pooled across measurements while preserving subject-level variability. Together, these components enable a principled analysis of spatial covariates and their association with survival outcomes.

The remainder of this paper is structured as follows: In Section 6.2, we introduce the methodology, describing the functional covariate modeling, hierarchical structure, and survival model. Section 6.3 presents simulation studies to evaluate model performance under different spatial structures and levels of measurement error. Section 6.4 applies our model to a real-world cancer imaging dataset, demonstrating its ability to improve survival predictions. Finally, Section 6.5 discusses the implications of our findings and potential directions for future research. We provide an open-source implementation of this model at github.com/jeliason/frame. All code used to generate the figures and analyses presented in this manuscript is available in the same repository.

6.2 Methods

6.2.1 Functional Covariate Construction and Uncertainty Quantification

A key feature of our model is the incorporation of uncertainty in the functional covariate derived from spatial summary statistics. While the framework generalizes to arbitrary spatial functions, we focus here on the *G-cross function*, a first-order summary statistic from spatial point pattern analysis used to quantify the proximity between two point types (A. Baddeley, Rubak, and Turner, 2015).

The G-cross function $G_{AB}(s)$ estimates the cumulative distribution function of distances

from points of type A to their nearest neighbors of type B, evaluated at distance s:

$$G_{AB}(s) = \mathbb{P}\left(\text{dist}(A \to B) \leq s\right),$$

where $\text{dist}(A \to B)$ is the distance from a typical type A point to its nearest type B neighbor. The function is bounded between 0 and 1 and reflects the degree of spatial clustering of type B around type A, with larger values indicating stronger proximity.

To propagate uncertainty in the estimated $G_{AB}(s)$ curves into the survival model, we employ a nonparametric bootstrap approach based on resampling type A points. For each bootstrap replicate b, we compute $\hat{G}_{AB}^{(b)}(s)$, and the standard deviation across replicates quantifies sampling variability:

$$\hat{\tau}_{ij}(s) = \text{SD}\left(\hat{G}_{AB}^{(b)}(s)\right).$$

6.2.2 Modeling the Spatially-Integrated Functional Covariate

Let $X_i(s)$ denote the latent, spatially-varying functional covariate for subject i at spatial location s, modeled via a B-spline expansion $X_i(s) = B_X(s)\theta_{X_i}$, where $B_X(s)$ is a row vector of B-spline basis functions and θ_{X_i} are subject-specific spline coefficients. To impose smoothness while allowing flexible variation, we place a second-order random walk prior (RW2) on the coefficients:

$$\theta_{X_i,k} - 2\theta_{X_i,k-1} + \theta_{X_i,k-2} \sim \mathcal{N}(0, \sigma^2_{\theta_X}).$$

Similarly, the spatially-varying regression coefficient function $\beta(s)$ is represented via its own B-spline expansion $\beta(s) = B_\beta(s)\theta_\beta$ with a second-order random walk prior:

$$\theta_{\beta,k} - 2\theta_{\beta,k-1} + \theta_{\beta,k-2} \sim \mathcal{N}(0, \sigma^2_{\beta,k}).$$

Each subject contributes M_i repeated measurements, indexed by j, each corresponding to a different sample (e.g., a distinct histological section). Observed curves $W_{ij}(s)$ are modeled as noisy realizations of $X_i(s)$ plus a smooth nuisance function:

$$W_{ij}(s) = X_i(s) + f_{ij}(s) + \epsilon_{ijs},$$

where $f_{ij}(s) = B_{\text{noise}}(s)\eta_{ij}$, with $\eta_{ij} \sim \mathcal{N}(0, \sigma^2_{\text{nuisance}} I)$, and $\epsilon_{ijs} \sim \mathcal{N}(0, \tau^2_{ij}(s))$.

Marginalizing over η_{ij} gives:

$$W_{ijs} \sim \mathcal{N}\Big(X_i(s),\ \tau_{ij}^2(s) + \sigma_{\text{nuisance}}^2 \cdot \|B_{\text{noise}}(s)\|^2\Big),$$

where $\|B_{\text{noise}}(s)\|^2 = B_{\text{noise}}(s)B_{\text{noise}}(s)^\top$.

6.2.3 Survival Modeling and Bayesian Inference

We model subject-level survival times using a Cox proportional hazards model, which relates covariates to the hazard function without specifying the baseline hazard. Let T_i denote the survival time for subject i, with event indicator $\delta_i \in \{0, 1\}$, where $\delta_i = 1$ indicates an observed event and $\delta_i = 0$ indicates right censoring. The hazard function under the Cox model is $h(t \mid \mu_i) = h_0(t)\exp(\mu_i)$, where $h_0(t)$ is the unspecified baseline hazard and μ_i is the linear predictor defined as:

$$\mu_i = \sum_{s=1}^{S} X_i(s)\beta(s)\Delta s,$$

with $X_i(s)$ representing the latent functional covariate at spatial location s, $\beta(s)$ the spatially varying regression function, and Δs the spacing between locations.

The Cox model assumes that the effect of the covariates on the hazard is multiplicative and time-invariant. The regression surface $\beta(s)$ can thus be interpreted in terms of the log hazard ratio (log-HR) contribution from each spatial location.

The partial likelihood for the Cox model is used in place of the full likelihood:

$$\mathcal{L}_{\text{partial}}(\beta) = \prod_{i:\delta_i=1} \frac{\exp(\mu_i)}{\sum_{j\in\mathcal{R}(T_i)}\exp(\mu_j)},$$

where $\mathcal{R}(T_i)$ is the risk set at time T_i.

6.2.4 Joint Posterior Distribution

The full Bayesian model jointly estimates the latent subject-level functional covariates, the spatially-varying regression surface, measurement error parameters, and survival parameters. Letting W_{ijs} denote the observed functional covariate values and $\hat{\tau}_{ij}^2(s)$ the associated variances, the joint posterior is:

$$
\begin{aligned}
p(\theta_X, \theta_\beta, \sigma_{\theta_X}, \sigma_{\beta,k}, \sigma_{\text{nuisance}} \mid W, \hat{\tau}^2_{ij}(s), T, \delta) \propto & \left[\prod_{i=1}^{N} p(\theta_{X_i} \mid \sigma_{\theta_X})\right] \cdot p(\theta_\beta \mid \sigma_{\beta,k}) \\
& \cdot p(\sigma_{\theta_X}) \cdot p(\sigma_{\beta,k}) \cdot p(\sigma_{\text{nuisance}}) \\
& \cdot \prod_{i=1}^{N} \prod_{j=1}^{M} \prod_{s=1}^{S} \mathcal{N}\Big(W_{ijs} \mid X_i(s),\ \hat{\tau}^2_{ij}(s) + \\
& \qquad \sigma^2_{\text{nuisance}} \cdot \|B_{\text{noise}}(s)\|^2\Big) \\
& \cdot \mathcal{L}_{\text{partial}}(\beta).
\end{aligned}
$$

Posterior inference is performed using Hamiltonian Monte Carlo as implemented in Stan. Prior distributions are listed in Table C.1.

6.3 Simulation Studies

Synthetic datasets were generated from the model in Section 6.2. For each subject, we sampled latent curves $X_i(s)$ and $\beta(s)$, computed $\mu_i = \int X_i(s)\beta(s)\,ds$, approximated by discrete sums, and simulated survival times from a Weibull distribution with shape $a = 2$ and scale $\lambda_i = \exp(\mu_i) \cdot 50$. Censoring times C_i were sampled from Uniform$(20, 200)$, with $T_i^{\text{obs}} = \min(T_i, C_i)$ and $\delta_i = \mathbb{I}(T_i \leq C_i)$.

Noisy observed curves $W_{ij}(s)$ were simulated by adding smooth nuisance functions and Gaussian noise to $X_i(s)$, with standard deviations $\tau_{ij}(s) = \log(1 + \exp(B_\tau(s)\theta_{\tau_{ij}}))$, where $\theta_{\tau_{ij}} \sim \mathcal{N}(\mu_\tau, \sigma^2_\tau)$.

We conducted one simulation study in which we varied the number of patients and repeated measurements and are planning another in which we vary the magnitude of nuisance and measurement noise. For each setting, we evaluated recovery of the spatially-varying effect $\beta(s)$ and subject-specific latent covariates $X_i(s)$. Estimation accuracy was assessed using the root mean squared error (RMSE) between the posterior mean and ground truth. For uncertainty quantification, we constructed simultaneous 95% credible bands and computed empirical coverage (the proportion of replications in which the true function was entirely contained within the band) and average band width (a proxy for interval precision). Together, these metrics summarize both point estimation quality and posterior calibration.

6.3.1 Effect of Sample Size on Model Efficiency

To examine the influence of dataset size, we varied the number of patients ($N \in \{20, 50, 100\}$) and the number of repeated measurements per patient ($J \in \{1, 2, 4\}$). All other simulation components were held constant. For each setting, we generated 100 replications. Each model was fit with 1000 warmup iterations followed by 1000 sampling iterations. Examples of recovered $X_i(s)$ and $\beta(s)$ are shown in Figures C.1 and C.2, respectively.

We observed distinct effects of patient- and image-level replication on estimation accuracy (Figure 6.1). Increasing the number of patients led to consistent improvements in recovering $\beta(s)$, reflecting improved population-level inference. In contrast, increasing the number of repeated measurements per patient primarily improved recovery of the subject-specific latent functions $X_i(s)$, with diminishing returns beyond two to four repeated curves. Similarly, interval width for $\beta(s)$ improves primarily with more patients (Figure C.3), while interval width for $X_i(s)$ decreases mainly with the number of repeated measurements (Figure C.4), highlighting the benefit of within-subject replication for individual-level uncertainty quantification. These trends are consistent with the hierarchical model structure, which pools information differently across population and individual levels. Lastly, we evaluated posterior calibration by computing the empirical coverage of 95% simultaneous credible intervals across 100 replications (Table C.2). Coverage for the spatial effect $\beta(s)$ was quite conservative, with coverages at 100% across all settings. Coverage for the patient-level functional covariate $X_i(s)$ was well-calibrated, with median values between 90% and 95%. These results demonstrate that the model's uncertainty quantification remains reliable (and, in the case of $\beta(s)$, somewhat too conservative) across a range of dataset sizes and measurement depths.

6.4 Results

6.4.1 Basel Breast Cancer Dataset

We applied our method to the publicly available Basel breast cancer dataset from Jackson et al. (2020), which contains high-dimensional single-cell imaging data from 281 patients at the University Hospital Basel. For each patient, a single 0.8-mm tumor core was profiled using multiplexed immunofluorescence staining across 29 markers and 4 morphological features. Cells were initially clustered into 71 phenotypic groups and subsequently merged into 27 biologically interpretable metaclusters representing tumor, stromal, immune, and endothelial compartments. Following preprocessing, we further collapsed these into major functional categories, including *Luminal-like Tumor*, *Basal-like Tumor*, *Proliferative Tumor*, *CAF-like*, *Fibroblasts*, *Macrophages*, *T cells*, *B cells*, and *Endothelial* cells. Each tumor core typically

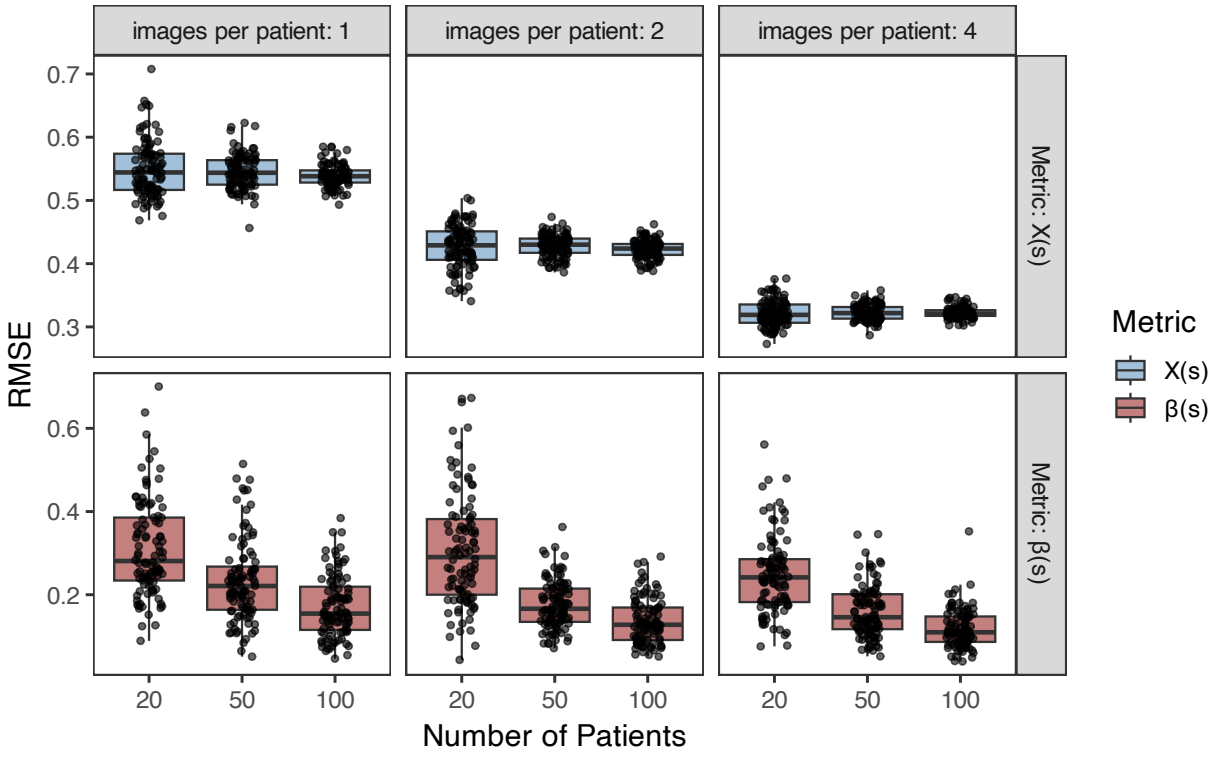


Figure 6.1: RMSE for recovery of the spatially-varying regression coefficient $\beta(s)$ and subject-specific latent functions $X_i(s)$ across varying numbers of patients and repeated measurements per patient. Each panel shows the distribution of RMSE across 100 simulated datasets. Estimation of $\beta(s)$ improves with increasing cohort size, while estimation of $X_i(s)$ benefits primarily from additional repeated measurements.

contained thousands of individual cells with spatial coordinates and phenotype labels.

6.4.2 Estimation of Functional Covariates

To quantify cell-type-specific spatial interactions in this dataset, we estimated the G-cross function between all pairs of cell types within each image core. The G-cross function, $G_{AB}(s)$, measures the cumulative distribution function of distances from a typical point of type A to its nearest neighbor of type B, providing a first-order summary of spatial proximity between cell types (A. Baddeley, Rubak, and Turner, 2015).

For each image, we constructed a marked spatial point pattern using cell coordinates and cell type labels. Cell types with fewer than four observed cells were excluded to ensure we could estimate the G-cross function. We then computed $G_{AB}(s)$ over a discretized distance grid from 0 to 80 microns in 5-micron increments for each valid type pair (A, B). To quantify uncertainty in the estimated spatial summary curves, we applied the nonparametric

bootstrap procedure described in Section 6.2.1 to derive the empirical estimate for $\hat{\tau}_{ij}(s)$.

6.4.3 Model Fitting and Posterior Inference

For each cell-type pair (A, B), we fit the proposed hierarchical Bayesian model to the Basel dataset using the `cmdstanr` interface to Stan (Stan Development Team, 2024; Gabry et al., 2024). The observed functional covariates—G-cross curves computed at 5-micron intervals from 0 to 80 microns—were standardized before model fitting. The posterior distribution was approximated using Hamiltonian Monte Carlo (HMC) with 1000 warmup and 1000 sampling iterations. Simultaneous credible intervals were computed using maximum standardized deviation bands, as detailed in Supplement C.3.1. In parallel, we fit a Cox proportional hazards model using selected values of the standardized G-cross curve (at $s = 20, 40, 60$ microns) as covariates, providing a benchmark for comparison against the proposed Bayesian functional model.

6.4.4 Visualization of Spatial Covariate and Effect Function Estimates

Figure 6.2 displays the estimated spatially varying regression functions $\beta(s)$ (in red) from our Bayesian functional survival model for selected pairs of cell types. Each panel corresponds to a directional relationship from a type 2 (neighboring) population to a type 1 (focal) population, with the spatial scale s ranging from 10 to 60 microns. The red solid line denotes the posterior mean of $\beta(s)$, and the shaded ribbon represents a simultaneous 95% credible interval. Overlaid in black are the point estimates and 95% confidence intervals from conventional Cox proportional hazards models fit using three discrete covariates $X_i(s)$ at spatial radii $s = 20, 40, 60$ microns.

The function $\beta(s)$ represents the log hazard ratio associated with a unit increase in the spatial summary covariate $X_i(s)$, which captures the degree of spatial interaction between the two cell types at distance s. Positive values of $\beta(s)$ indicate that increased proximity at that spatial scale is associated with higher hazard (worse prognosis), while negative values suggest a protective effect. For instance, the panel labeled *Epithelial Tumor* $\rightarrow$ *Epithelial Tumor* shows an increasing $\beta(s)$ at larger distances, suggesting that the presence of epithelial tumor cells near other epithelial tumor cells beyond 40 microns is associated with worse survival. Conversely, the *Macrophages* $\rightarrow$ *CAF-like* panel exhibits negative coefficients at longer spatial scales, potentially reflecting immune activity linked to improved outcomes. Other panels, such as *T cells* $\rightarrow$ *T cells*, show $\beta(s)$ near zero, indicating no detectable prognostic effect from clustering alone.

This functional modeling approach enables a nuanced view of how cell-cell spatial relationships contribute to survival outcomes across spatial scales, as well as smoothly quantifying uncertainty in those estimates across scales. Compared to the Cox model estimates at fixed radii, the Bayesian model provides smoother inference, pooling information across distances to better detect trends. In some cases, the Cox estimates align well with the Bayesian posterior (e.g., *T cells* → *Macrophages*), while in others they miss important features, such as the increasing trend at shorter scales seen in *Macrophages* → *Macrophages*. This highlights the value of modeling the full spatial interaction curve rather than relying on preselected distances.

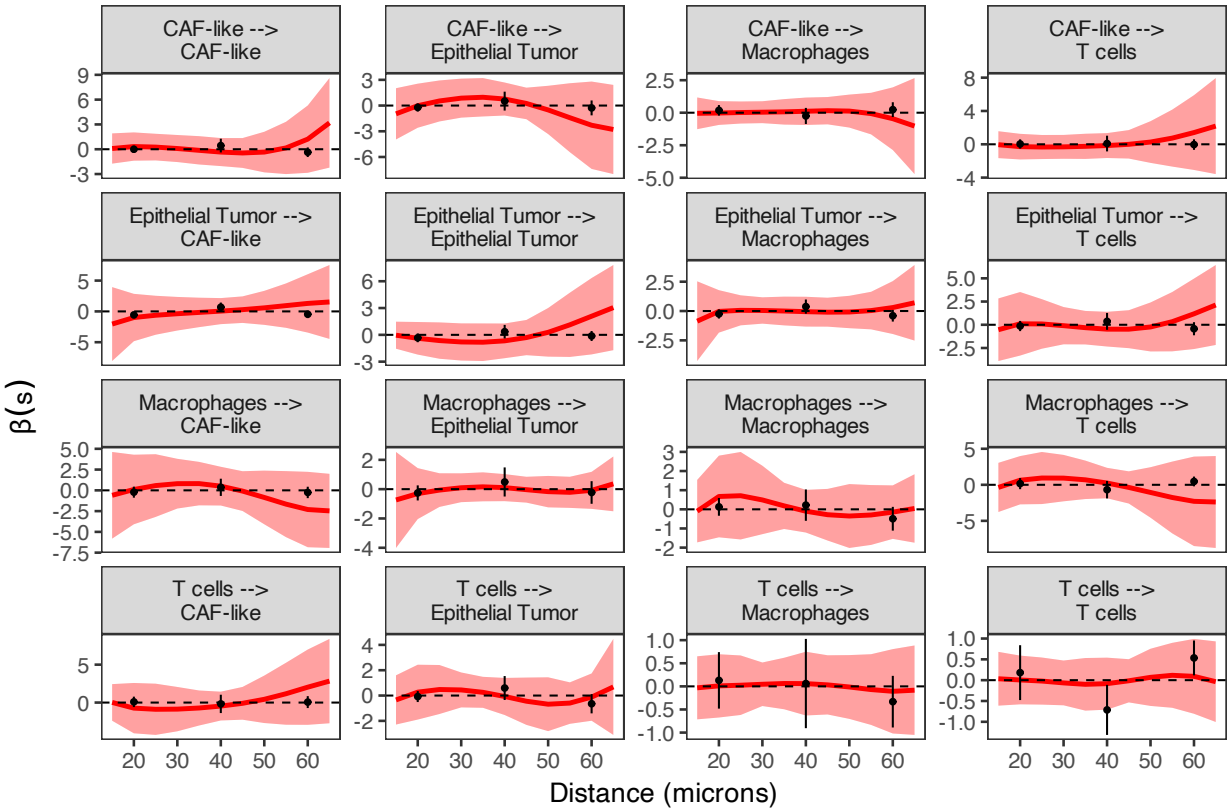


Figure 6.2: For each pair of cell types, posterior means and simultaneous 95% credible intervals were calculated for each spatial location using a maximum standardized residual approach. For a subset of cell type pairs, posterior summaries were plotted as red curves with ribbons. For comparison, we fit standard Cox models using three sampled spatial locations from the G-cross curves ($s = 20, 40, 60$ microns) as covariates, estimating pointwise log-hazard ratios and confidence intervals. These are overlaid as black point-ranges. Facets are organized by neighboring (Type 2) and focal (Type 1) cell type pairs.

Figure 6.3 displays the estimated latent G-cross functions $X_i(s)$ for selected directional cell-type interactions across six representative patients. The dashed bands show pointwise 2-

standard deviation bounds reflecting finite-sample variability in the observed G-cross curves. The Bayesian posterior intervals are consistently at least as wide, capturing not only this measurement error but also broader uncertainty in the entire functional estimate.

In patients with multiple biopsies containing the same pair of cell types, the model combines information across replicates to produce a single, patient-level estimate. For instance, patient 16 has two observed G-cross curves for the *Epithelial Tumor* $\rightarrow$ *Macrophages* interaction that differ in both shape and magnitude; the estimated posterior mean curve reflects a smooth compromise between them. Patient 104 shows similar variability across samples for the *Macrophages* $\rightarrow$ *Macrophages* interaction, and patient 205 exhibits divergent measurements for *Macrophages* $\rightarrow$ *T cells*. In each of these examples, the hierarchical model leverages the structure imposed by the B-spline basis to pool information across repeated measurements, resulting in a more stable estimate of the latent spatial covariate $X_i(s)$, while still capturing the associated uncertainty through its credible interval.

6.5 Discussion

In this work, we developed a Bayesian hierarchical model that incorporates spatially-derived functional covariates into survival analysis, with a focus on the G-cross function. This framework allows for the full interaction curve to influence survival outcomes, while also accounting for sampling variability and hierarchical data structure. By modeling the entire G-cross function as a smooth spatial covariate and estimating a corresponding spatially-varying effect function $\beta(s)$, our method captures heterogeneity in the prognostic value of cell-cell interactions across spatial scales. Through a propagation of measurement error via bootstrap-based estimates, the model yields interpretable estimates of both spatial effects and subject-level interaction curves with calibrated uncertainty.

Simulation results demonstrated that the model performs well across a range of sample sizes and replication levels. Increasing the number of patients improved estimation of the population-level spatial effect $\beta(s)$, while repeated measurements per subject helped recover individual-level latent covariates $X_i(s)$. Importantly, the model's simultaneous credible intervals exhibited appropriate coverage, supporting the validity of its uncertainty quantification even under moderate noise levels.

When applied to the Basel breast cancer dataset, our approach recovered interpretable spatial interaction patterns that were not readily captured by conventional Cox models using discretely sampled covariates. The functional effect surfaces $\beta(s)$ revealed scale-dependent associations between cell-cell proximity and patient outcomes, offering richer biological insights than fixed-radius models. The model also flexibly integrated repeated spatial sum-

maries across biopsies, yielding stable patient-level curves that respected finite-sample uncertainty while capturing meaningful trends. This feature is particularly important in cancer studies where tissue heterogeneity and sampling variation are common.

More broadly, our findings suggest that fully functional spatial summaries—when paired with appropriate error modeling and hierarchical structure—can serve as powerful predictors in clinical survival models. The proposed framework can be extended to other spatial summary functions, imaging modalities, or disease contexts, and supports future integration with downstream biological interpretation or therapy design.

Future work could explore automatic selection of basis complexity, alternative priors such as sparsity-inducing or structured shrinkage priors, or more flexible modeling of the sampling distribution of spatial functions. Overall, our results highlight the importance of careful statistical treatment of spatial covariates in survival analysis, and provide a path forward for uncertainty-aware modeling of complex spatial data in cancer research.

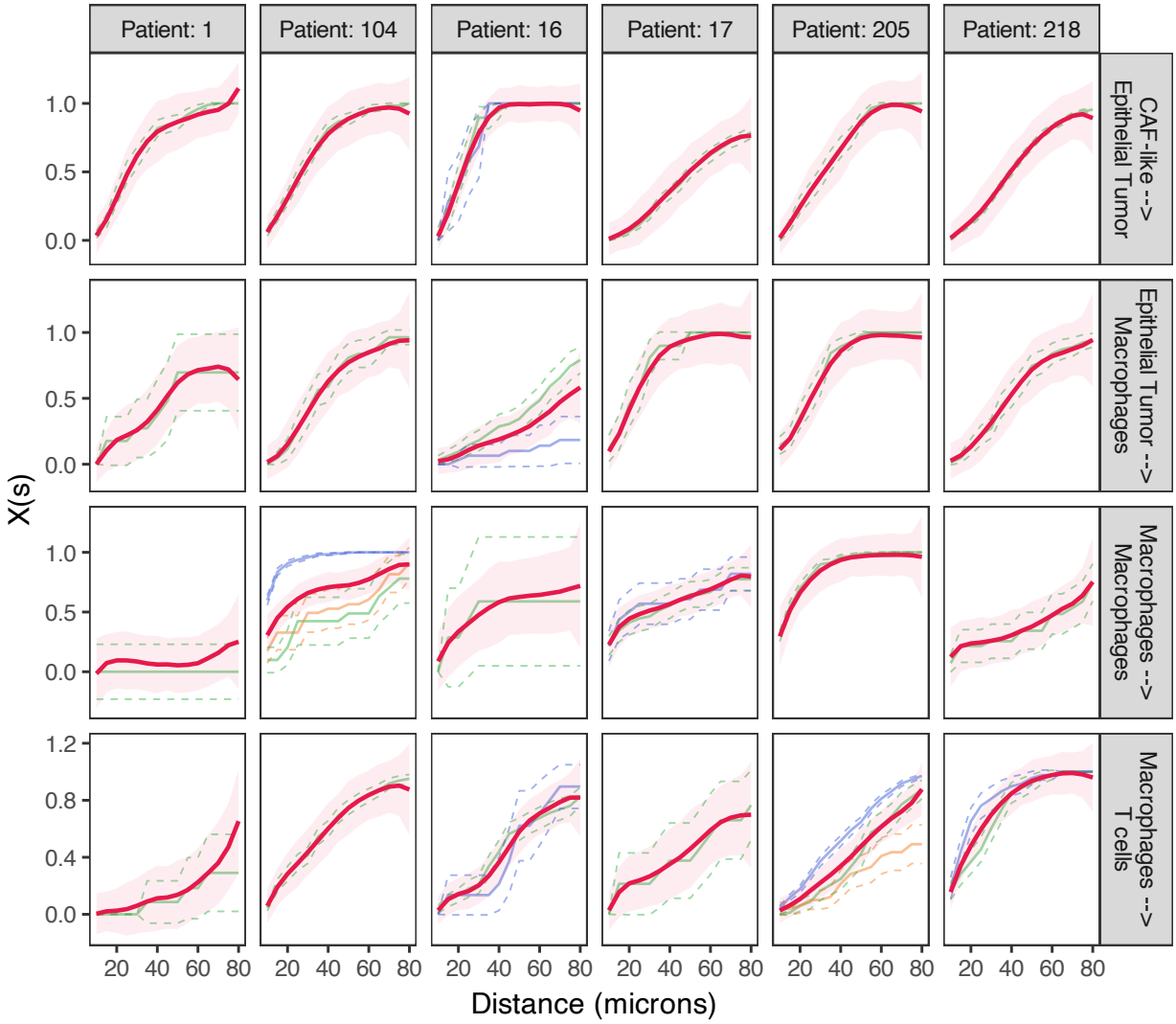


Figure 6.3: For a subset of patients, the posterior mean and 95% simultaneous credible interval for $X_i(s)$ were plotted alongside the observed noisy measurements. Observed curves are shown in light color, with dashed lines indicating ± 2 standard deviations based on bootstrap-estimated sampling error. The estimated posterior function for each patient is overlaid as a solid line with a light ribbon indicating the credible interval. Facets are arranged by interaction (e.g., Macrophages $\rightarrow$ T cells) and patient ID.

CHAPTER 7

Case Studies in Simulation-Based Inference for Mechanistic Tumor–Immune Modeling

As a complement to the spatial statistical methods developed in earlier chapters, this final chapter explores the use of simulation-based inference (SBI) for mechanistic modeling of tumor–immune dynamics. Whereas previous chapters focus on probabilistic modeling of spatial tissue architecture from multiplexed imaging data, the work presented here addresses a different but related challenge: inferring parameters in mechanistic models of tumor-immune interactions where the likelihood is intractable or difficult to compute. This is a common scenario in quantitative systems pharmacology (QSP), where models are typically specified through systems of differential equations that encode biological processes at the cellular or molecular scale.

SBI methods provide a flexible, likelihood-free framework for estimating parameters in such models by learning from forward simulations rather than requiring closed-form likelihoods. This chapter presents two case studies. The first is a proof-of-concept application of SBI to a simplified tumor–immune ordinary differential equation (ODE) model, where posterior distributions over parameters are estimated using neural simulation-based methods and benchmarked against traditional Hamiltonian Monte Carlo. The second case study applies similar techniques to a more complex QSP model of PD-1/PD-L1 checkpoint inhibition, highlighting both the strengths and limitations of SBI in high-dimensional biological systems. Together, these case studies demonstrate the potential of SBI to enable scalable, flexible inference in mechanistic pharmacology models while revealing challenges in identifiability and high-dimensional simulation.

7.1 Simulation-Based Inference for Simple Models of Tumor-Immune Dynamics: A Proof of Concept

To evaluate the applicability of SBI in quantitative systems pharmacology, I implemented a proof-of-concept study using a simplified mechanistic model of tumor–immune interactions. SBI refers to a family of likelihood-free Bayesian methods that rely on forward simulations from complex models rather than requiring closed-form likelihood functions (Cranmer, Brehmer, and Louppe, 2020). This approach is particularly well suited to biological systems where the models are mechanistic and stochastic, but the likelihood is either intractable or computationally expensive to evaluate. In this study, I used neural posterior estimation via BayesFlow—an amortized inference framework that learns a mapping from observed data to posterior distributions over parameters (Radev, Mertens, et al., 2022; Radev, Schmitt, et al., 2023). The purpose of this exercise was to assess how well SBI techniques can recover biophysically meaningful parameters from time-series data generated by ODE models. In parallel, I benchmarked the results against traditional Hamiltonian Monte Carlo (HMC) inference as implemented in Stan (Carpenter et al., 2017), allowing for a direct comparison between amortized and non-amortized Bayesian approaches.

The model under consideration describes the dynamic interaction between tumor cells and immune effector cells, capturing essential features of tumor growth and immune response. Letting $T(t)$ and $I(t)$ represent the tumor and immune cell populations at time t, respectively, the system is governed by a pair of nonlinear differential equations:

$$\frac{dT}{dt} = aT\left(1 - \frac{T}{K}\right) - bTI,$$
$$\frac{dI}{dt} = cTI - dI.$$

This formulation assumes that tumor cells grow logistically with intrinsic growth rate a and carrying capacity K, while immune cells act to reduce tumor burden through cytotoxic interactions at rate b. Conversely, tumor cells stimulate immune activation, modeled via a proliferation rate c, and immune cells decay naturally over time at rate d. These assumptions encode a bidirectional interaction where the tumor promotes immune engagement, and the immune response in turn suppresses tumor growth. The inverse problem of interest is to estimate the parameters $\theta = (a, K, b, c, d)$ based on noisy observations of both cell populations over time.

To simulate realistic experimental conditions, I constructed a forward model that included both deterministic dynamics and stochastic observation noise. Parameters were drawn from biologically plausible prior distributions and used to numerically integrate the ODE system

over a fixed time interval using the `odeint` routine from the SciPy package (Virtanen et al., 2020). The resulting trajectories of tumor and immune populations were then perturbed with additive Gaussian noise, representing the effects of measurement error. This defined a simulator that could generate synthetic training data for SBI. Examples of trajectories generated from various parameter settings are shown in Figure 7.1.

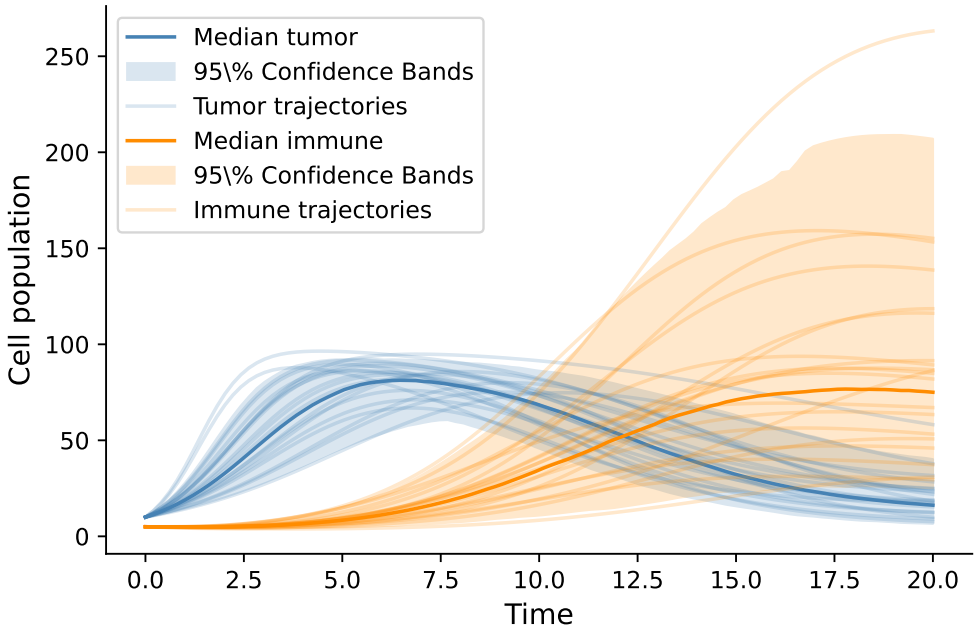


Figure 7.1: Simulated trajectories of tumor and immune populations across multiple parameter settings.

For the inference stage, I employed the BayesFlow framework to train a neural network capable of learning the mapping from observed time-series data to posterior distributions over model parameters. The architecture consisted of two components: a summary network and an inference network. The summary network used a time series transformer to automatically extract informative features from simulation trajectories, eliminating the need for hand-crafted summary statistics (Wen et al., 2023). This transformer architecture processes the entire temporal sequence with attention mechanisms, allowing the model to capture long-range dependencies and dynamic patterns across time points. The output of the summary network—a fixed-length embedding—was then passed to the inference network. The inference network used a flow-matching approach, in which a normalizing flow was trained to transform samples from a base distribution (e.g., a standard Gaussian) into samples from the posterior distribution over parameters conditioned on the learned embedding (Lipman et al., 2022). Together, these two networks formed an amortized inference engine capable of recovering parameter posteriors directly from raw simulation outputs. The training data

consisted of many independent simulations, each pairing a sampled parameter vector with its corresponding noisy trajectory. Once trained, the BayesFlow model was able to infer posterior distributions for new datasets without retraining, offering the practical advantage of amortized inference—that is, inference that is near-instantaneous at test time. Importantly, as mentioned previously, this approach did not require an explicit likelihood function and was able to learn directly from data generated by the forward simulator.

The performance of BayesFlow was evaluated through several complementary diagnostics to assess both accuracy and calibration. First, I visualized the resulting posterior distributions and compared the posterior means to ground-truth parameters using held-out validation simulations. For each parameter, I plotted the true value against samples drawn from the posterior distribution inferred by the model. These scatter plots revealed tight alignment between true and inferred values across all parameters, demonstrating strong recovery performance under amortized inference (Figure 7.2). To assess the calibration of the inferred posteriors, I applied simulation-based calibration (SBC) using the empirical cumulative distribution function (ECDF) of rank statistics (Talts et al., 2020). For each parameter, the posterior rank of the true value was computed across many simulated datasets, and the uniformity of these ranks was visualized using ECDF plots (Figure 7.3). In a well-calibrated model, the ECDF of ranks should lie within the 95% confidence bands under a uniform distribution. All parameters except for the noise variance on the tumor population σ_T exhibited approximately uniform rank distributions, indicating well-calibrated posteriors. The posterior of σ_T, however, denotes that the model is overconfident in this parameter. Finally, I compared the posterior predictive distribution to the prior predictive distribution and found that the former exhibited substantially reduced variance and improved alignment with the observed data (Figure 7.4). This tightening of the predictive distribution reflects the model's ability to constrain uncertainty in biologically meaningful ways through amortized inference.

To benchmark these results, I implemented the same mechanistic model in Stan using the CmdStanPy interface. Although Stan produced highly accurate posterior estimates (Figure 7.5), the computational cost was substantially higher for each individual dataset, since inference was performed from scratch for each new observation. This renders validation techniques like simulation-based calibration much less viable.

Overall, the results demonstrated that BayesFlow was able to match or exceed the accuracy of Stan in estimating parameters from synthetic data, while offering superior scalability and flexibility. The amortized nature of the BayesFlow approach enabled rapid inference across multiple datasets without re-running training, making it particularly well suited for applications involving large cohorts and high-throughput screening, as well as validation techniques such as simulation-based calibration. Moreover, SBI frameworks like BayesFlow

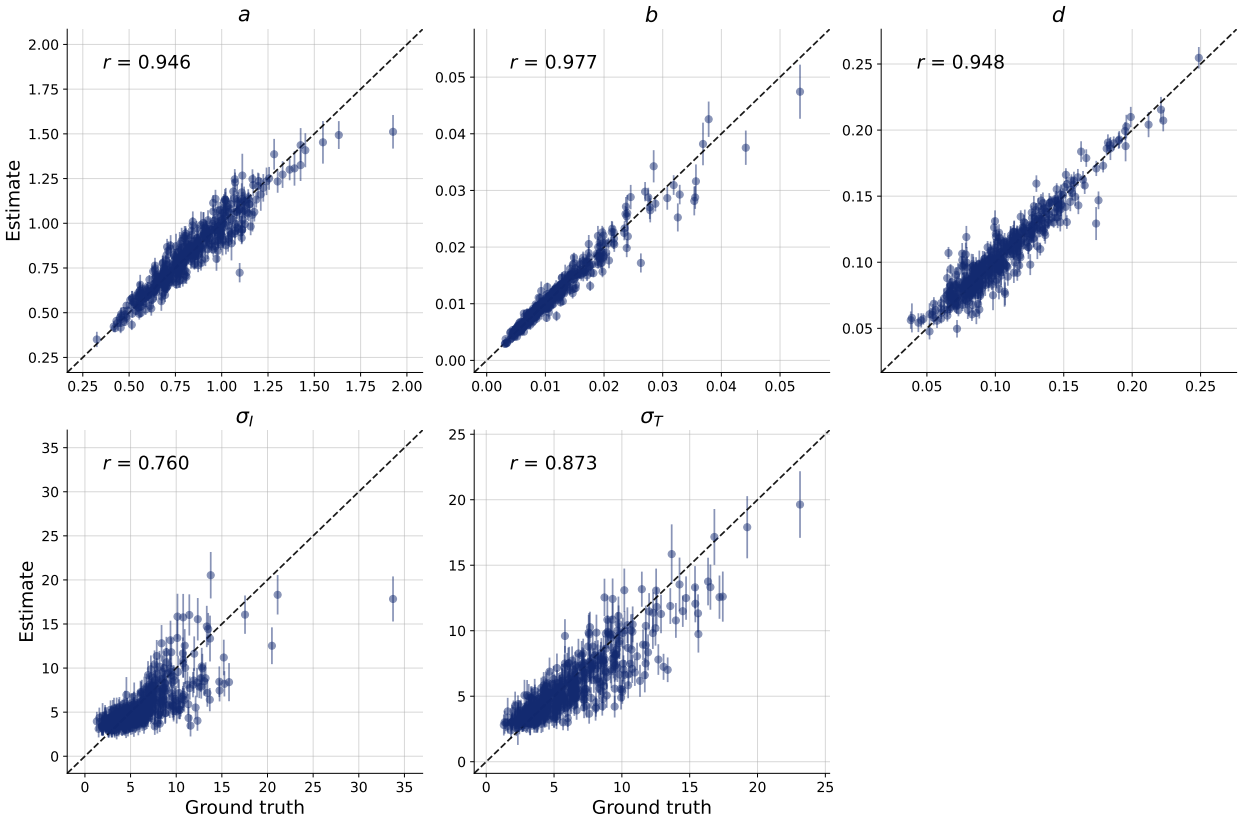


Figure 7.2: Posterior means versus ground-truth parameter values across held-out validation simulations. Each point represents the average posterior estimate from amortized inference using BayesFlow.

can accommodate observation processes that may be complex or non-differentiable, which can pose challenges for gradient-based inference methods such as HMC.

In conclusion, this case study highlights the potential of SBI as a powerful tool for parameter inference in mechanistic systems biology models. By directly leveraging forward simulations and bypassing the need for tractable likelihoods, SBI enables flexible and scalable inference, and stands as a valuable complement to traditional Bayesian methods in systems biology.

7.2 Simulation-Based Inference for a Mechanistic QSP Model of PD-1/PD-L1 Inhibition

To extend simulation-based inference techniques to more complex biological models, I then applied SBI for parameter recovery in a mechanistic QSP model of cancer immunotherapy. This work focuses on a reduced version of a spatial QSP model that describes immune

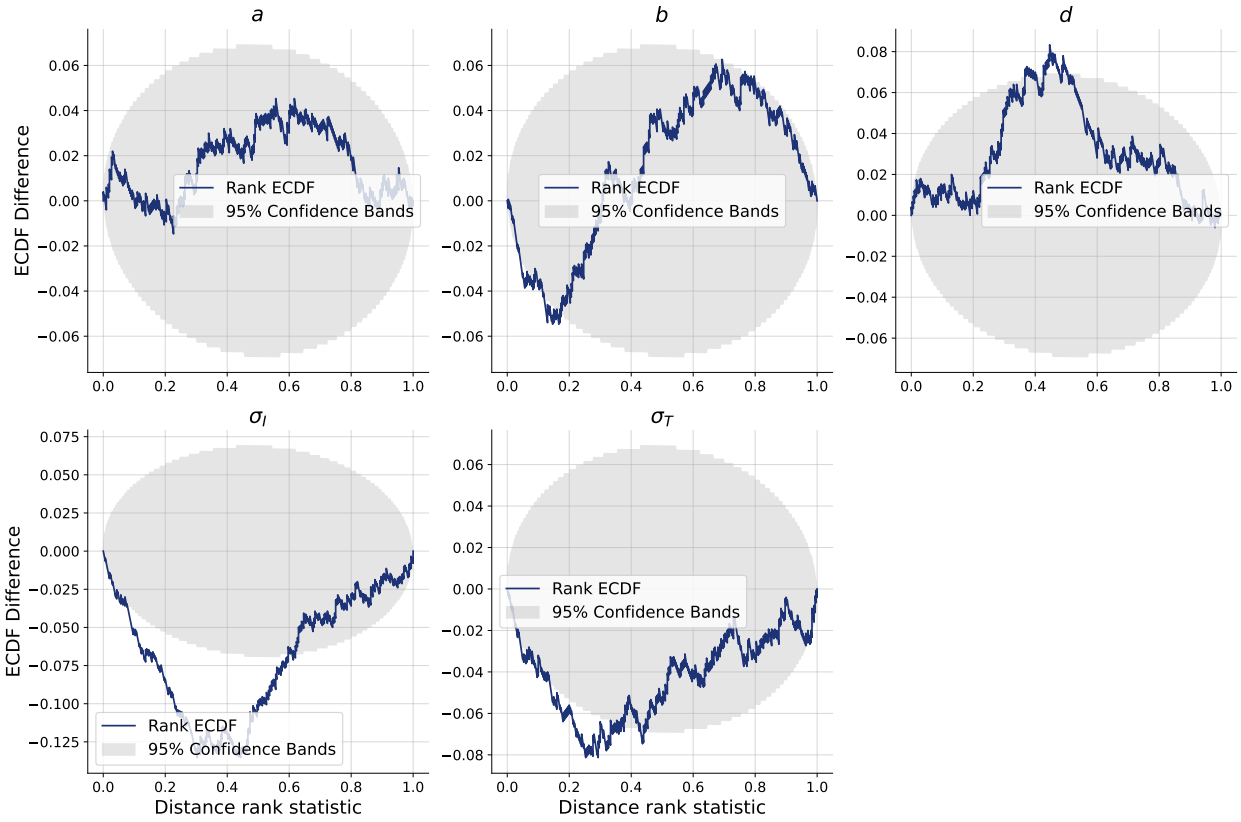


Figure 7.3: Simulation-based calibration using rank ECDFs for each inferred parameter.

checkpoint inhibition via the PD-1/PD-L1 pathway (Gong et al., 2021). To isolate the parameter inference task, I explicitly decoupled the spatial, agent-based component of the full model and instead focused on the core ODE system that governs dynamics of immune and tumor cell populations, cytokines, and immunotherapeutic agents.

The forward model maps biological parameters - such as tumor proliferation rates, immune infiltration dynamics, PD-1/PD-L1 binding kinetics, and T-cell exhaustion thresholds - to observable quantities over time. These outputs may include tumor burden, T-cell counts, and concentrations of signaling molecules such as IFN-γ or IL-2. Because the model encodes key mechanisms of immune checkpoint therapy, it provides a useful testbed for assessing how well inverse modeling techniques can recover parameters from complex biological dynamics.

In this project, building on the work in Section 7.1, I trained a transformer-based neural network to infer model parameters from simulated QSP trajectories. Each training example paired a multivariate time series of model outputs with the corresponding set of ground-truth parameters. These data were generated in advance by running large batches of simulations across a range of biologically plausible parameter values and initial conditions. The model was trained to learn a mapping from observed outputs to parameter distributions by trans-

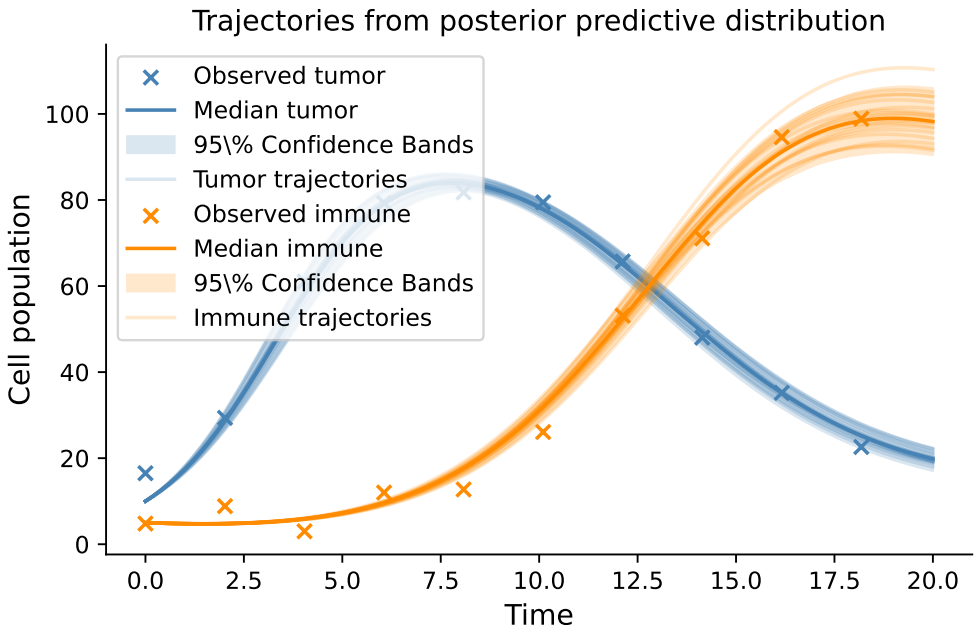


Figure 7.4: Posterior predictive distribution over time.

forming the parameters into a simpler latent space, where they follow a standard Gaussian distribution. The training loss encouraged the model to make this transformation as smooth and consistent as possible while still preserving the structure of the original parameter distributions. This approach allowed the network to learn a flexible approximation to the true posterior over parameters. Training and validation loss over epochs are shown in Figure 7.6.

Figure 7.7 shows simulation-based calibration results using rank ECDFs for each inferred parameter (Säilynoja, Bürkner, and Vehtari, 2022). Most parameters show good calibration, but some exhibit systematic deviations. For example, `k_Treg_transmig` and `k_Teff_transmig` rise above the confidence bands, indicating that the model is underconfident—its posteriors are too wide, with true parameter values falling near the center more often than expected. In contrast, parameters such as `Kd_p1_0_M1` and `K_C1_PDLX_Teff_PD1` dip below the bands, suggesting overconfident posteriors that are too narrow and place too little mass in the tails. Parameters like `Teff_PD1_tot` show minimal deviation and appear well-calibrated. These calibration patterns help diagnose which parameters are most prone to inference error and may guide future improvements to the model or training procedure.

Figure 7.8 shows posterior mean estimates versus ground-truth parameter values across held-out validation simulations, along with 80% credible intervals. Ideal recovery would align points along the diagonal. Several parameters were recovered with high accuracy, including `k_C_growth` and `Kd_p1_0_M1`, both showing strong correlation between true and

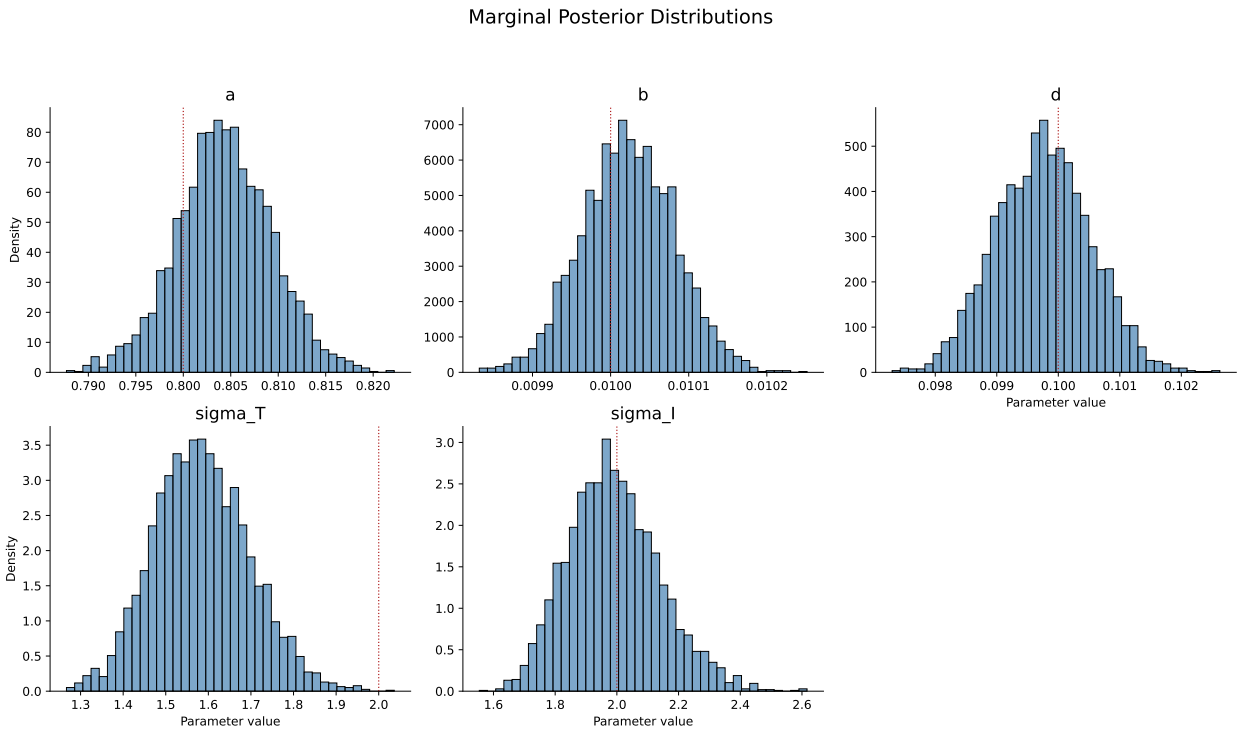


Figure 7.5: Posterior distributions as estimated via Stan.

estimated values ($r = 0.750$ and $r = 0.748$, respectively). Parameters such as `n_clone_p1_0`, `k_Treg_transmig`, and `k_Teff_transmig` also displayed moderate recoverability. In contrast, parameters like `cl_Nivo`, `Teff_PD1_tot`, and `Kd_PD1_PDL2` showed weak or no correlation with their true values, indicating that these quantities may be poorly identified from the available model outputs. This variation in recovery performance highlights which biological mechanisms are more easily inferred from data and which may require additional observations, model structure, or priors to improve identifiability.

Figure 7.9 shows the relationship between posterior contraction and posterior z-scores across inferred parameters (Schad, Betancourt, and Vasishth, 2020). Posterior z-scores are defined as the difference between the posterior mean and the true parameter value, scaled by the posterior standard deviation. Ideally, these values should be centered around zero and fall mostly within the range of [-3, 3], indicating well-calibrated uncertainty. Posterior contraction measures the reduction in uncertainty compared to the prior; values near 1 reflect strong information gain from the data, while values near zero suggest little or no learning. In this model, most parameters exhibit moderate contraction (e.g., around 0.4–0.6), with z-scores broadly centered near zero, indicating generally informative and stable inference. However, some parameters such as `k_Teff_transmig` and `k_Treg_transmig` show quite high contraction for some sets of posterior samples but also extreme z-scores, suggesting overcon-



Figure 7.6: Training and validation loss over epochs during model training.

fident estimates in those dimensions. Others, like `Kd_PD1_PDL2` and `cl_Nivo`, show near-zero contraction, implying that the data provided little information to constrain those parameters.

Overall, this study demonstrates that simulation-based inference can be effectively applied to a complex QSP model of immune checkpoint inhibition. The transformer-based summary and inference networks were able to recover some key parameters with reasonable accuracy and calibration, even in the presence of model noise and limited observables. At the same time, variability in calibration and contraction across parameters revealed important limitations—particularly for parameters with weak influence on the model outputs or strong prior uncertainty. These results point to the value of using posterior diagnostics such as rank ECDFs and z-score contraction to better understand the inferential landscape of mechanistic models. They also suggest that careful model design, targeted output selection, and the potential inclusion of additional priors or data sources will be important for improving parameter identifiability in future applications.

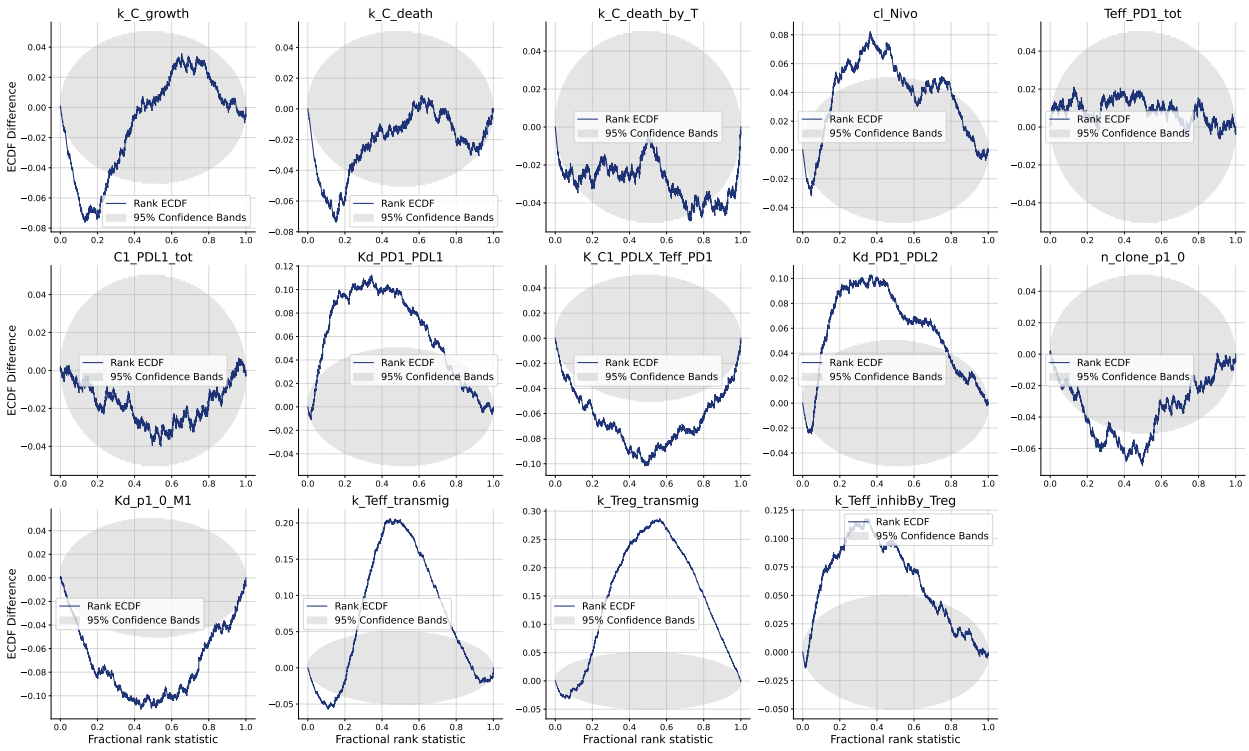


Figure 7.7: Simulation-based calibration using ECDF rank plots for each inferred parameter from the QSP model.

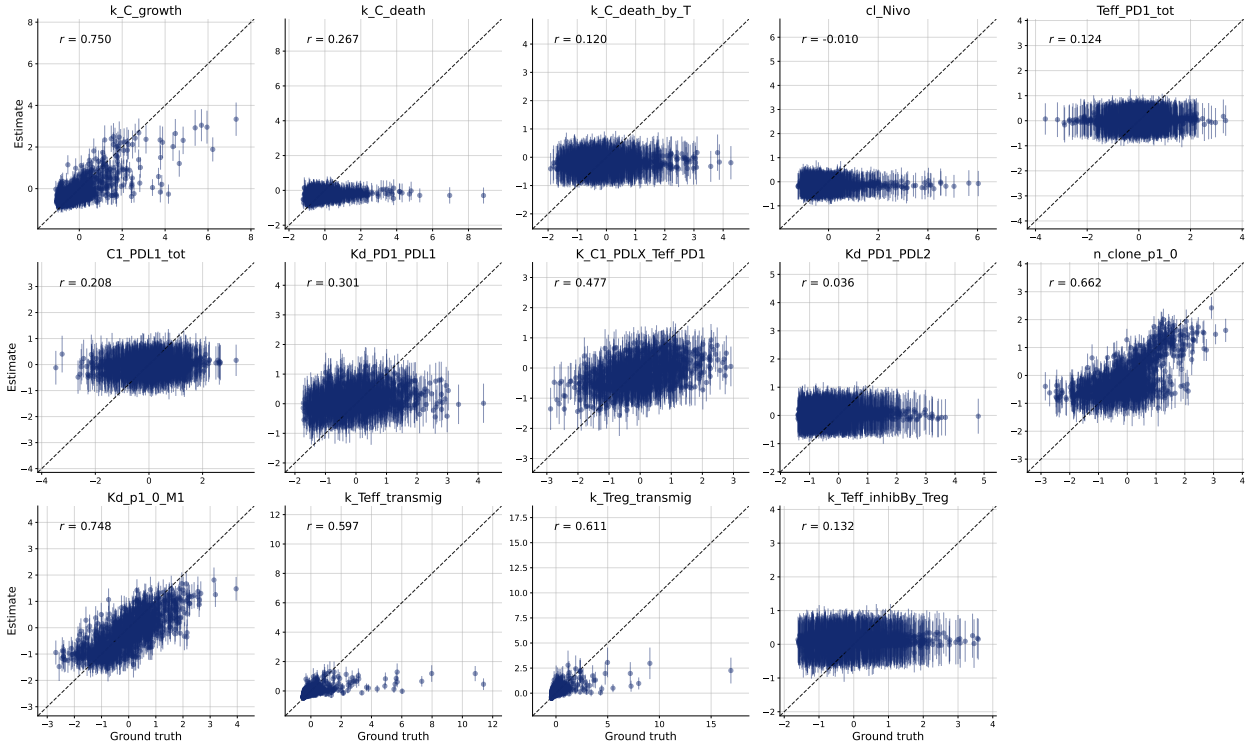


Figure 7.8: Posterior means versus true parameter values on held-out validation data in the QSP model.

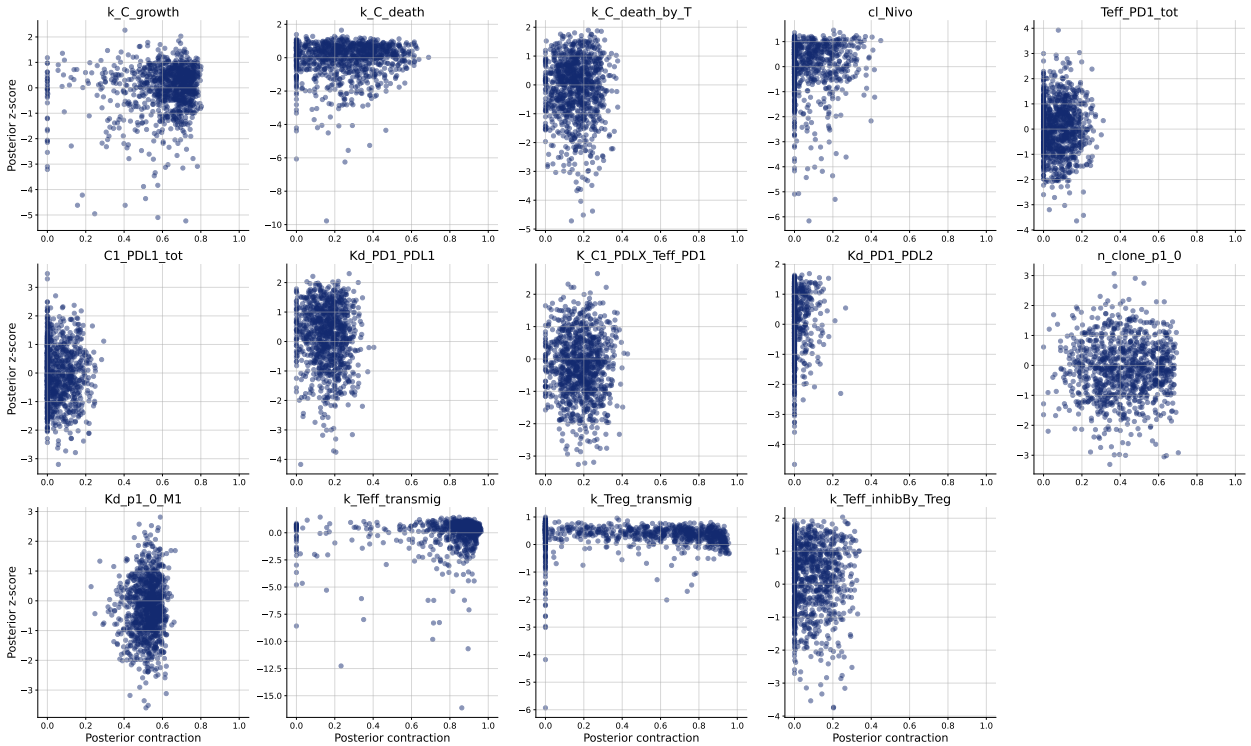


Figure 7.9: z-score contraction plot comparing prior and posterior distributions.

CHAPTER 8

Discussion

8.1 Summary of Contributions

This dissertation brought together a range of statistical approaches for modeling the spatial organization of the TME using multiplexed imaging data. The underlying goal throughout was to build models that were both rigorous and practical—models that reflected the structure of the data, respected biological complexity, and offered clearer insight into how cells interact within tissue. Although the methods varied across chapters, they were all grounded in the same motivation: to better understand the spatial biology of cancer through principled modeling.

The first part of the dissertation focused on applied spatial analysis. Chapter 2 examined immune architecture in metastatic melanoma, identifying spatial features linked to disease stage and survival. Chapter 3 applied a multitype Gibbs point process model to colorectal cancer data, demonstrating how probabilistic models could be used to characterize local interactions between cell types and relate them to clinical variables.

The second part moved into methodological development. Chapter 4 introduced SHADE, a multilevel spatial interaction model that captured directional relationships between cell types across images, patients, and cohorts. Chapter 5 presented BIPPS, a latent spatial factor model that estimated shared spatial patterns across patients by treating point patterns as multivariate spatial outcomes. Chapter 6 shifted the focus to clinical outcomes, developing a Bayesian survival model that included spatial summary functions as functional covariates, while explicitly accounting for measurement error.

Finally, Chapter 7 stepped outside the spatial modeling framework to explore simulation-based inference for mechanistic models of tumor–immune dynamics. While more exploratory in nature, this work shared the same interest in building statistical tools that could handle biological complexity in a transparent and flexible way.

Altogether, this dissertation offered a set of ideas and tools for making sense of high-

dimensional spatial data in cancer, with an emphasis on models that are interpretable, structured, and responsive to the data they are applied to.

8.2 Methodological Themes and Lessons Learned

Several methodological themes emerged across the work presented in this dissertation. While each chapter addressed a different modeling task, many of the same challenges recurred in various forms: balancing complexity with interpretability, handling multilevel structure in real-world datasets, accounting for uncertainty, and working within the constraints of both data and computation.

A persistent tension was the tradeoff between expressive models and statistical tractability. For instance, a fully joint asymmetric spatial interaction model, where all pairwise cell-type interactions are modeled simultaneously, could in principle capture a richer representation of the tumor microenvironment. However, such a model is difficult to specify cleanly and is ill-posed without strong constraints or prior information. Similarly, incorporating high-dimensional functional marker data would likely yield more biologically detailed interaction models, but the number of possible marker combinations and interaction types grows quickly. Modeling these relationships in a coherent, computationally feasible way remains a significant open challenge. In this context, simpler models that are well-structured and interpretable often offer a better balance between biological plausibility and statistical rigor.

The importance of multilevel modeling became apparent early in the process. Tissue imaging datasets naturally have hierarchical structure, with images nested within patients and patients grouped into clinically distinct cohorts. Ignoring this structure risks misrepresenting uncertainty and overstating the generalizability of image-level findings. At the same time, hierarchical models can be difficult to fit well in practice. Issues such as weak identifiability and pathologies like Neal's funnel can arise, particularly when data at higher levels are sparse (Neal, 2003; Stan Development Team, 2025). Careful prior specification and thoughtful model design can be very useful for producing stable and calibrated estimates in these settings.

Another methodological consideration involved measurement error in spatial covariates. Many spatial analyses treat estimated functions—such as nearest-neighbor distributions, like the G-cross function—as fixed inputs to downstream models, even though these estimates are subject to noise and sampling variability. Ignoring this uncertainty can lead to overconfident predictions and misleading inferences. However, quantifying measurement error in spatial statistics, particularly when they take functional form, is not straightforward. The bootstrap-

based approach used in Chapter 6 provided a practical solution, but more formal treatments of this uncertainty would be a useful direction for future work.

In Chapter 7, the focus shifted to simulation-based inference for mechanistic models of tumor–immune dynamics. This work highlighted the advantages of amortized inference methods, which can learn a mapping from simulated outputs to parameter posteriors and then be applied rapidly to new data. Compared to traditional Bayesian techniques like Hamiltonian Monte Carlo, these methods scale more easily and allow for richer validation strategies, such as simulation-based calibration. More importantly, this exploration was motivated by a broader realization: spatial imaging data alone is often insufficient to uncover causal relationships or to simulate the effects of therapeutic interventions. While spatial interactions can suggest biological associations, they rarely provide enough temporal or mechanistic information to understand how changes in one part of the system might affect another. Mechanistic models of dynamics, even simplified ones, offer a complementary perspective and open the door to more dynamic forms of hypothesis testing and prediction.

Finally, the issue of generalizability remained a central concern throughout the dissertation. Biological data, especially from tumor tissues, are highly heterogeneous across patients, disease types, and imaging platforms. The multilevel models developed here helped address this by structuring variation across individuals and groups. Still, many simplifying assumptions were necessary to make the models identifiable and computationally feasible. Capturing patient heterogeneity without sacrificing interpretability remains a key challenge, especially in applications where individualized inference is the goal.

These challenges highlight the gap between the questions we'd like to ask about the tumor microenvironment and what the data and models can realistically support. The methods in this dissertation were shaped by that gap. In many cases, I had to simplify the models or make compromises to get something that could be fit and interpreted. In the end, the focus was on making the models useful and interpretable, even if they couldn't capture every detail.

8.3 Biological and Clinical Insights

While the primary focus of this dissertation was methodological, many of the models developed here helped reveal biologically meaningful patterns with potential clinical relevance. The integration of spatial structure, hierarchical modeling, and uncertainty quantification provided a framework for identifying cellular organization patterns that may reflect underlying disease mechanisms or therapeutic vulnerabilities.

In melanoma, spatial interaction patterns among immune cells showed consistent associa-

tions with patient outcomes. Increased interactions between NK cells and B cells, as well as organized clustering between CTLs and helper T cells, were associated with improved survival. These patterns may indicate more coordinated immune responses in certain patients and could be predictive of response to immunotherapies. In contrast, excessive macrophage clustering and high B cell accumulation in inflammatory regions were linked to poorer outcomes, suggesting the presence of immunosuppressive niches that may require combination treatments to overcome.

The MGPP model applied to colorectal cancer data demonstrated how spatial modeling can move beyond pairwise statistics and account for a broader set of interactions within the tumor microenvironment. Unlike more ad hoc summary-based approaches, the MGPP framework allowed for joint modeling of multiple cell types and spatial covariates, along with associated uncertainty. This enabled a more integrated view of the TME as a structured, interacting system. The model offers a promising foundation for future work aiming to capture fine-grained spatial dependencies in a probabilistic and scalable way, which we feel we partially accomplished in developing the SHADE model.

In the SHADE analysis, cell-type-specific interaction patterns with stromal components highlighted potential mechanisms of immune exclusion and immune modulation. CTLs in DII patients showed increased clustering near CAFs across spatial ranges, consistent with CAF-mediated immune suppression. In contrast, CTL–CAF interactions in CLR patients peaked at intermediate distances, suggesting reduced stromal constraint. Similarly, memory CD4+ T cells in CLR tumors clustered near CAFs and epithelial–mesenchymal hybrid cells, raising the possibility of helper T cell involvement in modulating EMT plasticity or maintaining immune-permissive niches.

The BIPPS model applied to pancreatic cancer data revealed a striking contrast between IPMN and PDAC cases. Immune cells in IPMN tumors were more diffusely and evenly distributed, suggesting relatively intact immune surveillance. In PDAC, immune cells—especially T cells—were spatially segregated and excluded from tumor regions, reinforcing the idea that PDAC promotes an immunosuppressive architecture that facilitates immune evasion (Jamouss et al., 2025; Hernandez et al., 2022).

The functional survival model applied to breast cancer data illustrated how incorporating spatial interaction curves as functional covariates can yield more nuanced insights than traditional fixed-radius summaries. The estimated effect surfaces $\beta(s)$ captured scale-dependent associations between spatial proximity and outcomes, and the model flexibly combined multiple biopsies per patient to yield stable survival predictions. This is especially valuable in clinical studies, where tissue heterogeneity and sampling variability are common.

Finally, the SBI case studies demonstrated the feasibility of applying amortized Bayesian

inference to mechanistic QSP models. Despite limited observables and model noise, neural inference methods recovered several key parameters with good calibration. These results suggest that SBI can support parameter recovery and validation in mechanistic models of immunotherapy, particularly when paired with simulation-based diagnostics and uncertainty assessments.

As stated previously, even though the primary focus of this dissertation was methodological, the models were built to surface meaningful biological patterns—and some of the results point to associations that could be explored further in clinical or experimental studies.

8.4 Limitations and Open Questions

While the models developed in this dissertation address several key challenges in spatial and hierarchical modeling, there are important limitations to acknowledge, as well as open questions that remain unresolved.

First, many of the models rely on simplifying assumptions that made the inference problem more tractable, but also limited the biological complexity that could be captured. For instance, the spatial interaction models treated cell types as the only labels of importance, with very limited support for incorporating functional marker information. In reality, functional state often matters as much as, or more than, cell identity. Extending models like the MGPP model or SHADE to include functional markers is a natural next step, but presents challenges both statistically and biologically. When there are many functional markers, it's not clear how they should enter the model: as modifiers of spatial interaction, as outcomes influenced by spatial context, or perhaps both. Functional markers could also interact with one another and with cell types in complex ways.

Separately, uncertainty in cell type annotations—arising from measurement noise or ambiguous marker expression—can propagate into downstream analyses. One possible approach to address this uncertainty is to treat cell type annotations probabilistically, weighting pairwise interaction terms by the confidence in cell type assignment. This would better reflect the classification uncertainty but may introduce significant computational complexity, especially given the large number of possible cell type pairs. Addressing both of these challenges—integrating functional markers and accounting for uncertain annotations—will require modeling innovation and biological insight.

Moreover, the biological relationships between cell types are not independent. Certain cell types are more closely related biologically, sharing common lineages or functional similarities, which can be thought of as "phylogenetic" relationships. Current models often treat all cell types as equally distinct, ignoring this structure. Future models could incorporate these

relationships through hierarchical random effects that reflect the phylogenetic structure, potentially improving robustness and interpretability. However, this approach would also introduce new modeling and computational challenges, particularly in balancing complexity with tractability.

Second, data limitations placed meaningful constraints on what could be modeled reliably. Some datasets included only a handful of biopsies per patient, or had uneven representation across cohorts, which limited the depth of multilevel modeling. Additionally, the functional survival models were affected by heterogeneity in the spatial covariates across patients—particularly the estimated G-cross functions—which likely reflected both biological variability and unmeasured confounding. At the same time, the survival datasets these models were applied to had relatively few events, reducing the amount of information available to distinguish covariate effects. These issues made it difficult to fully capture the underlying variation in the data and may have contributed to the instability observed in estimating the standard deviation parameters for the spline coefficients. Often these collapsed toward zero, suggesting weak identifiability or a mismatch between the model's assumptions and the noise structure in the data. Improving this may require more informative priors, alternative parameterizations, or refinements in how measurement uncertainty is modeled.

Some of the most interesting open questions are conceptual. Asymmetric spatial interactions are biologically plausible, but modeling them jointly in a way that's statistically coherent is still difficult. From spatial data alone, it's not always clear whether we can meaningfully distinguish directionality—e.g., type A influencing type B differently than B influences A—since the distance structure is symmetric by nature. Still, progress on this front will likely depend more on methodological advances—specifically, developing ways to parameterize joint asymmetry in a statistically identifiable and biologically meaningful way—than on any single dataset or application.

Higher-order interactions beyond pairwise relationships are also biologically plausible and potentially important, such as those capturing spatial configurations involving three or more cells. These extensions could improve biological realism but introduce significant computational and modeling challenges due to the combinatorial explosion of possible configurations. Likewise, the local tissue microarchitecture—including structures such as ridges, gradients, edges, and voids—can strongly influence cellular interactions and neighborhood compositions. Careful analysis and thoughtful inclusion of these structural features as spatial covariates can already enhance the accuracy and biological realism of tumor microenvironment models, without necessarily requiring more sophisticated modeling frameworks.

Finally, the simulation-based inference work in Chapter 7 opens a number of paths for improving mechanistic modeling workflows. While SBI proved useful for recovering parameters

from synthetic data, there's still a lot of room to optimize these methods. Hyperparameter tuning, model reduction, sensitivity analysis, and better experimental design could all improve inference quality, especially for larger QSP models. As these methods move toward more personalized or predictive applications, ensuring good calibration and robust parameter recovery will be increasingly important.

In short, while the models in this dissertation offer useful starting points, they are far from final answers. Many of the most exciting directions—particularly those involving functional state, causal inference, dynamic behavior, and modeling of complex spatial relationships—remain open and will require further methodological development alongside close collaboration with biologists and clinicians.

8.5 Future Directions

The limitations outlined above point to several natural directions for future work—both in extending the models developed in this dissertation and in exploring new ways to combine spatial data, biological mechanisms, and predictive modeling.

One area of growing interest is the integration of spatial statistical modeling with mechanistic models that evolve over time. Spatial imaging provides a snapshot of a highly dynamic system, but without a model of the underlying processes, it is difficult to infer causality or predict the effects of intervention. In future work, I plan to build spatial–temporal models of tumor–immune dynamics, where spatial exclusion and immune suppression are thought to play central roles in disease progression. This will likely involve agent-based or hybrid models that incorporate both spatial structure and biological processes such as immune trafficking, cytokine signaling, and therapeutic response. Developing inference strategies for these complex systems—especially methods that are both scalable and well-calibrated—will be an ongoing challenge.

Another important direction is increasing the biological granularity of spatial statistical models. In particular, better methods are needed for incorporating functional markers into spatial models in a way that reflects their biological role. For example, markers related to immune activation, checkpoint inhibition (e.g., PD-1/PD-L1), or inflammation may modulate how cell types interact, or may themselves be shaped by spatial context. One possibility is to embed biological mechanisms directly into spatial interaction terms—modeling, for instance, how inflammatory signaling might increase the likelihood of certain cell–cell proximities or alter the shape of spatial interaction curves. These kinds of biologically structured models could provide a more mechanistic interpretation of spatial patterns, but they also require careful statistical treatment to remain identifiable and computationally feasible in

the presence of high-dimensional marker data.

There is also strong potential for applying these models in more translational settings, including personalized medicine and clinical trial design. Spatial models that capture variation across patients and relate tissue architecture to outcomes could inform patient stratification, biomarker discovery, or even therapeutic targeting. As these methods continue to improve—especially in terms of uncertainty quantification, model calibration, and scalability—they may become increasingly useful tools for guiding clinical decision-making.

Comparisons with existing methods will also be an important step in validating these models. One strategy I demonstrated in this dissertation is to use simulation studies where data are generated under hierarchical models (such as SHADE) and then fit with both the true model and simpler models that ignore hierarchy. This can demonstrate the value of explicitly accounting for multilevel structure. Further benchmarking could include comparisons with conventional summary statistics or applying the models to publicly available datasets with known biological patterns.

Finally, the work in SBI opens up many possibilities for model validation and exploration in mechanistic systems. Future work in this area will involve tuning neural architectures, improving identifiability through experimental design or model reduction, and developing more rigorous calibration techniques. These directions are especially important for large-scale QSP models, where the number of parameters can be high and model behavior is often sensitive to small changes. Bringing together the flexibility of SBI with the structure and interpretability of mechanistic modeling remains a promising and largely untapped area of research.

Overall, the next phase of this work will involve finding new ways to connect statistical modeling, biological insight, and dynamic simulation, with the broader aim of developing models that are both scientifically grounded and clinically useful.

8.6 Closing Thoughts

This work came from a desire to connect statistical modeling with real biological questions, especially around how the tumor microenvironment shapes disease. Throughout my PhD, I've been drawn to the challenge of trying to model that complexity in a way that's both rigorous and grounded in the data. Building statistical models that could reflect something meaningful about this system was often hard, but also something that brought me a lot of pride and joy. What I enjoyed most was the thinking and problem-solving that came with it—trying to match the structure of the model to the structure of the biology, and seeing where that could take me.

APPENDIX A

Supplementary Material for The SHADE Model for Asymmetric Multilevel Spatial Interactions

A.1 Supplementary Figures

A.1.1 The impact of dummy points on inference

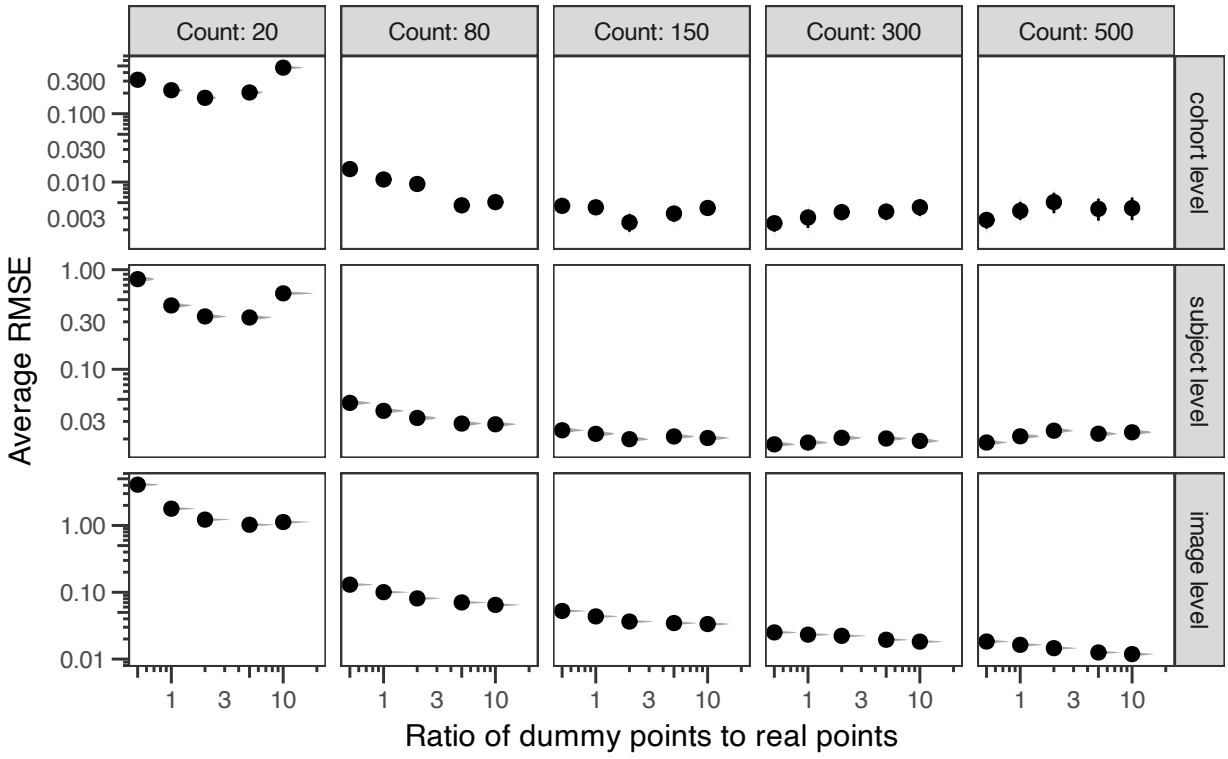


Figure A.1: Average RMSE for spatial interaction coefficients across hierarchical levels: image-level ($\delta^{(m,p)}_{t_1 \to t_2}$), patient-level ($\gamma^{(n,p)}_{t_1 \to t_2}$), and cohort-level ($\psi^{(g,p)}_{t_1 \to t_2}$), under different numbers of cells per type and dummy-to-real point ratios.

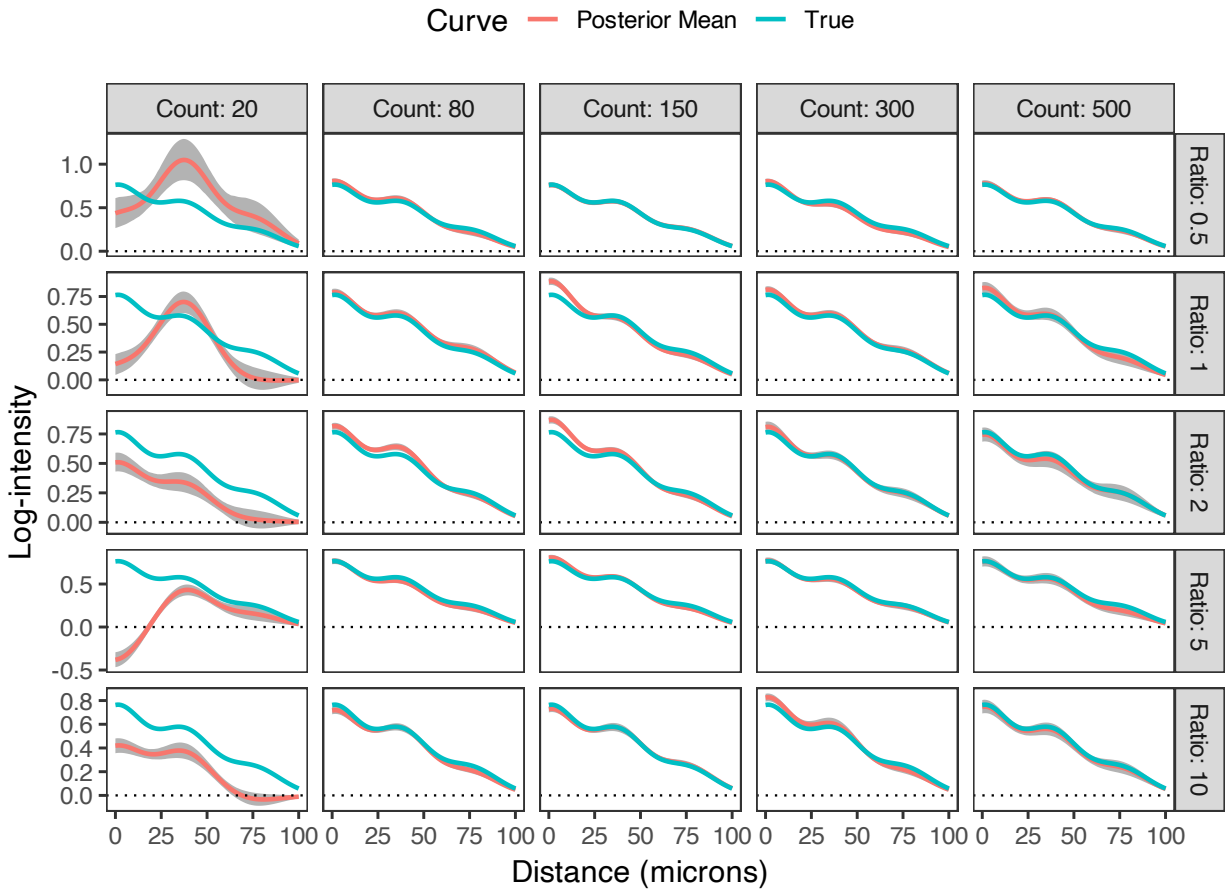


Figure A.2: Example of estimated cohort-level SIC from dummy point simulations.

A.1.2 Number of patients and images per patient

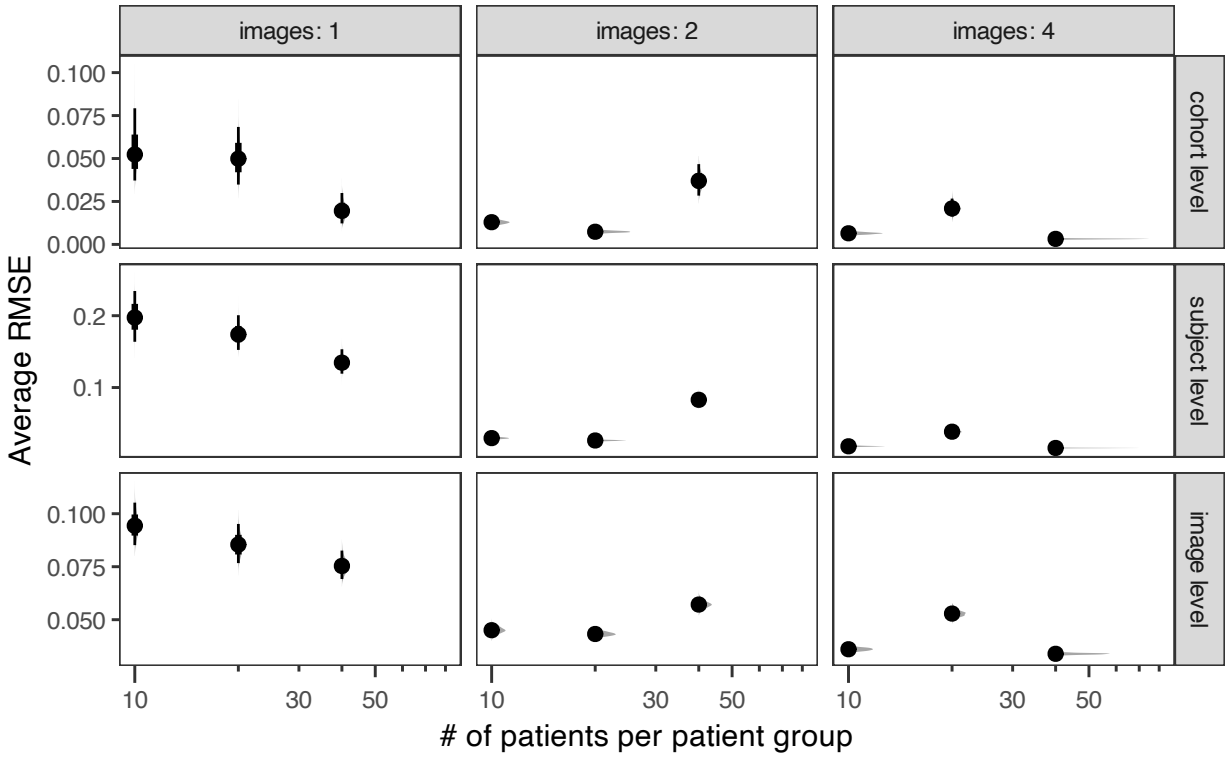


Figure A.3: Average RMSE for spatial interaction coefficients as a function of dataset size, varying the number of patients and number of images per patient. Results are shown for image-level ($\delta^{(m,p)}_{t_1 \to t_2}$), patient-level ($\gamma^{(n,p)}_{t_1 \to t_2}$), and cohort-level ($\psi^{(g,p)}_{t_1 \to t_2}$) parameters.

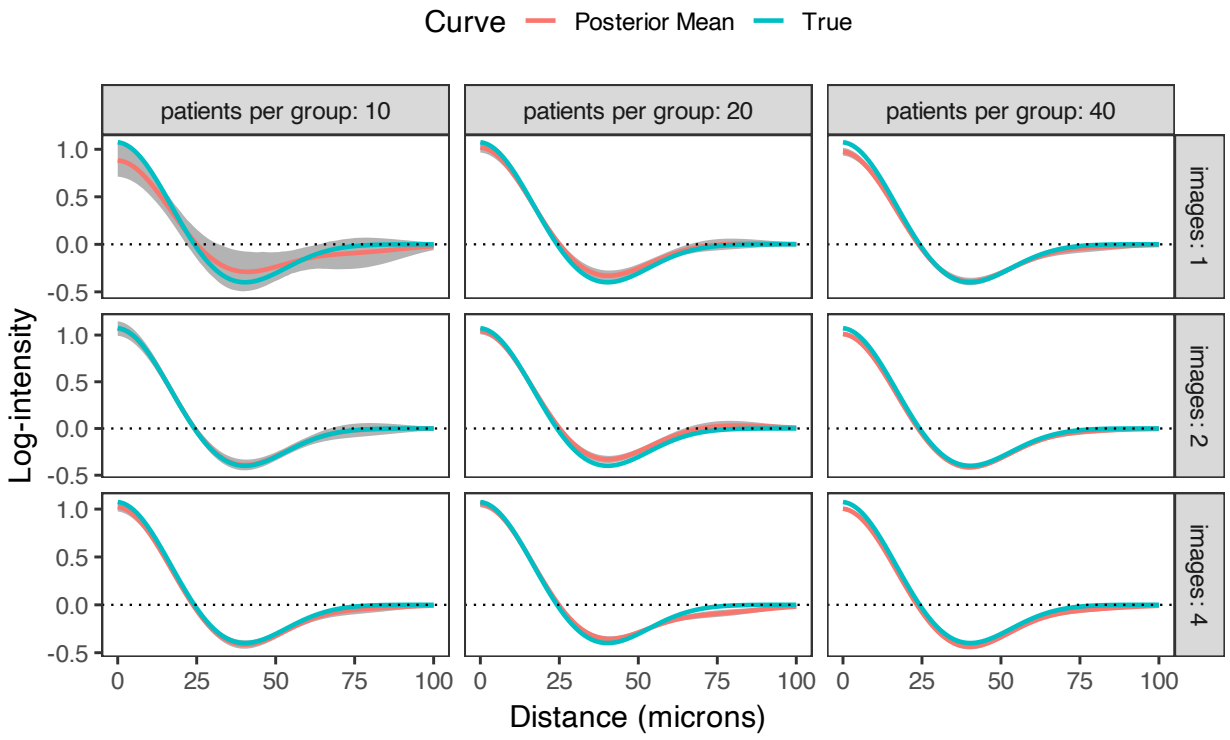


Figure A.4: Example of an estimated cohort-level SIC from the dataset size simulation study, illustrating how increasing the number of patients and images per patient reduces the variance in SIC estimates.

A.1.3 CRC data analysis

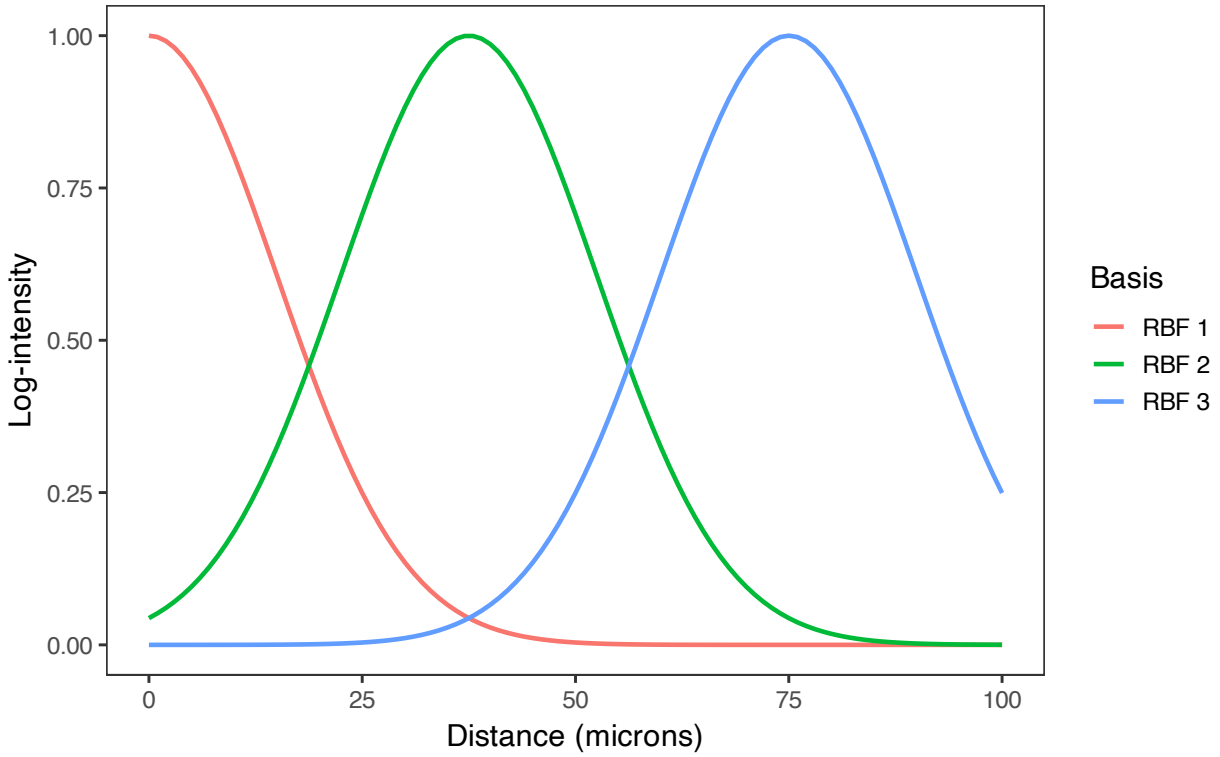


Figure A.5: Radial basis functions $\phi_p(s)$ used to compute distance-based interaction features $\boldsymbol{q}_{A_k}(v)$ in the CRC analysis. These basis functions define the spatial resolution of the SICs.

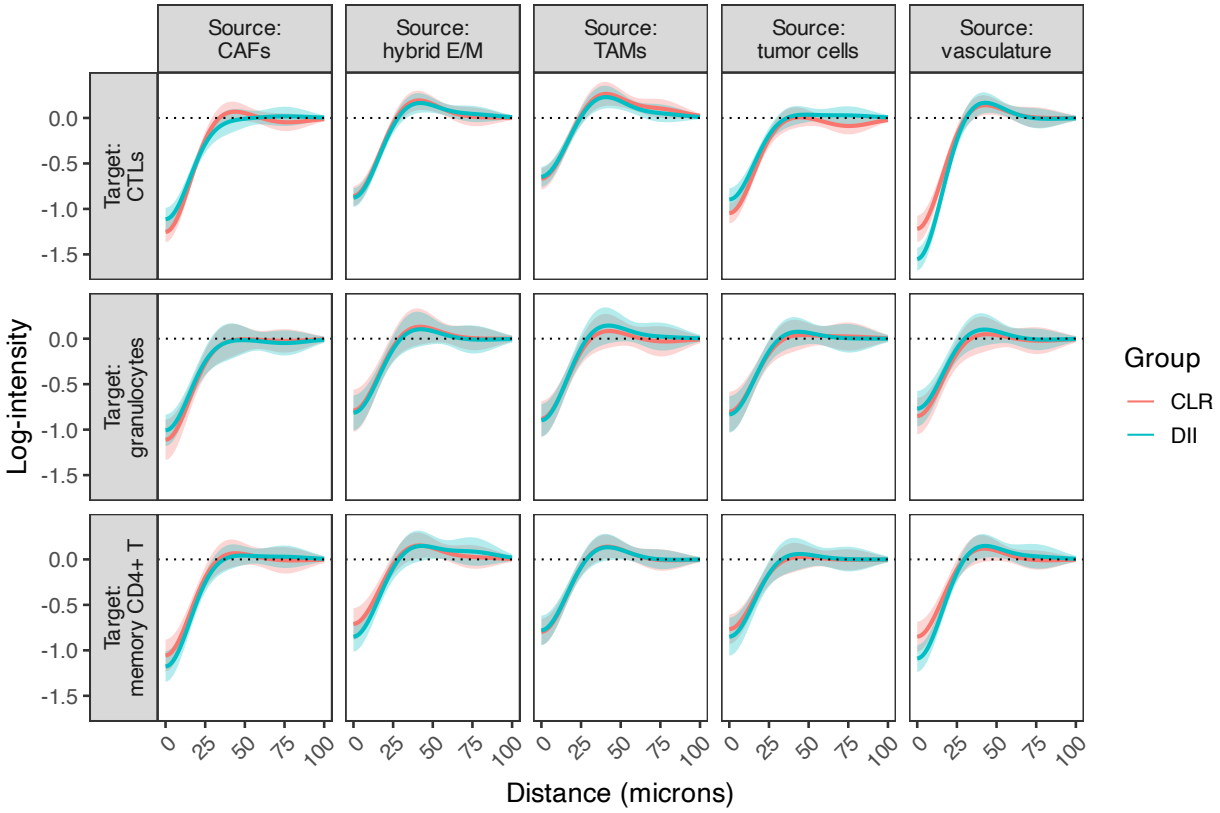


Figure A.6: Cohort-level SICs (parameterized by $\psi^{(g,p)}_{t_1 \to t_2}$) estimated for all source–target cell type pairs in the CRC dataset, stratified by CLR and DII patient groups.

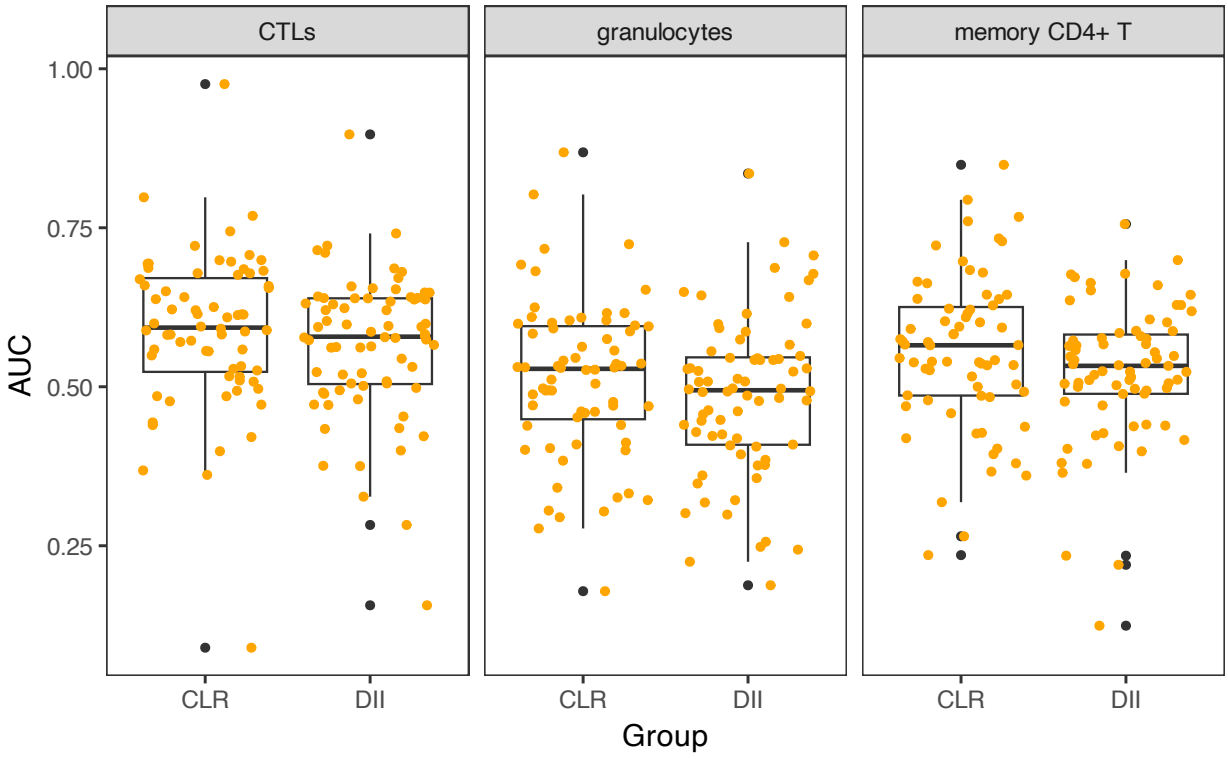


Figure A.7: Distribution of AUCs by cell type and patient group, evaluating prediction performance of SHADE's conditional intensity model when predicting the spatial organization of each target cell type based on the spatial distribution of all other types.

A.2 Priors for Variance Parameters

We place half-normal priors on the standard deviations of the spatial interaction coefficients at each level of the hierarchy:

$$\sigma_{\text{cohort},p} \sim \text{Half-Normal}(\xi_p, \xi_p), \tag{A.1}$$

$$\sigma_{\text{patient},p} \sim \text{Half-Normal}(1.5, 10), \tag{A.2}$$

$$\sigma_{\text{image},p} \sim \text{Half-Normal}(1, 10). \tag{A.3}$$

where is a user-specified location/scale (in our CRC analysis, we used $\xi_p \in \{5, 3, 1\}$).

A.3 Simulation Study Details

A.3.1 Generating Target Points

Target cell locations are drawn from an inhomogeneous Poisson process with intensity:

$$\lambda(v) = \exp\left(\sum_{k=1}^{K}\sum_{p=1}^{P} \boldsymbol{q}_{A_k}^{\top}(v)\boldsymbol{\delta}_{A_k}^{(m)} + \beta_0^{(m)}\right), \tag{A.4}$$

where $\beta_0^{(m)}$ is a normalization offset to ensure that the expected number of target points matches N_{points} in image m. It is computed as:

$$\beta_0^{(m)} = \log\left(\frac{N_{\text{points}}}{\int_W \exp\left(\sum_{k=1}^{K}\sum_{p=1}^{P} \boldsymbol{q}_{A_k}^{\top}(v)\boldsymbol{\delta}_{A_k}^{(m)}\right) dv}\right). \tag{A.5}$$

A.3.2 Default Parameter Settings

Unless otherwise specified, the following defaults were used in simulation experiments:

- **Spatial domain:** $S = 1500$.
- **Number of source cell types:** $K = 2$.
- **Basis functions:** $P = 3$ radial basis functions with bandwidth 15 and support up to 75 microns.
- **Points per cell type per image:** $N_{\text{points}} = 150$.
- **Variance parameters:**

- $\sigma_{\text{cohort},p} = 0.5$
- $\sigma_{\text{patient},p} = 0.1$
- $\sigma_{\text{image},p} = 0.1$

- **Number of patients and images:** variable across experiments.

A.4 Quantifying Variability in Spatial Interactions

We define SICs hierarchically across cohort, patient, and image levels. For a given source–target cell type pair (A, B), to capture hierarchical variation, we define *patient-level* and *image-level* SICs as:

$$\text{SIC}^{(n)}_{A_k \to B}(s) = \sum_{p=1}^{P} \gamma^{(n,p)}_{A_k} \phi_p(s) \tag{A.6}$$

$$\text{SIC}^{(m,n(m))}_{A_k \to B}(s) = \sum_{p=1}^{P} \delta^{(m,p)}_{A_k} \phi_p(s) \tag{A.7}$$

where $\phi_p(s)$ are spline basis functions, and $\gamma^{(n,p)}_{A,B}$ and $\delta^{(m,p)}_{A,B}$ are patient- and image-level coefficients, respectively. To quantify heterogeneity in SICs across patients and images, we compute robust measures of variability at each spatial distance s, and summarize them by taking the median across all distances. Specifically:

Between-patient (within-cohort) variability. Let $\mathcal{C}$ index cohorts, and let $\mathcal{N}_c$ denote the set of patients belonging to cohort $c \in \mathcal{C}$. For each spatial distance s, we compute the cohort-level median SIC as:

$$\overline{\text{SIC}}^{(c)}_{A_k \to B}(s) = \text{median}_{n \in \mathcal{N}_c} \left(\text{SIC}^{(n)}_{A_k \to B}(s) \right)$$

The between-patient (within-cohort) variability for a given cell type pair (A, B) is then defined as:

$$\text{MAD}_{\text{patient}}(A, B) = \text{median}_s \left\{ \text{MAD}_n \left[\text{SIC}^{(n)}_{A_k \to B}(s) - \overline{\text{SIC}}^{(c(n))}_{A_k \to B}(s) \right] \right\} \tag{A.8}$$

Between-image (within-patient) variability. Similarly, for each patient n, we compute the patient-level median SIC:

$$\overline{\text{SIC}}^{(n)}_{A_k \to B}(s) = \text{median}_{m:n(m)=n} \left(\text{SIC}^{(m,n)}_{A_k \to B}(s) \right)$$

The between-image variability is defined as the MAD of image-level SICs from the patient median:

$$\mathrm{MAD}_{\mathrm{image}}(A, B) = \mathrm{median}_s \left\{ \mathrm{MAD}_m \left[\mathrm{SIC}^{(m,n(m))}_{A_k \to B}(s) - \overline{\mathrm{SIC}}^{(n(m))}_{A_k \to B}(s) \right] \right\} \tag{A.9}$$

These robust statistics provide interpretable summaries of spatial heterogeneity at each level of the hierarchy while mitigating sensitivity to outliers and small-sample variability. We report these values for each cell type pair and visualize them using heatmaps (Figure 4.9).

APPENDIX B

Supplementary Material for The BIPPS Model for Shared Spatial Dependencies Across Patients

B.1 Simulation Details

Both sets of simulations used the following sets of equations to simulate from this model:

$$
\begin{aligned}
\varphi_j &\sim \text{Uniform}(\phi_{\min}, \phi_{\max}), \quad j = 1, \ldots, k \\
L_j &= \text{Cholesky}\left(\exp(-\varphi_j D)\right), \quad j = 1, \ldots, k \\
\varepsilon_{i,j} &\sim \mathcal{N}(0, I_n), \quad i = 1, \ldots, N, \quad j = 1, \ldots, k \\
V_i &= \begin{bmatrix} L_1\varepsilon_{i,1} & L_2\varepsilon_{i,2} & \ldots & L_k\varepsilon_{i,k} \end{bmatrix} \\
\beta_0 &= \boldsymbol{\alpha}\mathbf{1}_{1\times q} \\
X_i &= \mathbf{1}_{n\times 1}, \quad i = 1, \ldots, N \\
W_i &= V_i\boldsymbol{A}^T + X_i\beta_0, \quad i = 1, \ldots, N \\
\boldsymbol{y}_i &\sim \text{Poisson}\left(\exp(W_i)\right), \quad i = 1, \ldots, N
\end{aligned}
$$

$$
\boldsymbol{A} = \begin{bmatrix}
\lambda_{11} & 0 & 0 & \ldots & 0 \\
\lambda_{21} & \lambda_{22} & 0 & \ldots & 0 \\
\vdots & \vdots & \ddots & \ldots & 0 \\
\lambda_{k1} & \lambda_{k2} & \ldots & \ldots & \lambda_{kk} \\
\vdots & \vdots & \ddots & \ldots & \vdots \\
\lambda_{q1} & \lambda_{q2} & \ldots & \ldots & \lambda_{qk}
\end{bmatrix}
$$

where:

- q is the number of simulated cell types, k is the number of latent factors, N is the number of images to be simulated and n is the total number of grid pixels
- $\lambda_{ii} \sim \text{Uniform}(0.5, 1)$ for diagonal elements and $\lambda_{ij} \sim \text{Uniform}(-0.7, 0.7)$ for below-diagonal elements.
- $D \in \mathbb{R}^{n \times n}$ is the pairwise distance matrix.
- $L_j \in \mathbb{R}^{n \times n}$ are the Cholesky factors of the covariance matrices.
- $V_i \in \mathbb{R}^{n \times k}$ are the latent factor matrices.
- $\boldsymbol{A} \in \mathbb{R}^{q \times k}$ is the factor loading matrix.
- X_i is the design matrix.
- β_0 is the intercept term representing the baseline intensity offset.
- $W_i \in \mathbb{R}^{n \times q}$ is the transformed latent space.
- $\boldsymbol{y}_i \in \mathbb{R}^{n \times q}$ is the observed count data modeled via a Poisson likelihood.

B.2 Supplemental Results

B.2.1 Recovering Cross-Correlations When the Number of Latent Factors is Misspecified

In this section, we examine the impact of misspecifying the number of latent factors, k, on the estimation of cross-correlation functions. The true number of latent factors is denoted as $k^* = 3$, and we compare cases where k is under-specified ($k = 2$), correctly specified ($k = 3$), and over-specified ($k = 6$). We also analyze the stability and reliability of these estimates using metrics such as effective sample size (ESS), Gelman-Rubin diagnostic ($\hat{R}$), and Widely Applicable Information Criterion (WAIC).

Figure B.1 shows an example of estimated cross-correlation curves when the number of latent factors is under-specified ($k = 2$). Compared to the true model ($k^* = 3$), the estimates exhibit large deviations for many pairs, indicating underfitting. In contrast, Figure B.2 presents the cross-correlation estimates for the correctly specified model ($k = 3$), demonstrating improved recovery of the underlying spatial dependence structure. When the model is over-specified ($k = 6$), as shown in Figure B.3, the cross-correlation estimates remain largely stable but exhibit slight increases in deviations from the true curves.

To further assess estimation differences, Figure B.4 visualizes the difference in two sets of cross-correlation curves estimated using $k = 2$. This comparison highlights the degree of variability introduced when the number of latent factors is underspecified.

We evaluate the efficiency of cross-correlation estimates by examining their effective sample sizes (ESS) at different spatial distances (Figure B.5). The ESS provides insight into the quality of posterior samples, with lower values indicating higher autocorrelation and potential inefficiencies in MCMC sampling. Additionally, Figure B.6 presents the Gelman-Rubin diagnostic ($\hat{R}$) across varying k, where values close to 1 indicate convergence.

To assess model fit, Figure B.7 shows the distribution of WAIC values across different choices of k. The lowest WAIC values correspond to the correctly specified model ($k = 3$), supporting its selection as the optimal choice. However, models with $k = 6$ have WAIC values very close to those from models with $k = 3$.

Next, Figure B.8 provides an example of an MCMC trace plot for estimated mean-centered cross-correlations at a spatial distance of $h = 767.6\mu m$, when $k = 2$. The trace plot reveals the mixing behavior of the MCMC chains, which is crucial for ensuring robust inference.

Finally, we show the model fitting times for various k in Figure B.9, demonstrating the linear increase in fitting time as k increases.

These results collectively illustrate how misspecifying k influences cross-correlation estimation and highlight the importance of model selection criteria such as WAIC in guiding the choice of an appropriate number of latent factors.

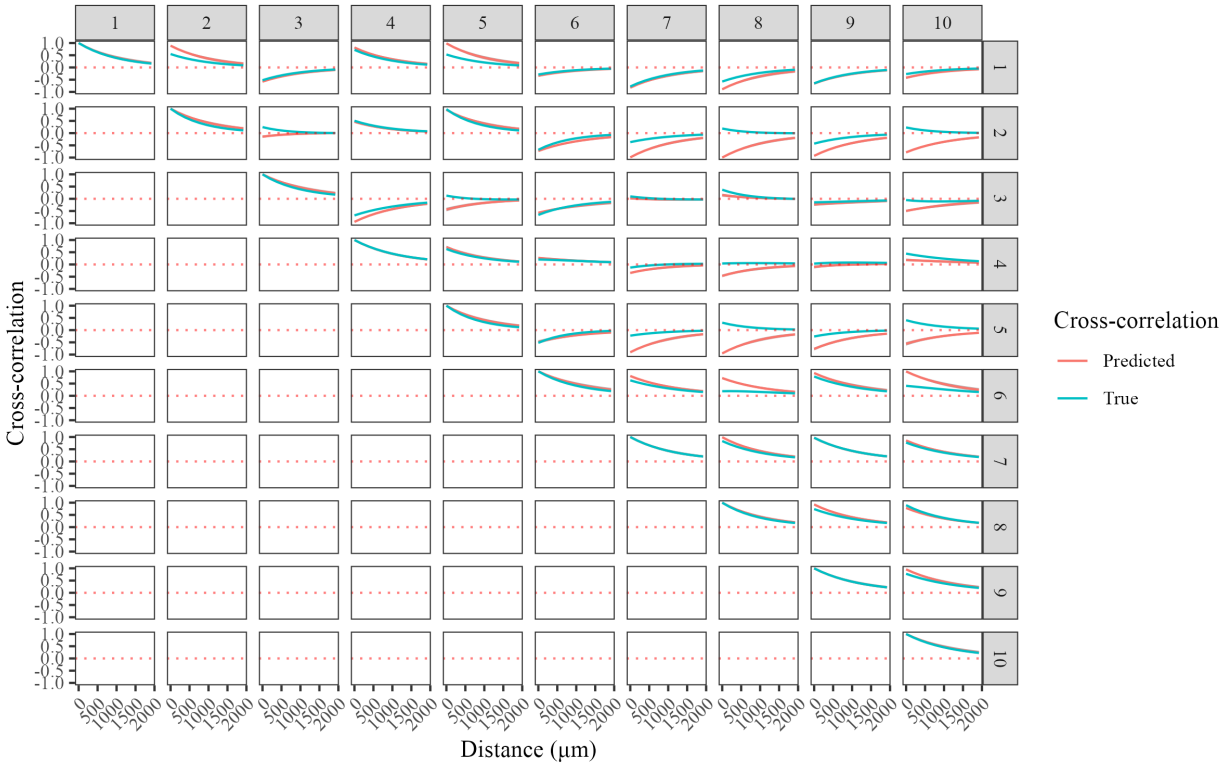


Figure B.1: An example of cross-correlation curves estimated using $k = 2$, when $k^* = 3$.

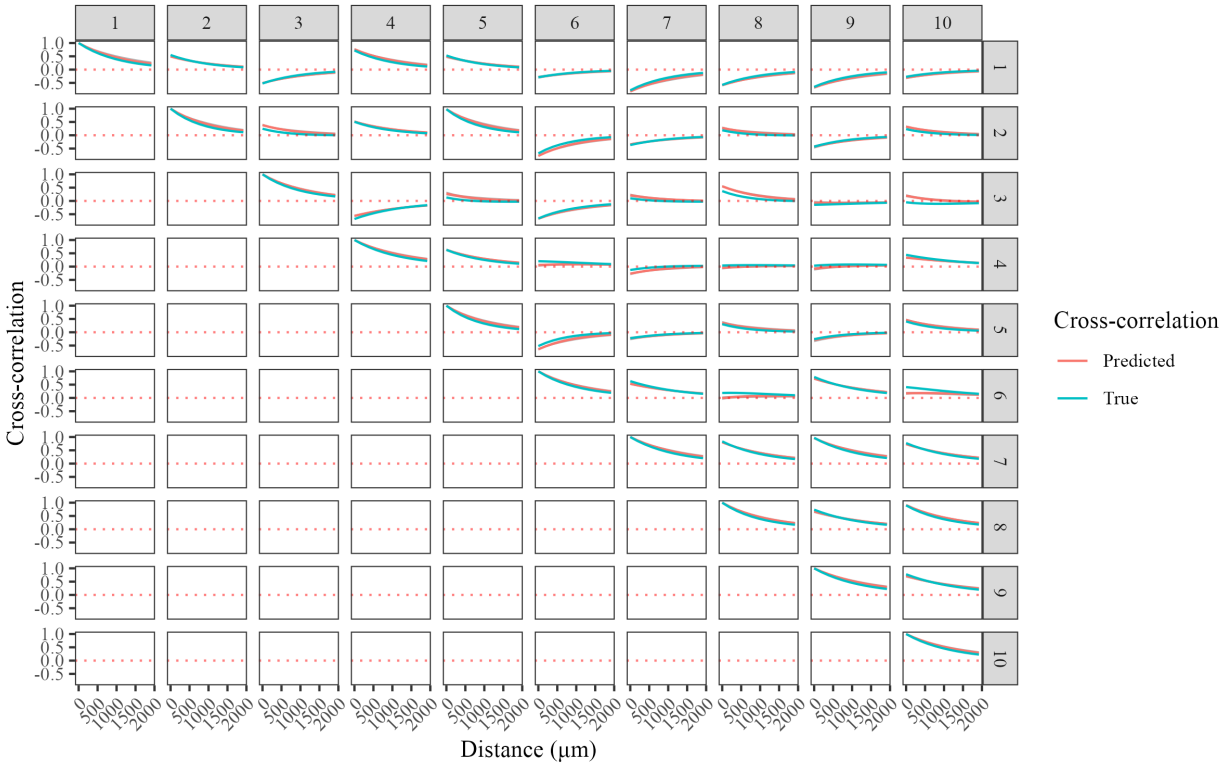


Figure B.2: An example of cross-correlation curves estimated using $k = 3$, when $k^* = 3$.

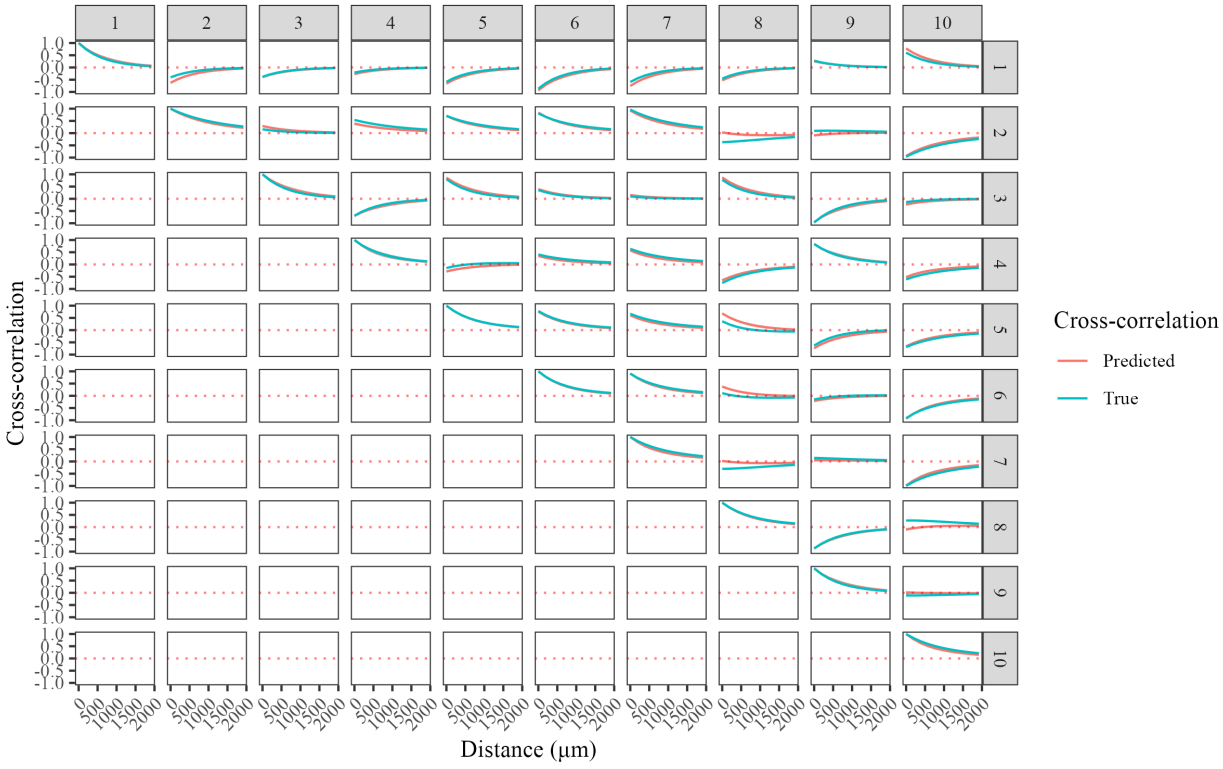


Figure B.3: An example of cross-correlation curves estimated using $k = 6$, when $k^* = 3$.

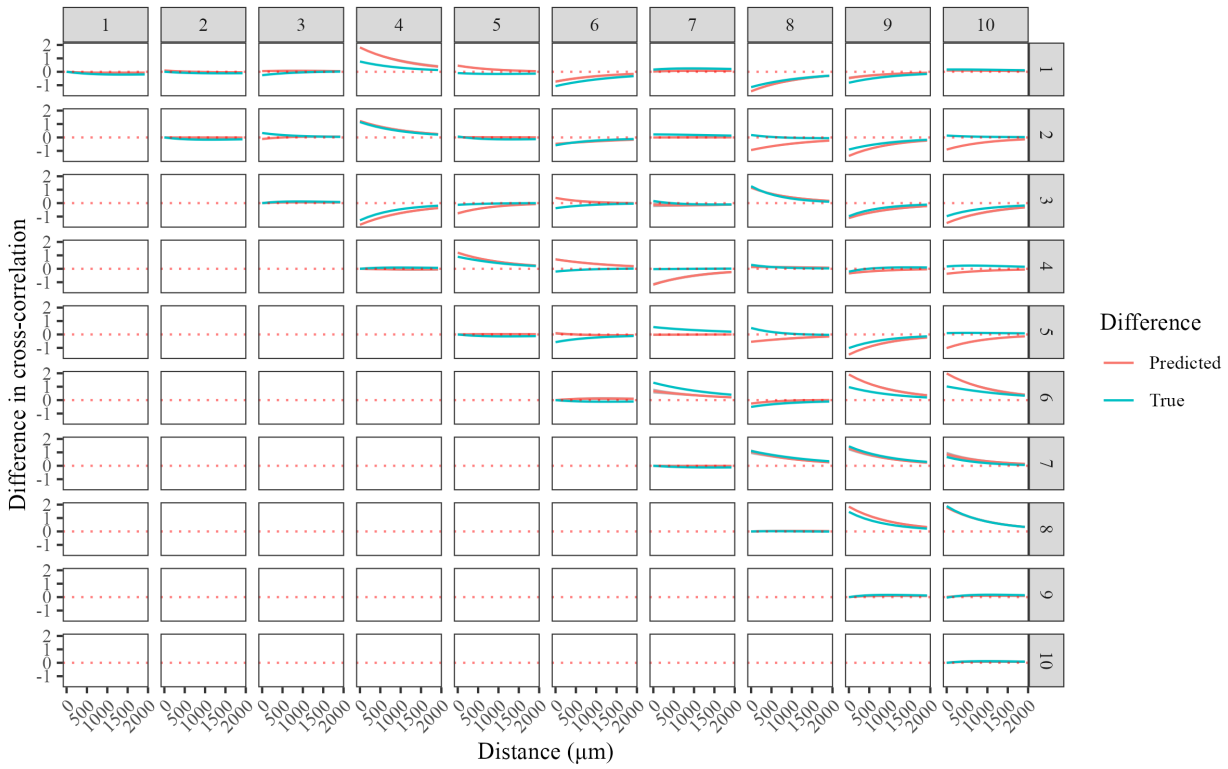


Figure B.4: An example of the difference in two sets of cross-correlation curves that were both estimated using $k = 2$, when $k^* = 3$.

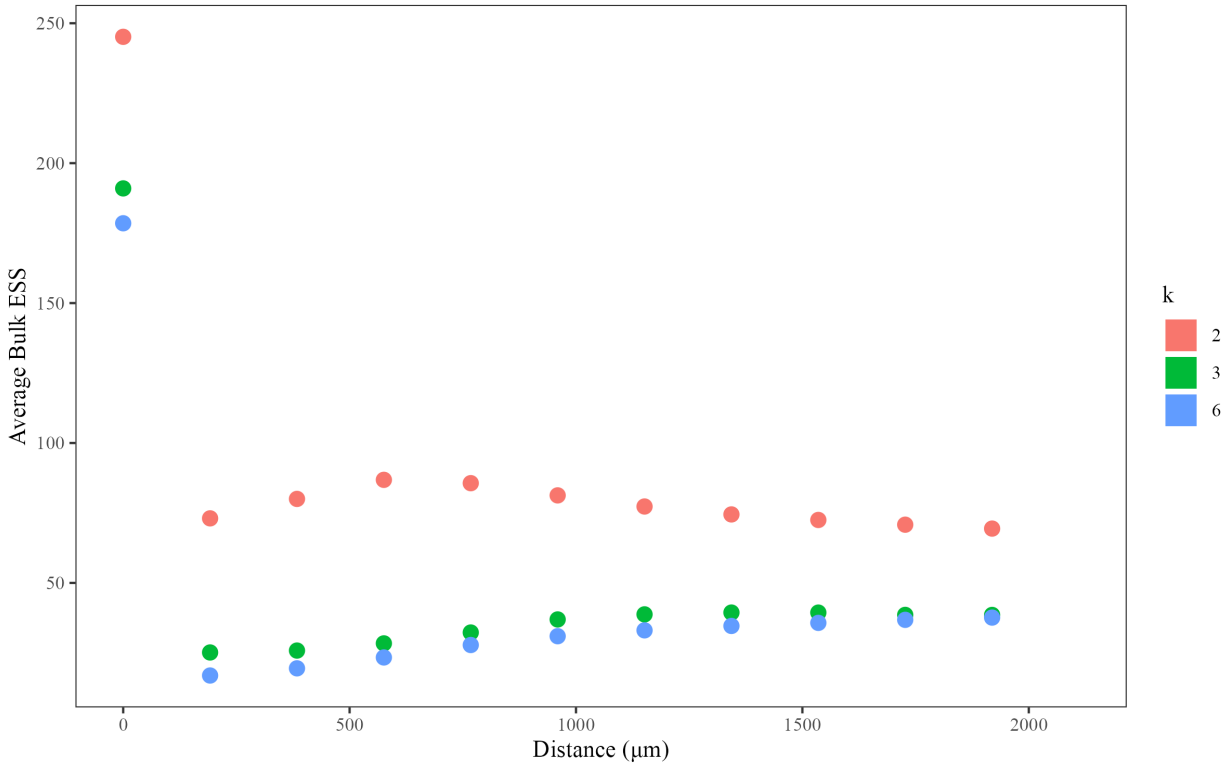


Figure B.5: Effective sample sizes (ESS) of cross-correlation estimates at different spatial distances, over varying k.

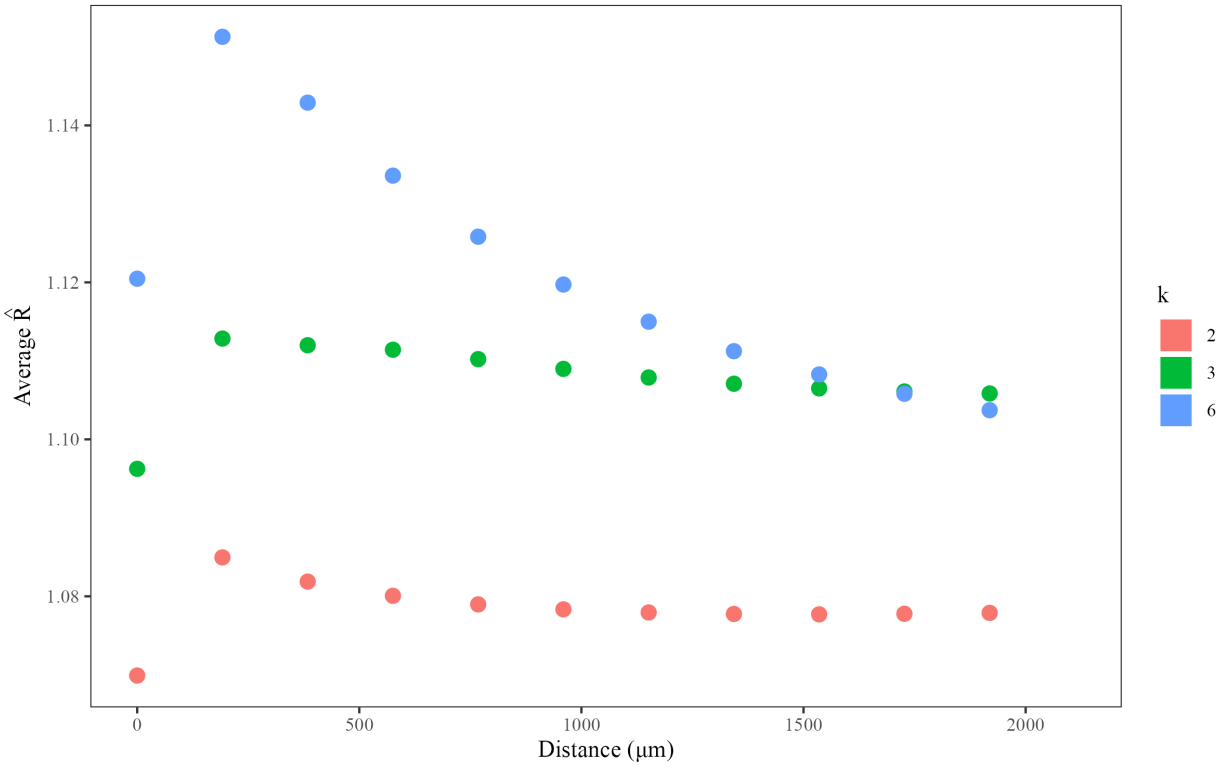


Figure B.6: $\hat{R}$ of cross-correlation estimates at different spatial distances, over varying k.

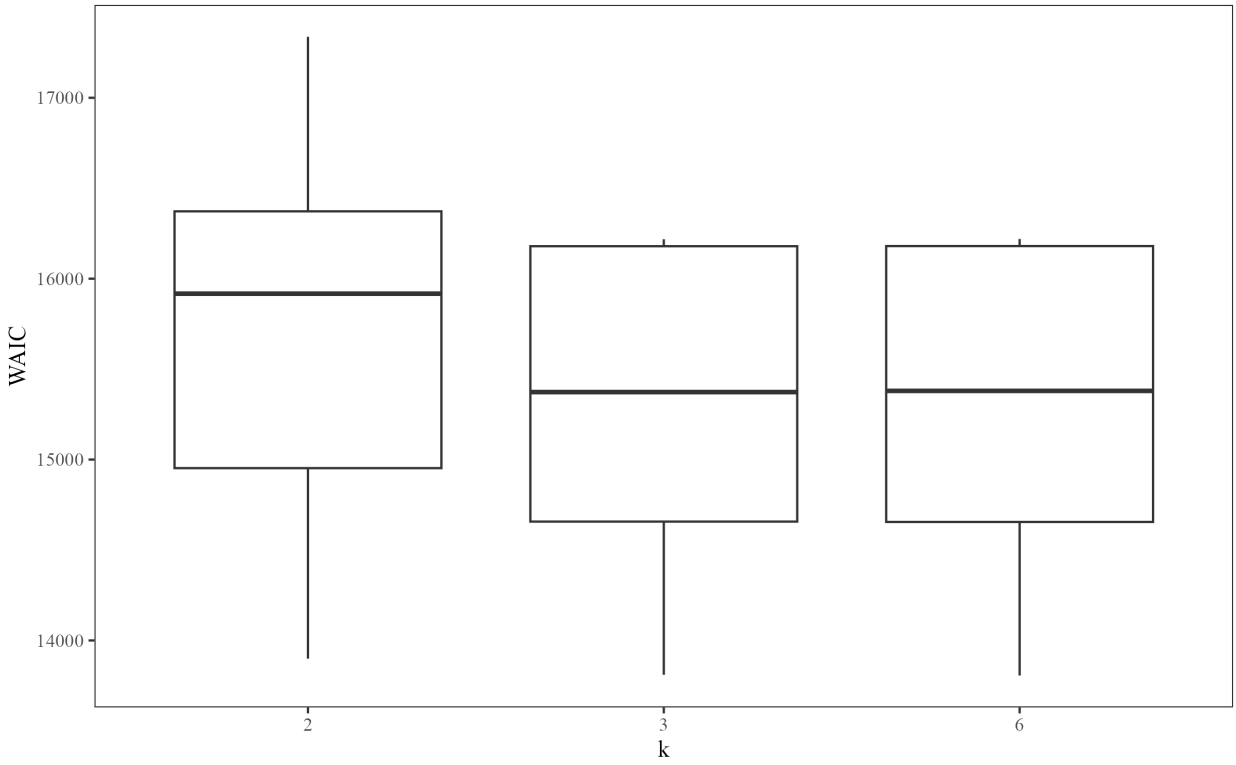


Figure B.7: The distribution of WAIC for different k, when k^*=3.

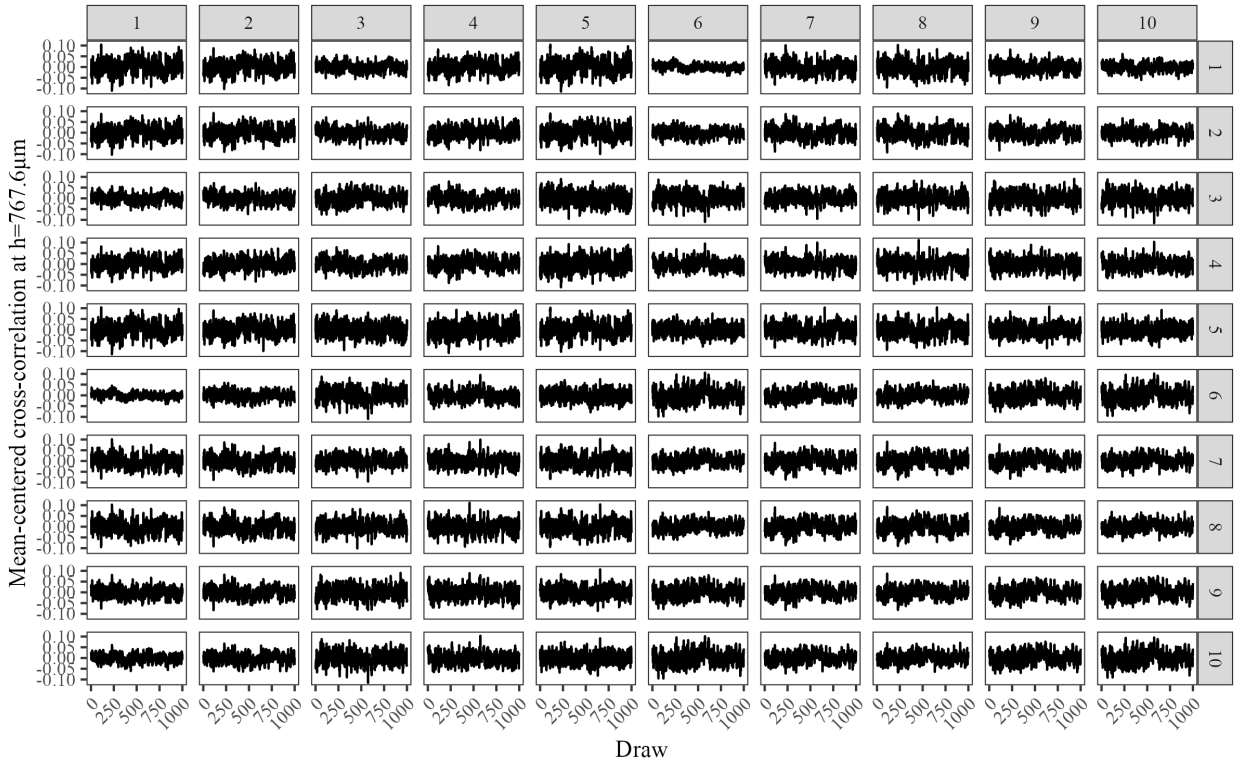


Figure B.8: Example of MCMC trace plot for estimated mean-centered cross-correlations at $h = 767.6\mu m$, using $k = 2$.

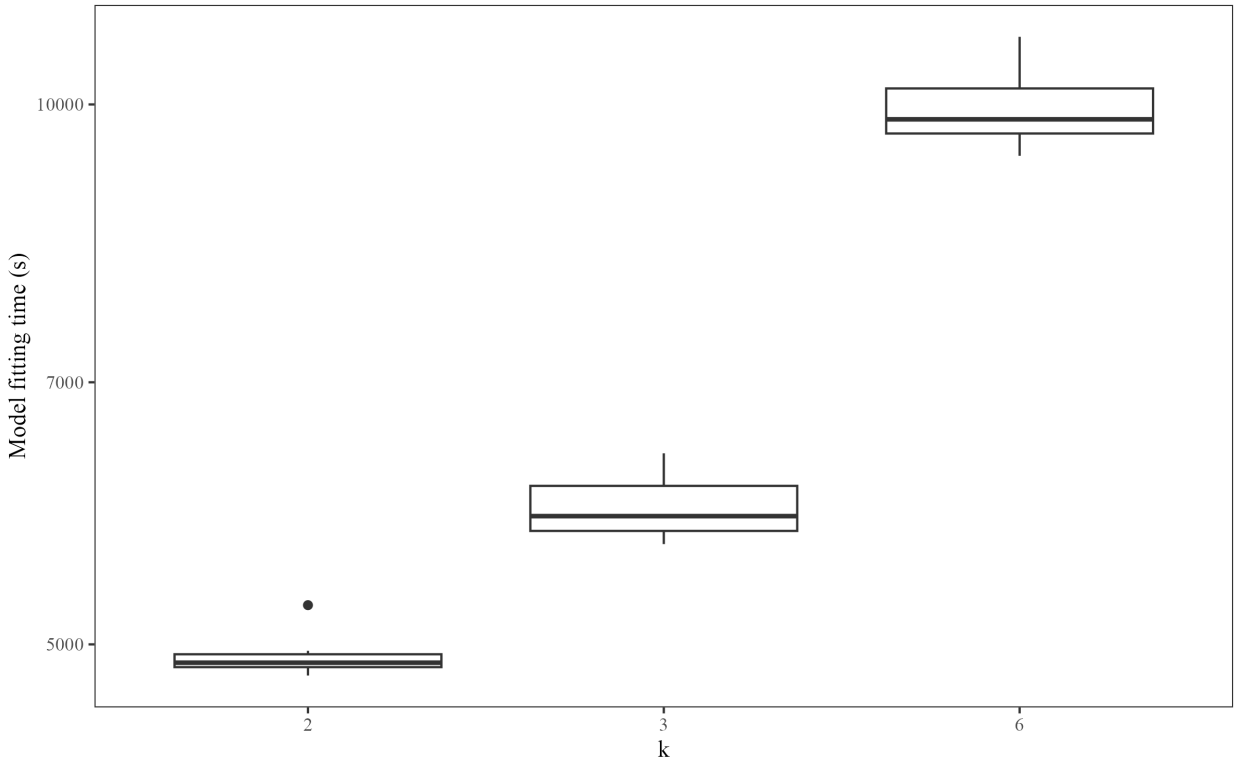


Figure B.9: Model fitting times for simulations at various k.

B.2.2 The Effect of Grid Size on Cross-Correlation Recovery

In this section, we assess how the choice of grid size, n_x, affects the recovery of cross-correlation functions. The grid size determines the resolution of spatial intensity estimates, impacting both estimation accuracy and computational efficiency. We consider three different grid sizes: $n_x \in 12, 24, 48$.

Figure B.10 compares the estimated cross-correlation curves for these different grid sizes. The results demonstrate that there are not large differences in estimated cross-correlation curves across various values of n_x.

To evaluate sampling efficiency, Figure B.11 presents the effective sample sizes (ESS) for different spatial distances across varying n_x. Lower grid resolutions tend to yield larger ESS values, suggesting more efficient MCMC sampling and reduced autocorrelation in posterior estimates.

Convergence diagnostics are shown in Figure B.12, which displays the Gelman-Rubin diagnostic ($\hat{R}$) across different grid sizes. Values of $\hat{R}$ close to 1 indicate satisfactory convergence, and we observe that coarser grids ($n_x = 12$) maintain the best convergence properties.

Lastly, we show the model fitting times for various n_x in Figure B.13, demonstrating the quadratic increase in fitting time as n_x increases.

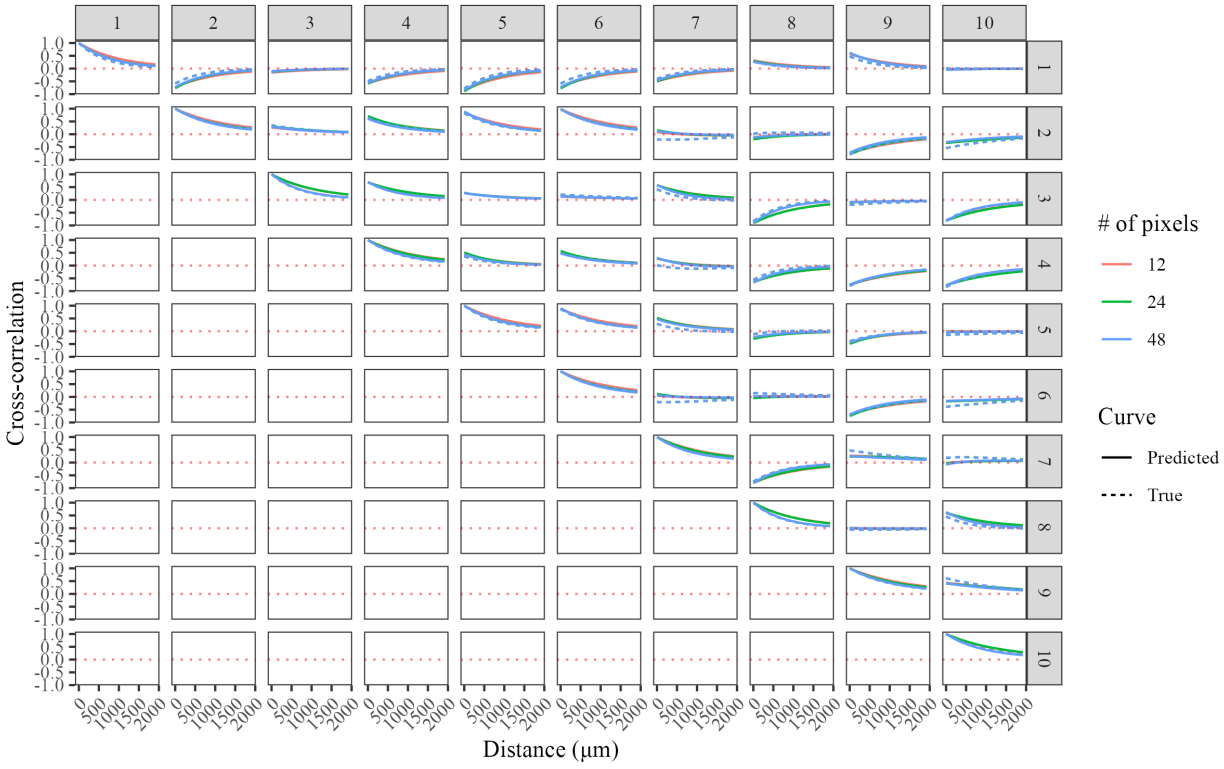


Figure B.10: A comparison of cross-correlation curves estimated using $n_x \in \{12, 24, 48\}$

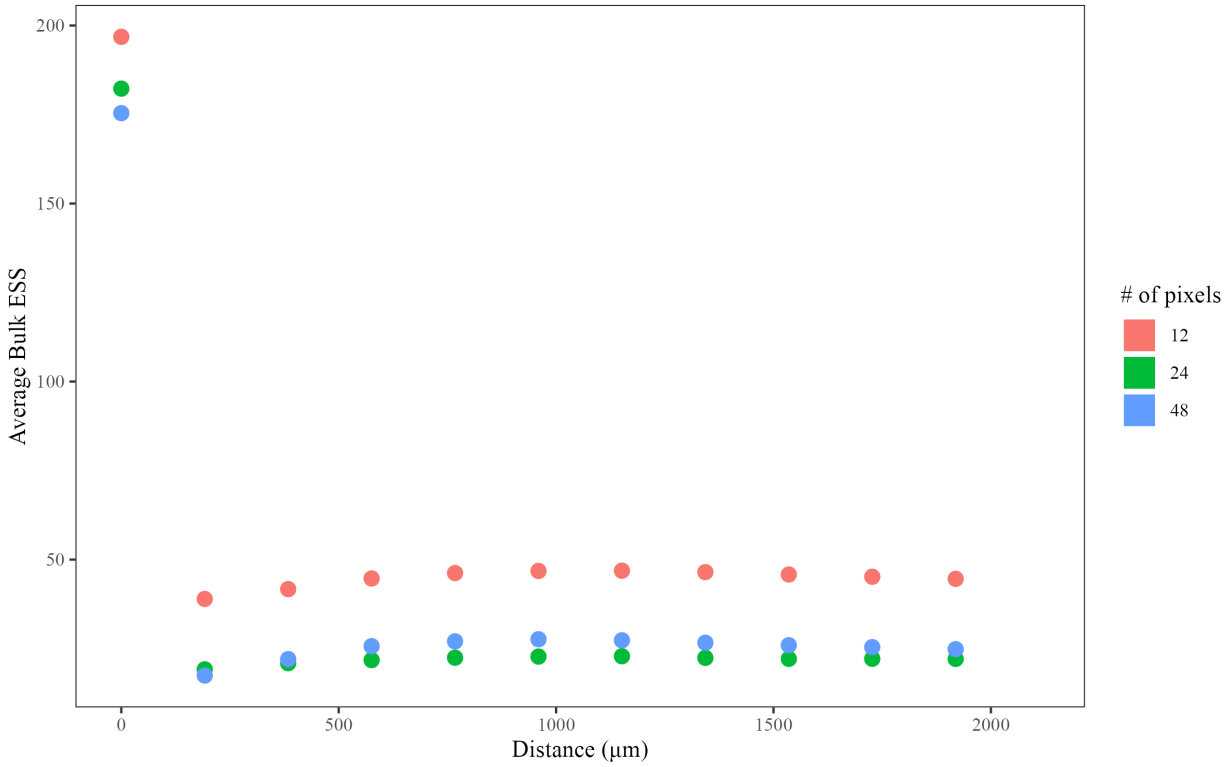


Figure B.11: Effective sample sizes (ESS) of cross-correlation estimates at different spatial distances, over varying n_x.

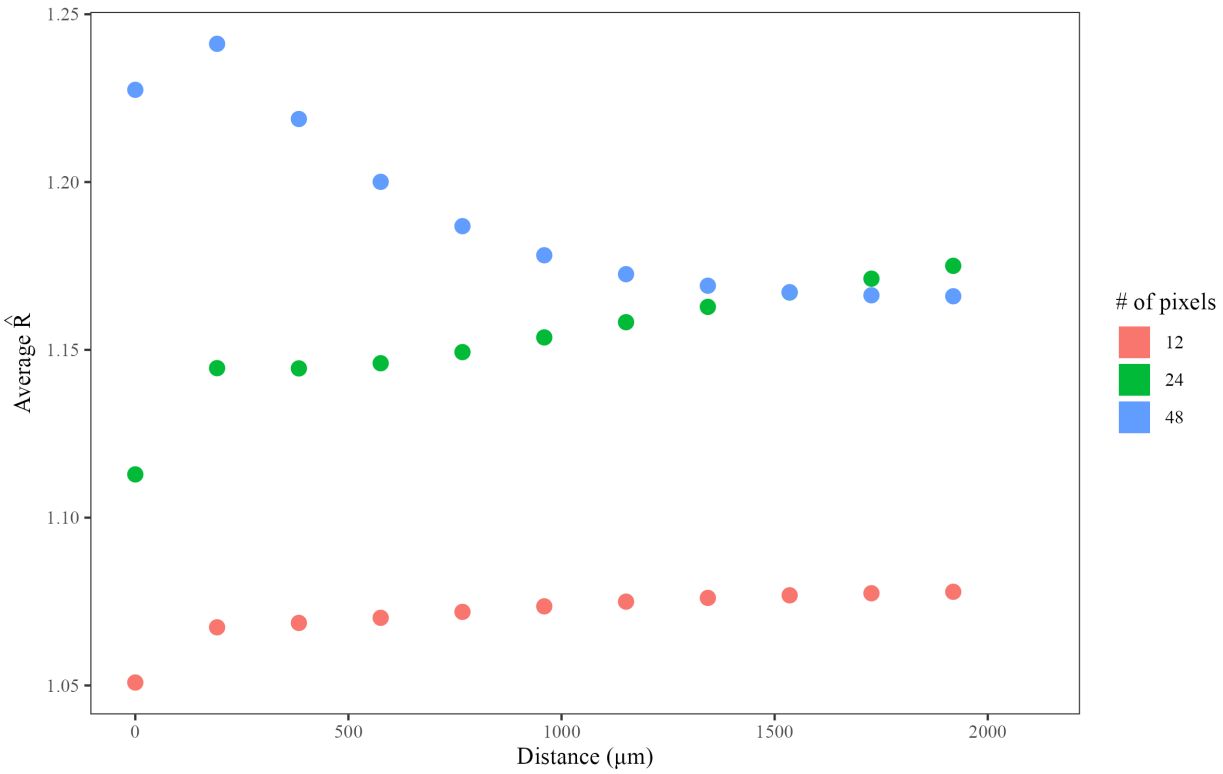


Figure B.12: $\hat{R}$ of cross-correlation estimates at different spatial distances, over varying n_x.

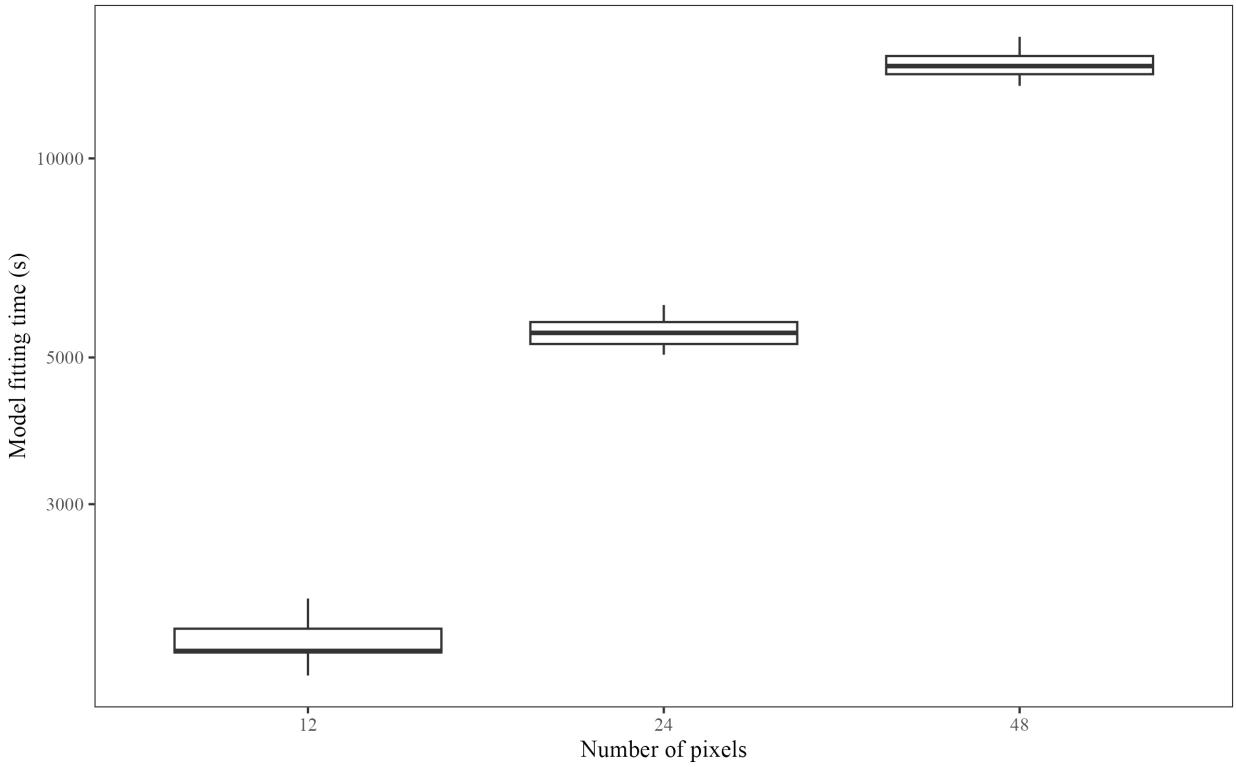


Figure B.13: Model fitting times for simulations at various n_x.

B.2.3 Choosing k for PANC Analysis

In this section, we evaluate the selection of the number of latent factors, k, for pancreatic cancer (PANC) analysis by examining model convergence and the stability of cross-correlation estimates. We analyze two patient groups: PDAC (pancreatic ductal adenocarcinoma) and IPMN (intraductal papillary mucinous neoplasm).

Table B.1 shows the model fitting times for various k in both patient groups.

Figures B.14 and B.15 display the Gelman-Rubin diagnostic ($\hat{R}$) values for cross-correlation estimates at different spatial distances across various k in the PDAC and IPMN patient groups, respectively. Convergence is indicated by $\hat{R}$ values close to 1. Our results show that convergence results are mixed for various values of k.

Figures B.16 and B.17 presents the estimated cross-correlations at various spatial distances for different choices of k in the PDAC and IPMN patient groups. The results indicate that cross-correlation estimates remain relatively stable for moderate to large values of k, whereas overly small values introduce larger deviations in the estimates.

Table B.1: Model fitting times on images from pancreatic cancer patient groups for various k.

	Model fitting time (s)	
k	**PDAC**	**IPMN**
2	11934	7522
3	14852	9253
4	18388	11307
5	21354	14258

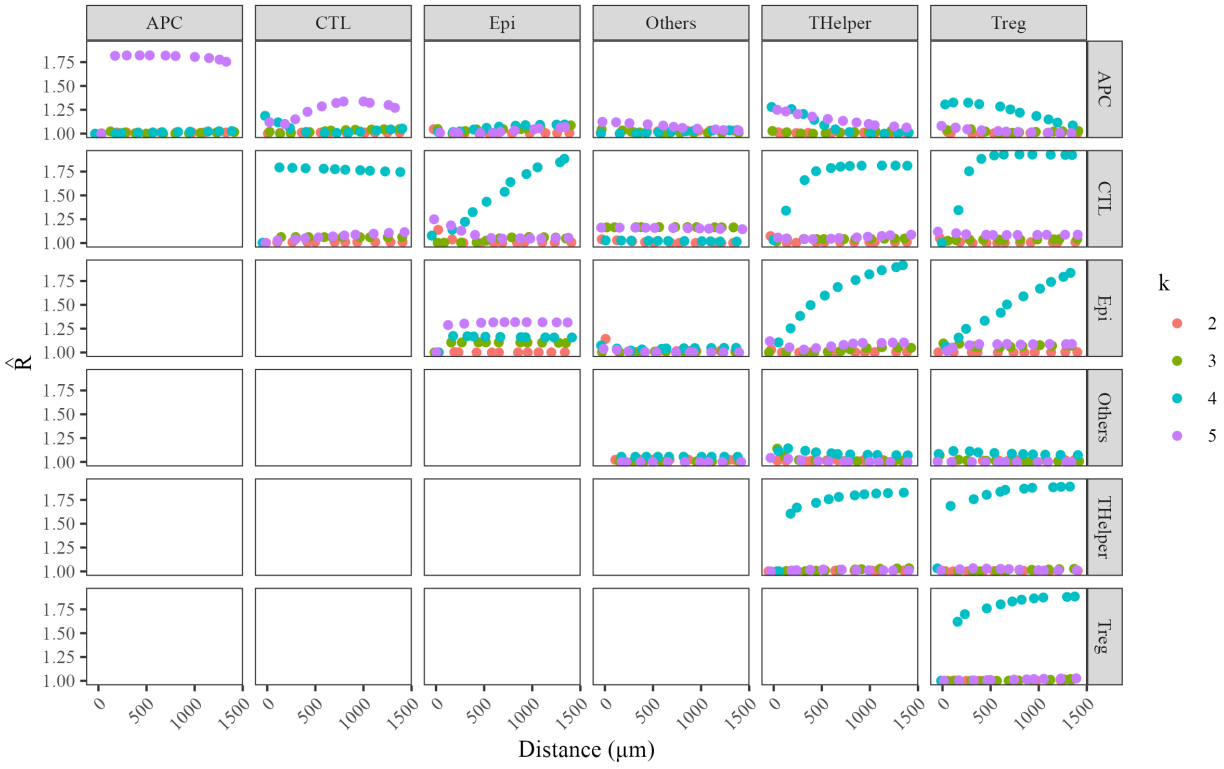


Figure B.14: $\hat{R}$ values to determine convergence of MCMC sampling of cross-correlations at various distances, under various k, in the PDAC patient group. Values closer to 1 indicate better convergence.

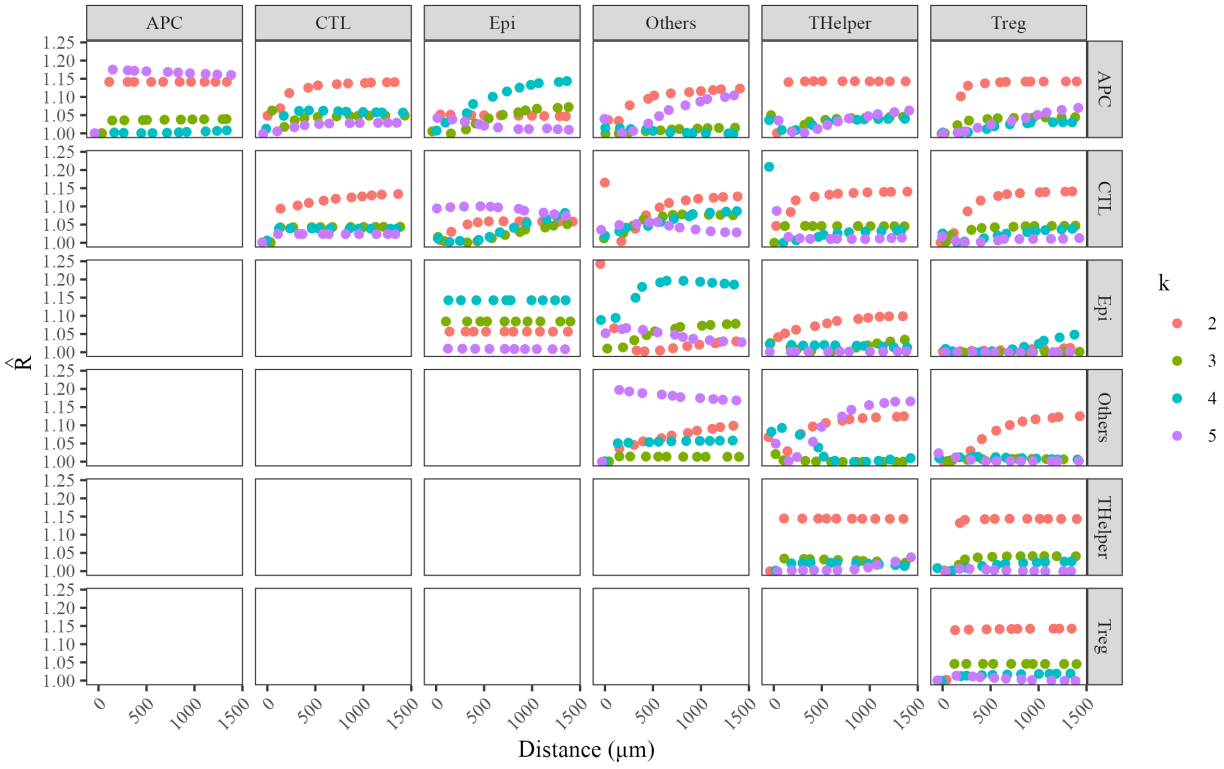


Figure B.15: $\hat{R}$ values to determine convergence of MCMC sampling of cross-correlations at various distances, under various k, in the IPMN patient group. Values closer to 1 indicate better convergence.

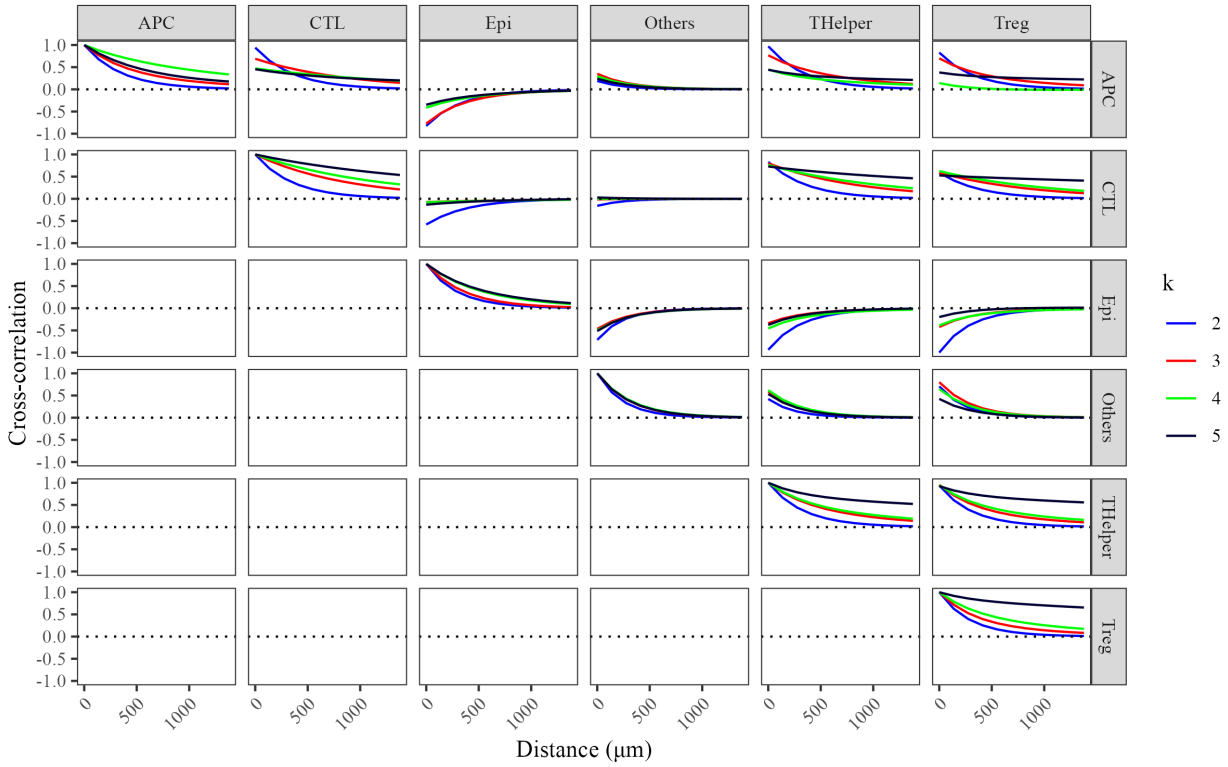


Figure B.16: Estimated cross-correlations at various distances, estimated at various k, in the PDAC patient group.

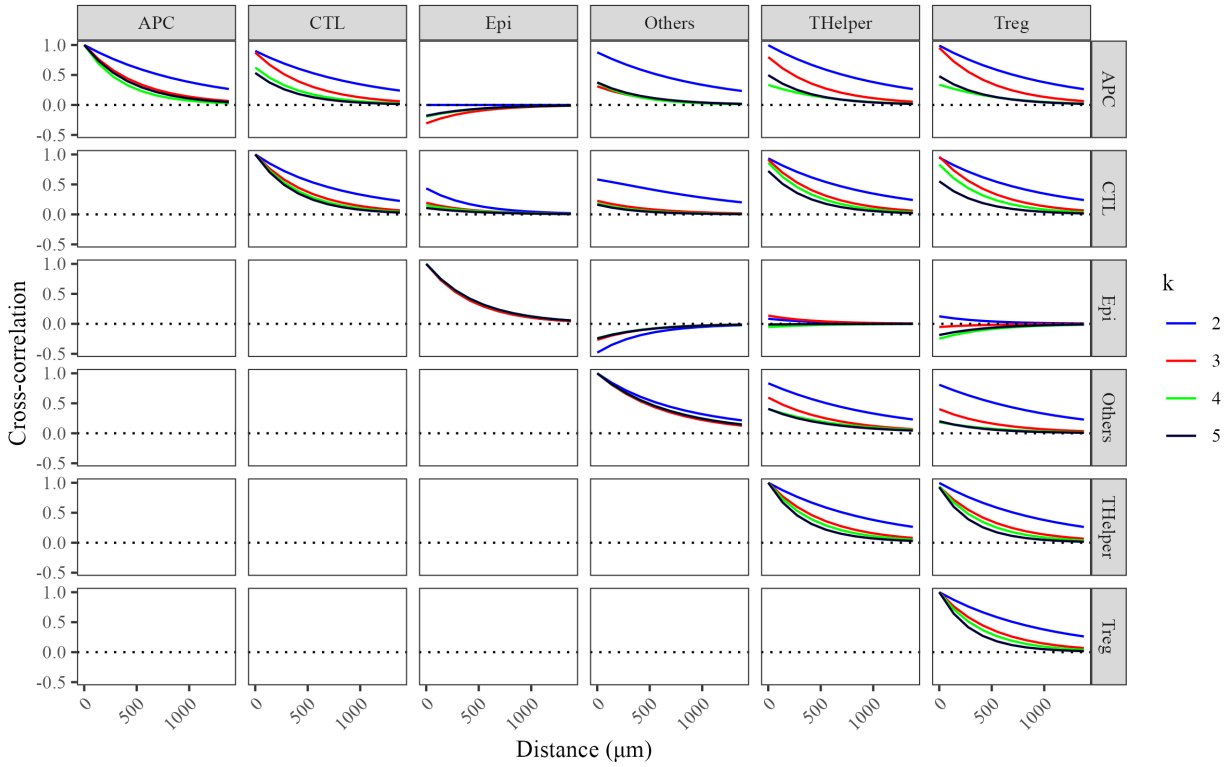


Figure B.17: Estimated cross-correlations at various distances, estimated at various k, in the IPMN patient group.

B.2.4 Analysis of PANC Dataset

In this section, we analyze the cross-correlation differences between patient groups in the pancreatic cancer (PANC) dataset.

Figure B.18 shows the esimated mean of the zero-distance cross-correlations between each pair of cell types for both patient groups. One interesting thing to note is the exclusion of epithelial cells from all other cell types, in both groups.

Figure B.19 presents the estimated difference in cross-correlation between the IPMN and PDAC patient groups for the same selected cell types as shown in the main article. These differences provide insight into how spatial cellular organization varies between disease subtypes, highlighting potential tumor-immune interactions that differentiate immune-permissive and immune-exclusive microenvironments.

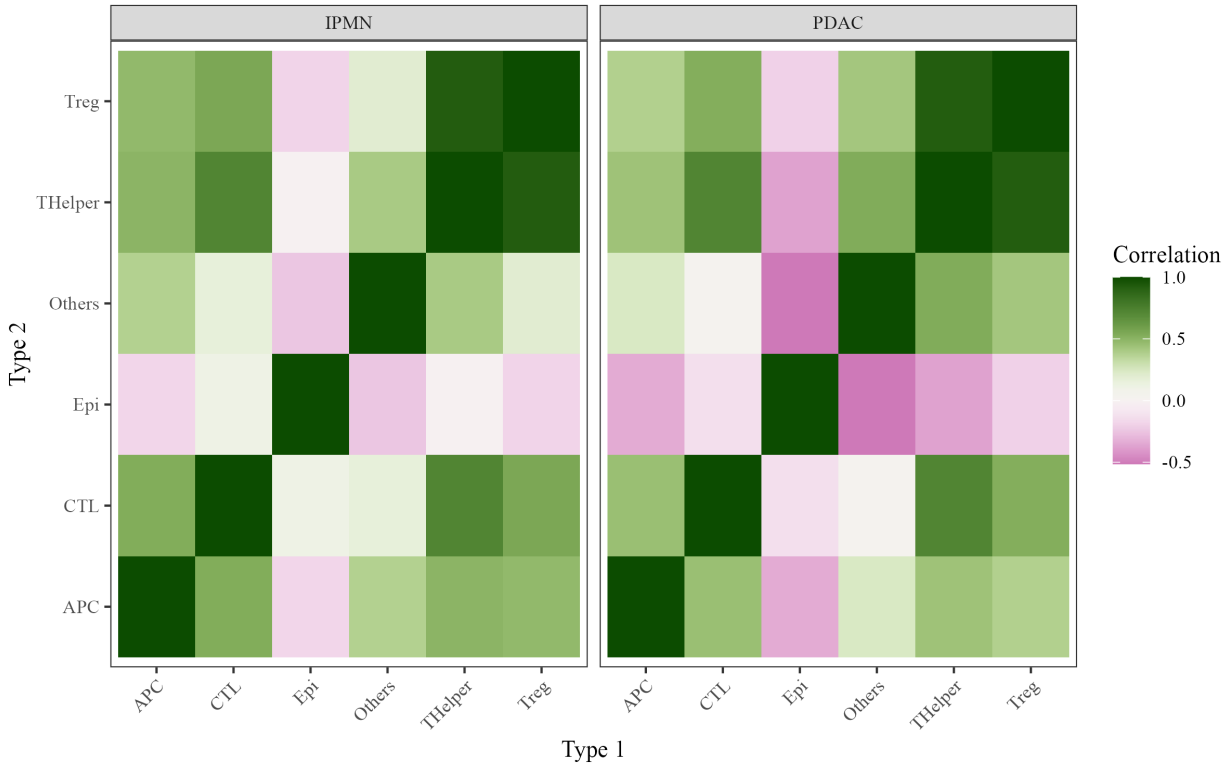


Figure B.18: Estimated posterior mean of the zero-distance ($h = 0$) cross-correlation matrix between the different cell types in both patient groups from the pancreatic cancer dataset.

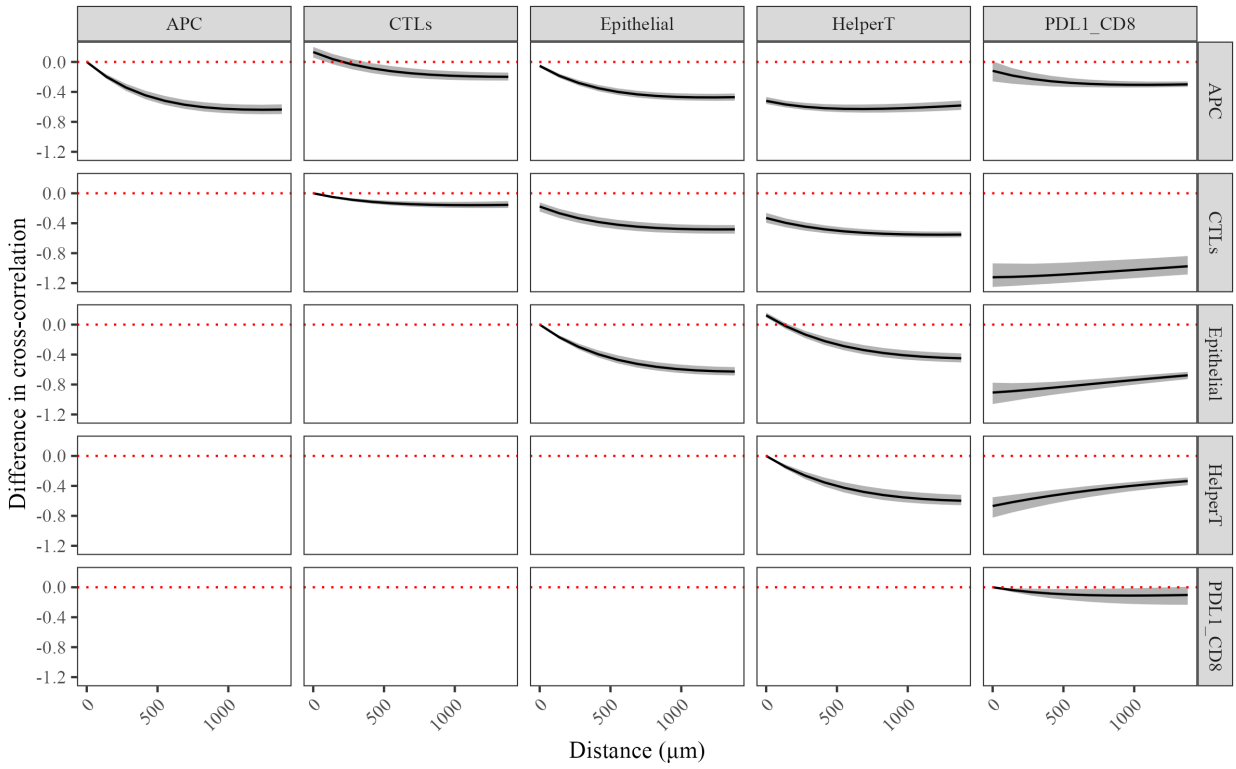


Figure B.19: The estimated difference in cross-correlation between patient groups for the select cell types. The black line in each plot indicates the posterior expectation, while the shaded region indicates the pointwise 95% credible interval.

B.2.5 Choosing k for CRC Analysis

In this section, we assess the selection of the number of latent factors, k, for colorectal cancer (CRC) analysis by evaluating model convergence and the stability of cross-correlation estimates across two patient groups: CLR and DII.

Table B.2 shows the model fitting times for various k in both patient groups.

Figures B.20 and B.21 present the Gelman-Rubin diagnostic ($\hat{R}$) values for cross-correlation estimates at different spatial distances under varying k in the CLR and DII patient groups, respectively. Values closer to 1 indicate better convergence, and the results suggest that low-to-moderate values of k lead to more stable MCMC performance.

Figures B.22 and B.23 display estimated cross-correlations at different spatial distances for various k values in the CLR and DII patient groups, respectively. These estimates demonstrate spatial cross-correlation curves estimated from models with low values of k ($k = 2, 4$) have larger deviations, while those with $k \geq 6$ seem to have converged.

Finally, Figure B.24 illustrates the differences in estimated cross-correlations between the two patient groups across varying k. Again, it seems that spatial cross-correlation curves estimated from models with low values of k ($k = 2, 4$) have larger deviations, while those with $k \geq 6$ seem to have converged.

Table B.2: Model fitting times on images from colorectal cancer patient groups for various k.

	Model fitting time (s)	
k	**CLR**	**DII**
2	10046	11410
4	14703	15898
6	21533	23235
8	30255	32648
10	43254	45180



Figure B.20: $\hat{R}$ values to determine convergence of MCMC sampling of cross-correlations at various distances, under various k, in the CLR patient group. Values closer to 1 indicate better convergence.



Figure B.21: $\hat{R}$ values to determine convergence of MCMC sampling of cross-correlations at various distances, under various k, in the DII patient group. Values closer to 1 indicate better convergence.

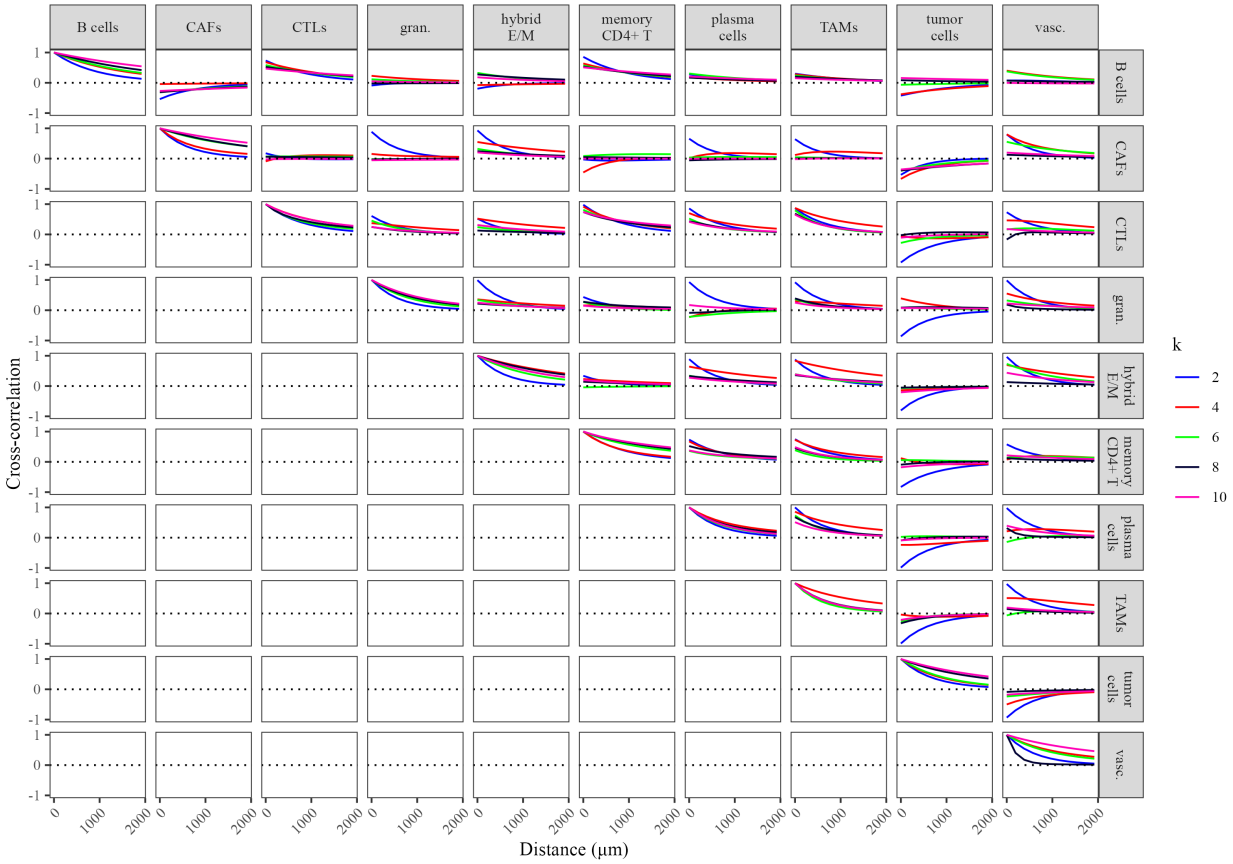


Figure B.22: Estimated cross-correlations at various distances, estimated at various k, in the CLR patient group.

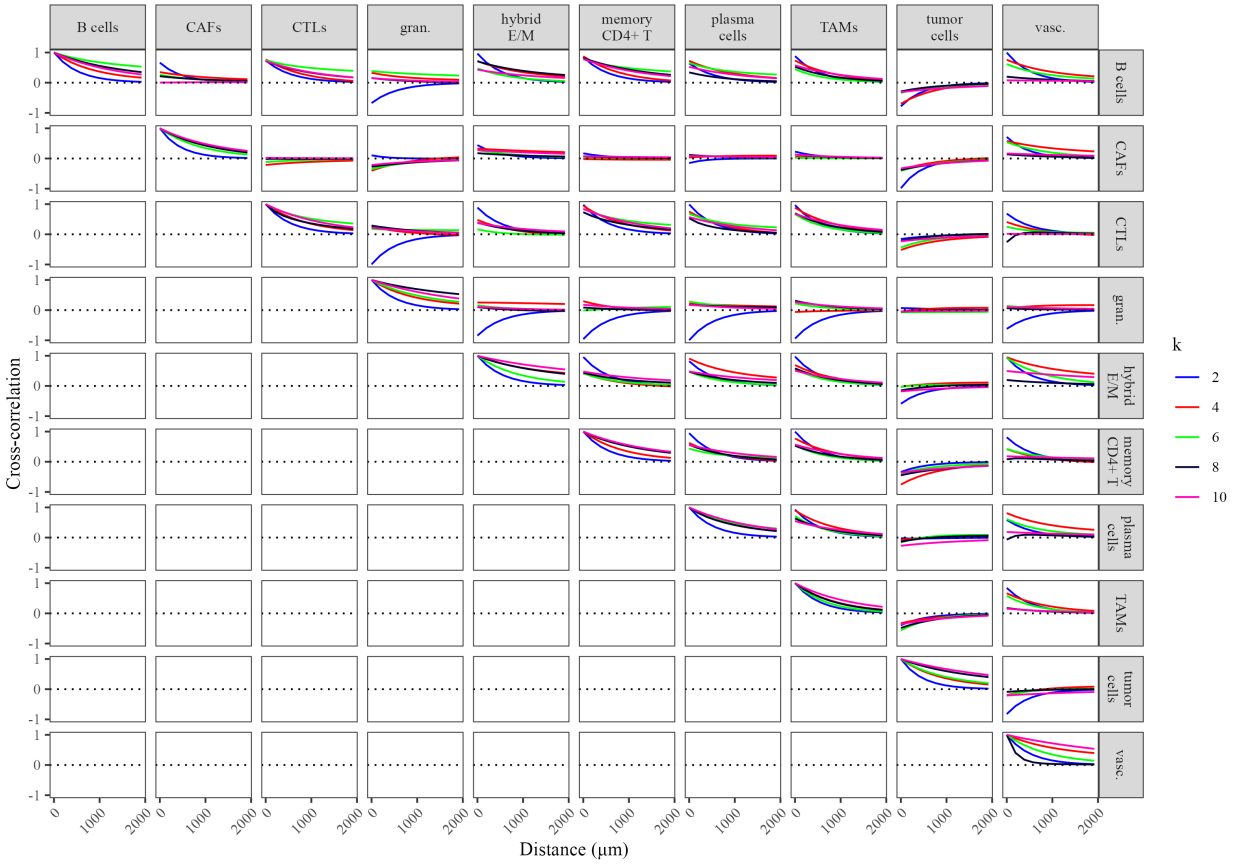


Figure B.23: Estimated cross-correlations at various distances, estimated at various k, in the DII patient group.

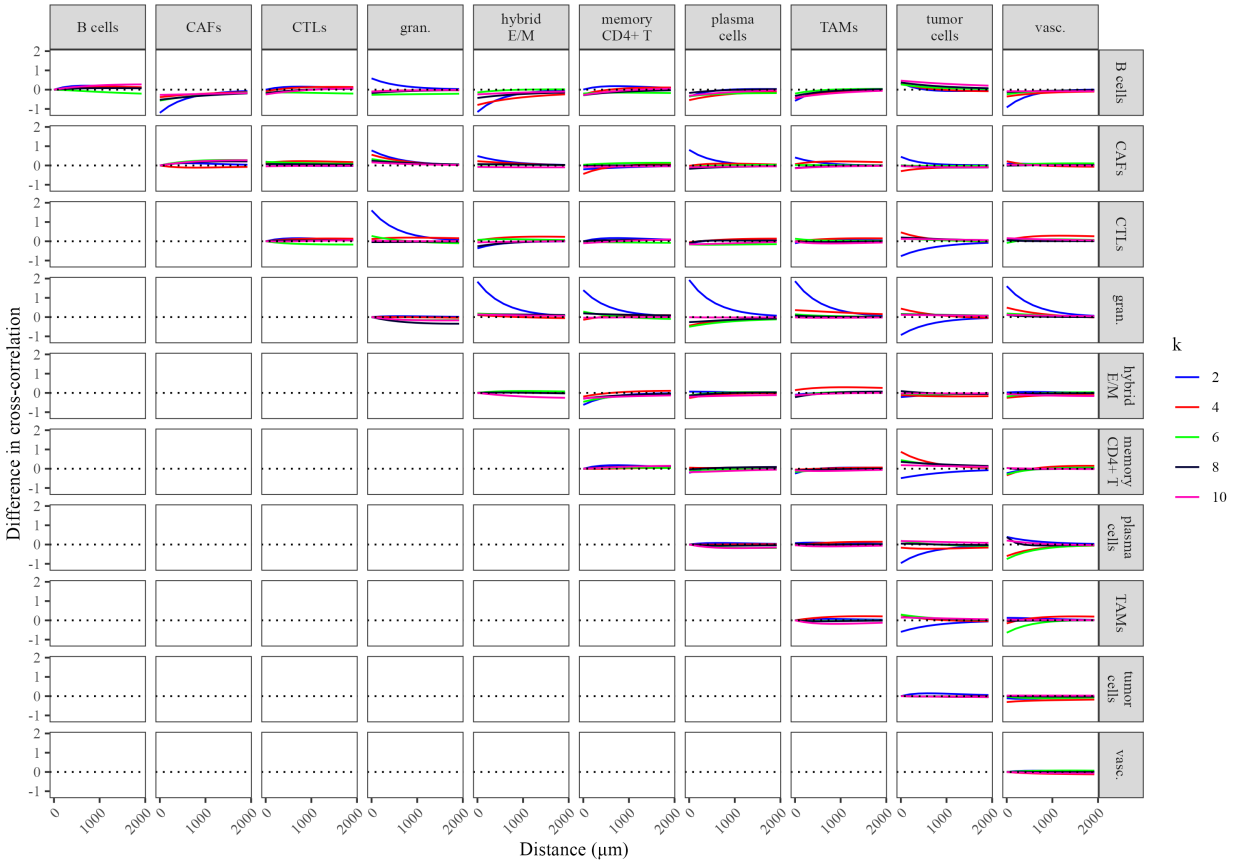


Figure B.24: Difference in estimated cross-correlations at various distances, estimated at various k, between the two patient groups.

B.2.6 Analysis of CRC Dataset

In this section, we analyze spatial cross-correlations in the colorectal cancer (CRC) dataset, comparing cellular interactions between the CLR and DII patient groups.

Figure B.25 presents the estimated spatial cross-correlations for all cell-type combinations in both patient groups.

Figure B.26 displays the estimated difference in cross-correlation between the two patient groups for the same selected cell types as in the main article. This comparison highlights key differences in spatial organization between the CLR and DII groups, potentially revealing distinct immune or stromal interactions associated with each subtype.

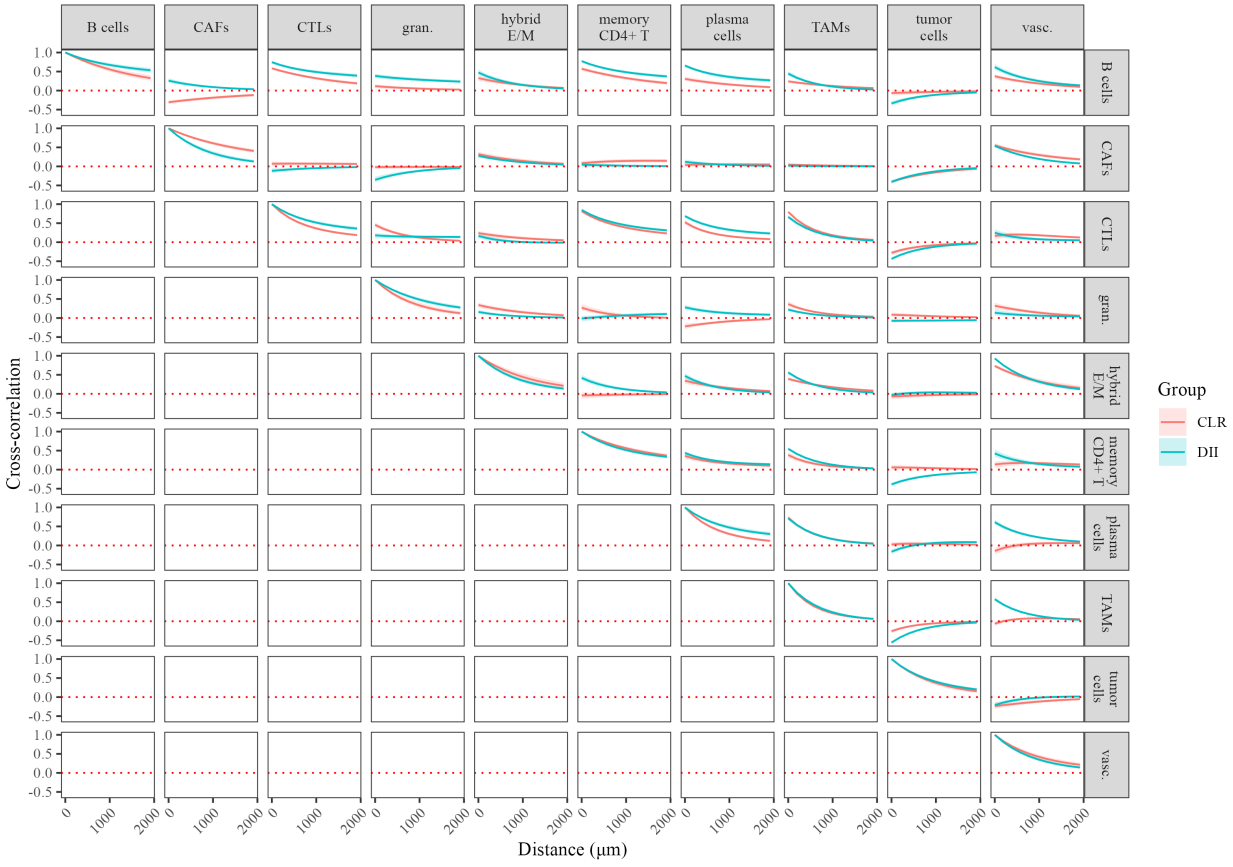


Figure B.25: The estimated spatial cross-correlations between select pairs of cell-type combinations from patients in the CLR group and the DII group. The colored line in each plot indicates the posterior expectation, while the shaded region indicates the pointwise 95% credible interval.

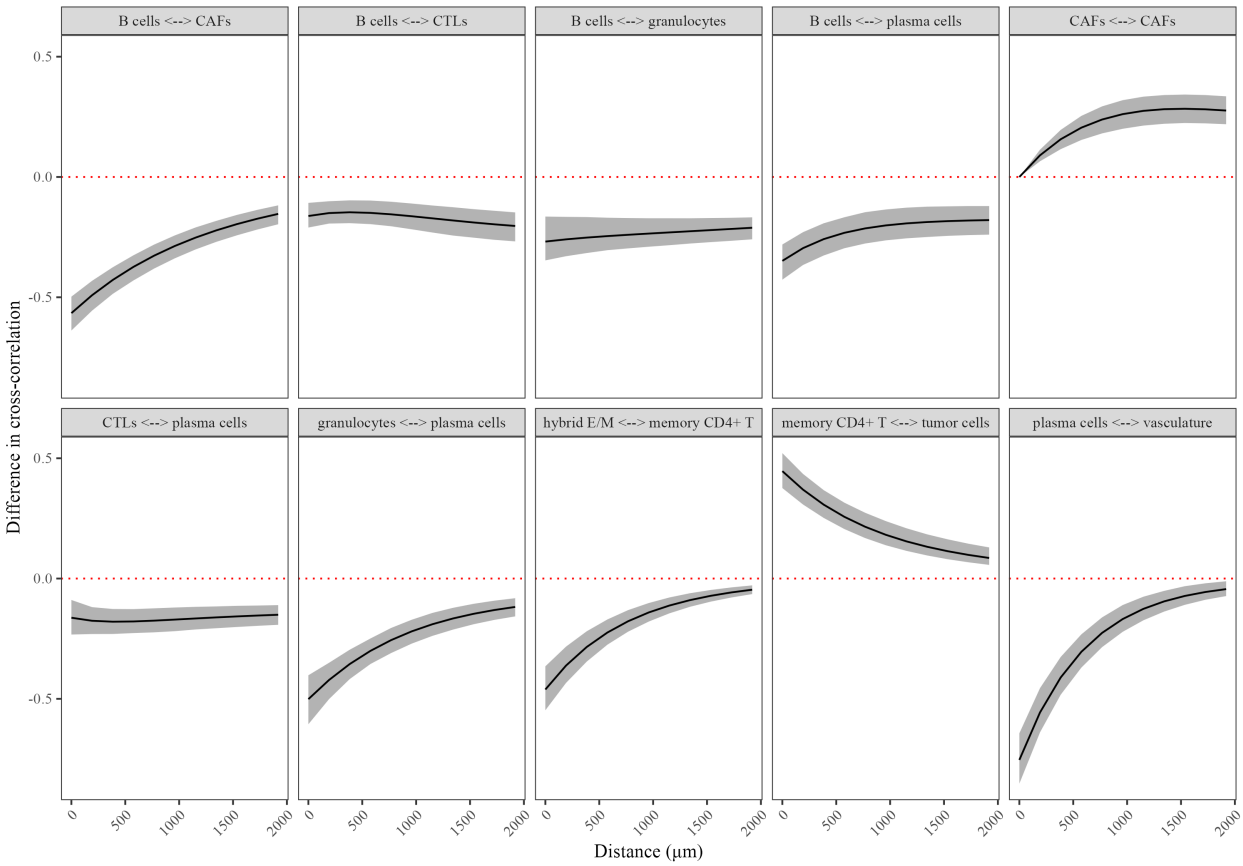


Figure B.26: The estimated difference in cross-correlation between patient groups for the same selected cell types as in the main article. The black line in each plot indicates the posterior expectation, while the shaded region indicates the pointwise 95% credible interval.

B.3 Markov-chain Monte Carlo for posterior sampling

We seek to sample from the posterior distribution of the following model[1]

$$\begin{aligned} y_{ij}(\boldsymbol{l}) \mid \lambda_{ij}(\boldsymbol{l}) &\sim \text{Poisson}(\lambda_{ij}(\boldsymbol{l})), & \boldsymbol{\lambda}_i(\boldsymbol{l}) &= \left[\lambda_{i1}(\boldsymbol{l}) \cdots \lambda_{iq}(\boldsymbol{l})\right]^\top \\ \log \boldsymbol{\lambda}_i(\boldsymbol{l}) &= \boldsymbol{\alpha}_i + \boldsymbol{X}_i(\boldsymbol{l})\boldsymbol{B} + \boldsymbol{w}_i(\boldsymbol{l}), & \boldsymbol{w}_i(\boldsymbol{l}) &= \boldsymbol{A}\boldsymbol{v}_i(\boldsymbol{l}), \\ \boldsymbol{v}_i(\boldsymbol{l}) &= \left[v_{i,1}(\boldsymbol{l}), \ldots, v_{i,k}(\boldsymbol{l})\right], ^\top & v_{i,j}(\cdot) &\sim GP(\boldsymbol{0}, \rho(\cdot,\cdot;\varphi_j)), \end{aligned} \tag{B.1}$$

where $y_{ij}(\boldsymbol{l})$ is the number of cells of type j at location $\boldsymbol{l}$ in the ith image. The MCMC algorithm we use for fitting our Bayesian hierarchical model is a straightforward multi-subject generalization of Algorithm 3 in Peruzzi and Dunson, 2024. The key differences are as follows:

- The joint update of the jth column of $\boldsymbol{B}$ (labeled $\boldsymbol{B}_{\cdot j}$) and jth row of $\boldsymbol{A}$ (labeled $\boldsymbol{A}_{j\cdot}$) in step 1 targets the full conditional posterior

$$N(\boldsymbol{F}_j; \boldsymbol{0}, \boldsymbol{V}_F) \prod_{i=1}^{N} \prod_{\boldsymbol{l}} \text{Poisson}(y_{ij}(\boldsymbol{l}); \lambda_{ij}(\boldsymbol{l})),$$

  where $\boldsymbol{F}_j = (\boldsymbol{B}_{\cdot j}^\top, \boldsymbol{A}_{j\cdot})^\top$, we fix $\boldsymbol{V}_F = \text{diag}\{10^3\}$, and we take advantage of conditional independence across subjects when performing the first product.

- Because the Poisson density has no dispersion parameter, step 2 is not performed

- Step 3 targets the density

$$\prod_i N(\boldsymbol{v}_i; \boldsymbol{0}, \tilde{\boldsymbol{C}}_{\boldsymbol{\varphi}}),$$

  where $\boldsymbol{v}_i$ is the vector that collects the spatial factors at all spatial locations for subject i and $\tilde{\boldsymbol{C}}_{\boldsymbol{\varphi}}$ is the covariance matrix induced by the spatial factor model using MGP processes. In practice, this density is efficiently evaluated as a product of small dimensional Gaussian densities, without any need to directly compute $\tilde{\boldsymbol{C}}_{\boldsymbol{\varphi}}$ or its inverse. Refer to Peruzzi and Dunson, 2024, Section 2, for details on MGP models, and Section 4 for additional information on MGPs used in the context of spatial factor and coregionalization models such as those implemented here.

- Step 4 is unchanged, but it is now enclosed in an additional loop that cycles through images and which can be performed in parallel. The updated step 4 thus samples the image-specific latent effects in parallel for all images in the data.

[1] This section was adapted from a manuscript coauthored with Michele Peruzzi and Arvind Rao. The original draft of this section was written primarily by Michele Peruzzi.

B.4 Implementation details

The R package is implemented in C++ using the Armadillo library (Sanderson and Curtin, 2016) via RcppArmadillo (Eddelbuettel and Sanderson, 2014). Parallel operations are performed via OpenMP (Dagum and Menon, 1998). For fast linear algebra, we recommend to link R with efficient BLAS/LAPACK libraries (E. Anderson et al., 1999) such as OpenBLAS (X. Zhang, 2020). All real-data analyses were run on a workstation equipped with a 16-core AMD Ryzen 9 7950X3D CPU, 192GB memory, using Intel Math Kernel Library version 2024.2 (Intel Corporation, 2024) for fast linear algebra. Simulation studies were run on an HPC platform - each simulation was run with 4 cores on a 3.0 GHz Intel Xeon Gold 6154 with 32GB of memory, again using the Intel Math Kernel Library version 2024.2.

APPENDIX C

Supplementary Material for Bayesian Survival Modeling with Functional Spatial Covariates

C.1 Figures

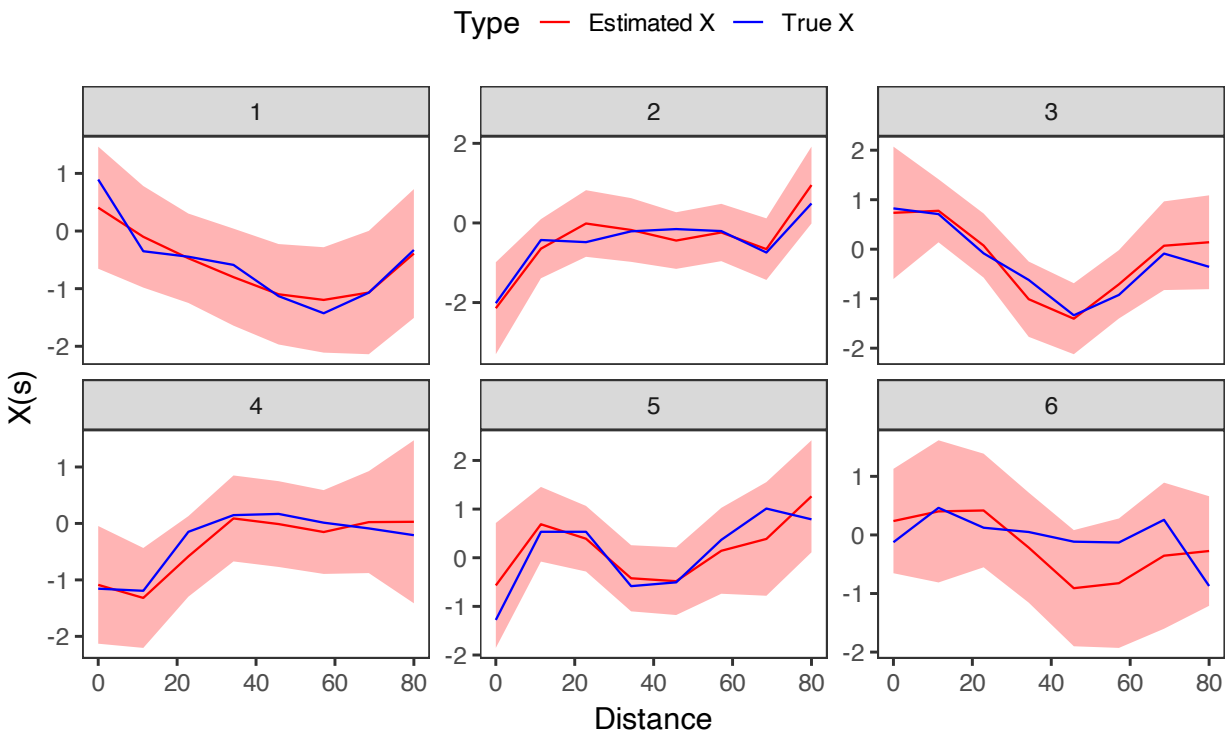


Figure C.1: Examples of posterior estimates of patient-level functional covariates $X_i(s)$ with 95% simultaneous credible bands, for a simulation with $N = 50$ patients and $M = 2$ images per patient. Each facet displays one simulated patient. The true curve $X_i(s)$ (blue) is overlaid with the posterior mean estimate (red) and shaded band denoting the 95% uniform credible interval.

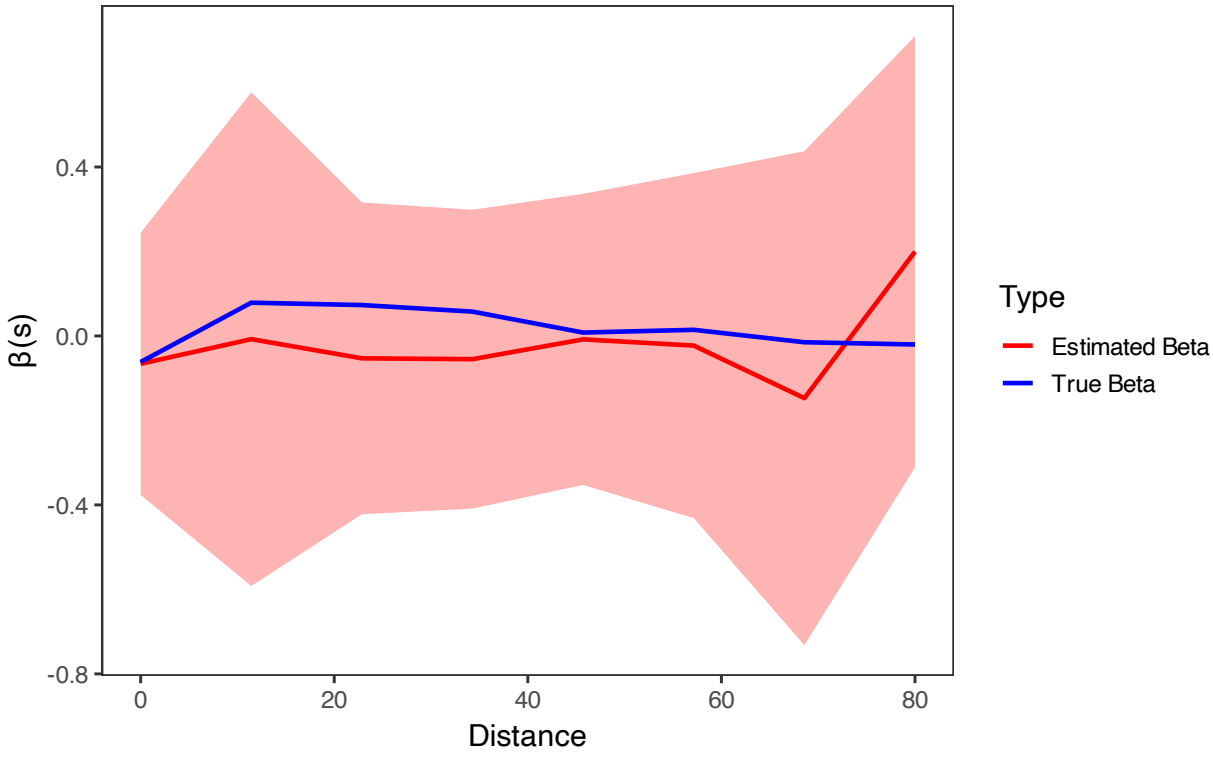


Figure C.2: Posterior estimate of the spatially-varying effect $\beta(s)$ with 95% simultaneous credible band, for a simulation with $N = 50$ patients and $M = 2$ images per patient. The true effect (blue) is overlaid with the posterior mean estimate (red) and a shaded region denoting the 95% uniform credible interval.

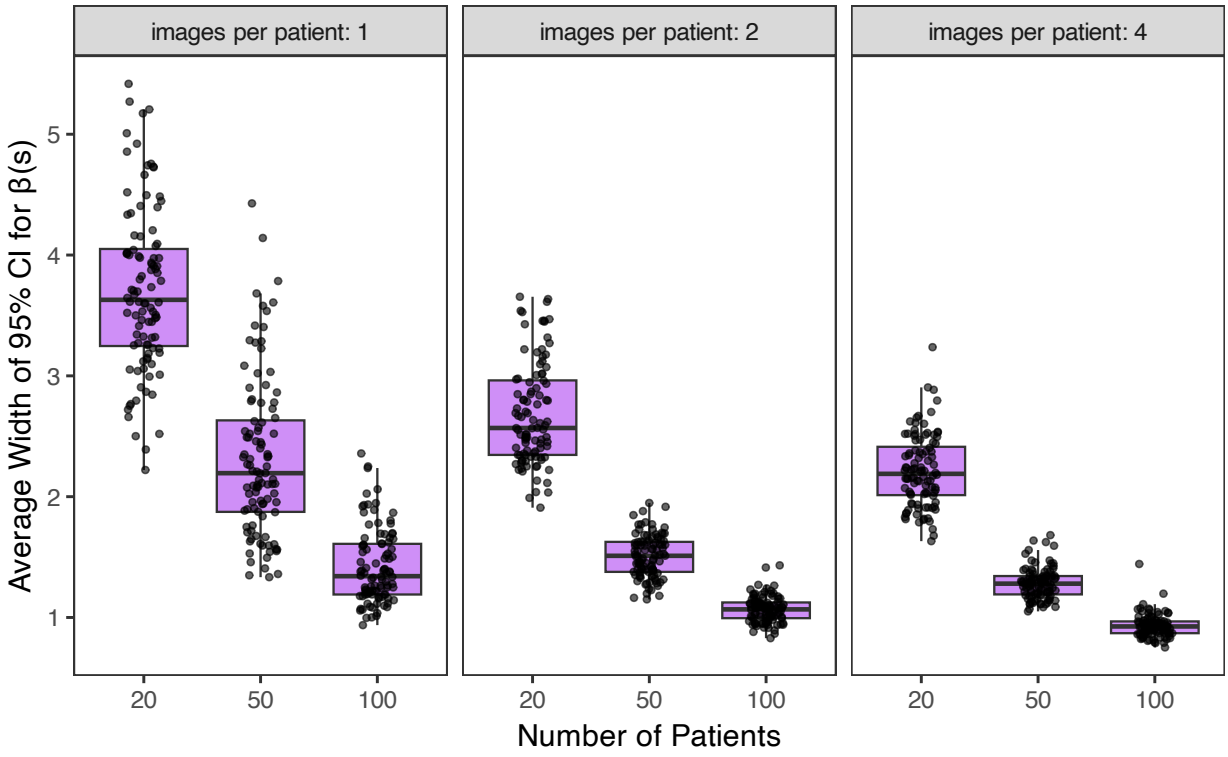


Figure C.3: Average width of the simultaneous 95% credible intervals for the spatially-varying coefficient function $\beta(s)$. Both increased patient count and repeated measurements lead to narrower intervals, indicating improved precision in estimating population-level spatial effects.

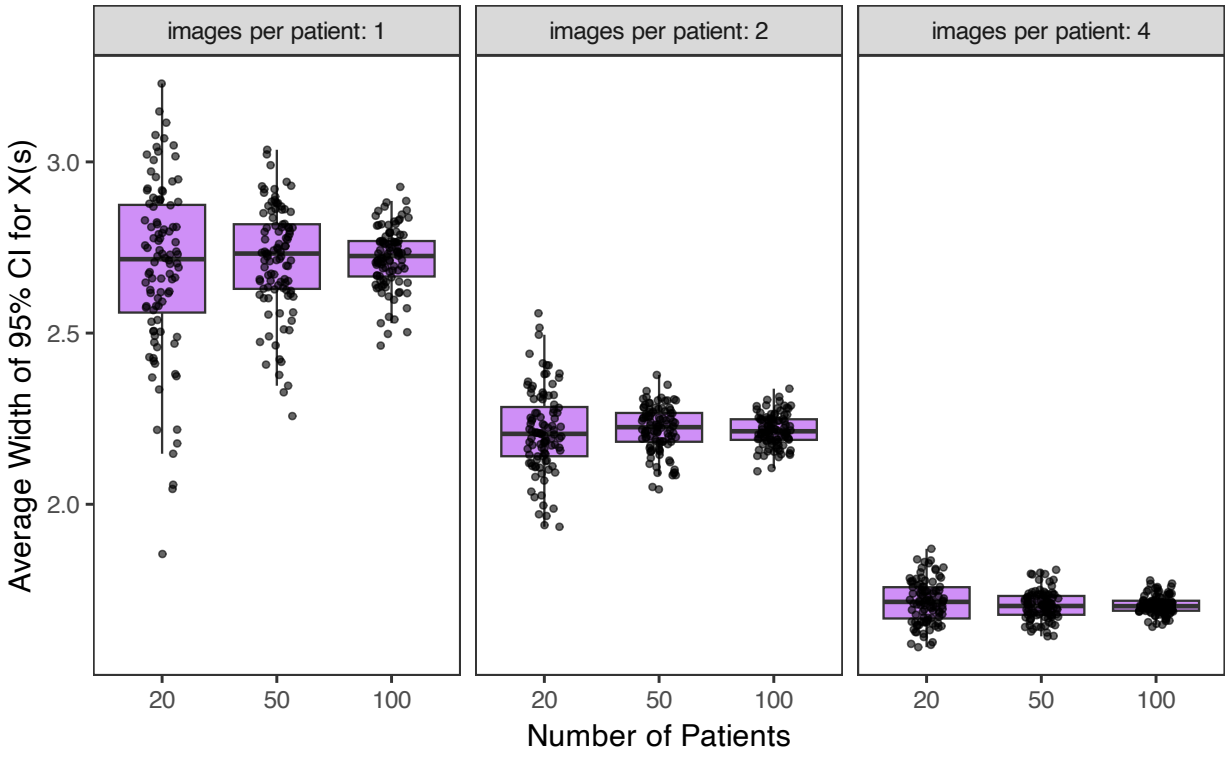


Figure C.4: Average width of the simultaneous 95% credible intervals for patient-specific functional covariates $X_i(s)$, across varying numbers of patients and repeated measurements. Increasing the number of images per patient substantially reduces interval width, reflecting improved uncertainty quantification for individual-level functions. The number of patients has minimal additional effect on this quantity.

C.2 Tables

Table C.1: Prior distributions used in the Bayesian survival model.

Parameter	**Prior Distribution**
Subject-level spline coefficients θ_{X_i}	$\mathcal{N}(0, \sigma^2_{\theta_X} I)$
Spline coefficients for $\beta(s)$, θ_β	$\mathcal{N}(0, \sigma^2_\beta I)$
Weibull shape parameter α	$\text{Gamma}(2, 0.5)$
Log scale parameter for survival $\log \kappa$	$\mathcal{N}(0, 1)$
Standard deviation of θ_{X_i}, σ_{θ_X}	$\mathcal{N}^+(0, \text{scale}^2_x)$
Standard deviation of $\theta_{\beta,k}$, $\sigma_{\beta,k}$	$\mathcal{N}^+(0, \text{scale}^2_\beta)$
Standard deviation of nuisance function, σ_{nuisance}	$\mathcal{N}^+(0, \text{scale}^2_\eta)$

Table C.2: Simultaneous credible interval coverage across dataset sizes. Median 95% simultaneous credible interval coverage across 100 simulations for both the spatial effect $\beta(s)$ and the patient-level covariate $X_i(s)$.

Number of Patients	Number of Images	95% CI Coverage (β)	95% CI Coverage (X)
20	1	1.00	0.90
20	2	1.00	0.95
20	4	1.00	0.95
50	1	1.00	0.92
50	2	1.00	0.94
50	4	1.00	0.94
100	1	1.00	0.93
100	2	1.00	0.95
100	4	1.00	0.94

C.3 Methods

C.3.1 Simultaneous Credible Bands for Functional Coefficients

To quantify uncertainty in the estimated spatially varying coefficient $\beta(s)$, we constructed *simultaneous 95% credible bands* using posterior samples from our Bayesian model. While pointwise credible intervals can quantify uncertainty at individual spatial locations, they do not account for multiple comparisons across the domain and tend to overstate confidence when evaluating the function globally. In contrast, simultaneous credible bands provide a global probabilistic guarantee that the entire function lies within the band over its domain with high posterior probability.

Let $\hat{\beta}(s)$ denote the posterior mean and $\hat{\sigma}(s)$ the posterior standard deviation of $\beta(s)$, evaluated over a fixed grid of spatial locations $s \in \mathcal{S}$. For each posterior draw $\beta^{(j)}(s)$, we computed the standardized residual

$$Z^{(j)}(s) = \frac{\beta^{(j)}(s) - \hat{\beta}(s)}{\hat{\sigma}(s)},$$

and then calculated the *maximum absolute standardized deviation* across the domain:

$$T^{(j)} = \sup_{s \in \mathcal{S}} \left| Z^{(j)}(s) \right|.$$

This statistic captures the largest fluctuation, in units of local posterior uncertainty, that the sampled function exhibits relative to the posterior mean. By repeating this computation over posterior draws, we obtain the empirical distribution of the supremum deviation under the posterior. The critical value c for the simultaneous band is then defined as the $(1-\alpha)$-quantile of the empirical distribution of $T^{(j)}$:

$$c = \text{quantile}_{1-\alpha}\left(\{T^{(1)}, \ldots, T^{(J)}\} \right).$$

The resulting *simultaneous credible band* is:

$$\beta(s) \in \left[\hat{\beta}(s) \pm c \cdot \hat{\sigma}(s) \right], \quad \text{for all } s \in \mathcal{S}.$$

This construction ensures that with posterior probability $1-\alpha$, the entire function $\beta(s)$ remains within the band across the full spatial domain. It provides a conservative yet interpretable summary of global uncertainty, correcting for the multiple comparisons that arise when interpreting the function across many spatial locations.

This approach is conceptually analogous to Scheffé's method for simultaneous inference in classical linear models, but adapted to the function space and implemented through posterior simulation (Scheffé, 1970). The simultaneous credible band is particularly useful for assessing whether the functional effect is consistently positive, negative, or null over specific spatial regions, and for identifying features such as consistent deviations from zero that are unlikely to be due to noise. It also serves as a diagnostic for the informativeness of the data and the adequacy of the smoothing prior: wider bands reflect higher posterior uncertainty, particularly in regions with sparse or noisy data.

BIBLIOGRAPHY

Adler, Frederick R. et al. (Feb. 10, 2022). "Editorial: From Ecology to Cancer Biology and Back Again". In: *Frontiers in Ecology and Evolution* 10, p. 840375. ISSN: 2296-701X. DOI: 10.3389/fevo.2022.840375.

Akoya (Dec. 2021). *inForm Tissue Analysis Software - Akoya*. en-US.

Allam, Mayar et al. (Sept. 1, 2022). "Spatially variant immune infiltration scoring in human cancer tissues". In: *npj Precision Oncology* 6.1. Number: 1 Publisher: Nature Publishing Group, pp. 1–21. ISSN: 2397-768X. DOI: 10.1038/s41698-022-00305-4.

American Cancer Society (2021). *Cancer Facts and Figures 2021*. Tech. rep. Atlanta: American Cancer Society.

— (2022). *Cancer Facts and Figures 2022*. Tech. rep. Atlanta: American Cancer Society.

— (2024). *Cancer Facts and Figures 2024*. Tech. rep. Atlanta: American Cancer Society.

Amin, Mahul B. et al. (Mar. 2017). "The Eighth Edition AJCC Cancer Staging Manual: Continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging". In: *CA: a cancer journal for clinicians* 67.2, pp. 93–99. ISSN: 1542-4863. DOI: 10.3322/caac.21388.

Anderson, E. et al. (1999). *LAPACK Users' Guide*. Third. Philadelphia, PA: Society for Industrial and Applied Mathematics.

Anderson, Nicole M. and M. Celeste Simon (Aug. 2020). "The tumor microenvironment". In: *Current Biology* 30.16, R921–R925. ISSN: 09609822. DOI: 10.1016/j.cub.2020.06.081.

Baddeley, A. J. and M. N. M. van Lieshout (Dec. 1, 1995). "Area-interaction point processes". In: *Annals of the Institute of Statistical Mathematics* 47.4, pp. 601–619. ISSN: 1572-9052. DOI: 10.1007/BF01856536.

Baddeley, Adrian, Jean-François Coeurjolly, et al. (2014). "Logistic regression for spatial Gibbs point processes". In: *Biometrika* 101.2. Publisher: [Oxford University Press, Biometrika Trust], pp. 377–392. ISSN: 0006-3444.

Baddeley, Adrian, Ege Rubak, and Rolf Turner (Nov. 11, 2015). *Spatial Point Patterns: Methodology and Applications with R*. Google-Books-ID: rGbmCgAAQBAJ. CRC Press. 830 pp. ISBN: 978-1-4822-1021-7.

Baddeley, Adrian and Rolf Turner (Jan. 26, 2005). "spatstat: An R Package for Analyzing Spatial Point Patterns". In: *Journal of Statistical Software* 12, pp. 1–42. ISSN: 1548-7660. DOI: 10.18637/jss.v012.i06.

Baker, Alfie T. et al. (Jan. 2021). "Cancer-Associated Fibroblasts and T Cells: From Mechanisms to Outcomes". In: *The Journal of Immunology* 206.2, pp. 310–320. ISSN: 0022-1767. DOI: 10.4049/jimmunol.2001203.

Balogh, Kristen N., Dennis J. Templeton, and Janet V. Cross (2018). "Macrophage Migration Inhibitory Factor protects cancer cells from immunogenic cell death and impairs anti-tumor immune responses". In: *PloS One* 13.6, e0197702. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0197702.

Banerjee, Sudipto, Bradley P. Carlin, and Alan E. Gelfand (Sept. 2014). *Hierarchical Modeling and Analysis for Spatial Data*. 2nd ed. New York: Chapman and Hall/CRC. ISBN: 978-0-429-13717-4. DOI: 10.1201/b17115.

Barua, Souptik et al. (Jan. 2018). "A Functional Spatial Analysis Platform for Discovery of Immunological Interactions Predictive of Low-Grade to High-Grade Transition of Pancreatic Intraductal Papillary Mucinous Neoplasms". In: *Cancer Informatics* 17. Publisher: SAGE Publications Ltd STM, p. 1176935118782880. ISSN: 1176-9351. DOI: 10.1177/1176935118782880.

Behanova, Andrea, Anna Klemm, and Carolina Wählby (Jan. 28, 2022). "Spatial Statistics for Understanding Tissue Organization". In: *Frontiers in Physiology* 13, p. 832417. ISSN: 1664-042X. DOI: 10.3389/fphys.2022.832417.

Ben-Said, Mariem (Aug. 5, 2021). "Spatial point-pattern analysis as a powerful tool in identifying pattern-process relationships in plant ecology: an updated review". In: *Ecological Processes* 10.1, p. 56. ISSN: 2192-1709. DOI: 10.1186/s13717-021-00314-4.

Benimam, Mohamed M. et al. (Mar. 2025). "Statistical analysis of spatial patterns in tumor microenvironment images". en. In: *Nature Communications* 16.1. Publisher: Nature Publishing Group, p. 3090. ISSN: 2041-1723. DOI: 10.1038/s41467-025-57943-y.

Benjamini, Yoav and Yosef Hochberg (1995). "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing". In: *Journal of the Royal Statistical Society. Series B (Methodological)* 57.1, pp. 289–300. ISSN: 0035-9246.

Bindea, Gabriela et al. (Oct. 2013). "Spatiotemporal Dynamics of Intratumoral Immune Cells Reveal the Immune Landscape in Human Cancer". English. In: *Immunity* 39.4. Publisher: Elsevier, pp. 782–795. ISSN: 1074-7613. DOI: 10.1016/j.immuni.2013.10.003.

Binnewies, Mikhail et al. (May 2018). "Understanding the tumor immune microenvironment (TIME) for effective therapy". en. In: *Nature Medicine* 24.5. Publisher: Nature Publishing Group, pp. 541–550. ISSN: 1546-170X. DOI: 10.1038/s41591-018-0014-x.

Black, Sarah et al. (Aug. 2021). "CODEX multiplexed tissue imaging with DNA-conjugated antibodies". In: *Nature Protocols* 16.8, pp. 3802–3835. ISSN: 1750-2799. DOI: 10.1038/s41596-021-00556-8.

Bodenmiller, Bernd (Apr. 2016). "Multiplexed Epitope-Based Tissue Imaging for Discovery and Healthcare Applications". English. In: *Cell Systems* 2.4. Publisher: Elsevier, pp. 225–238. ISSN: 2405-4712, 2405-4720. DOI: 10.1016/j.cels.2016.03.008.

Boedtkjer, Ebbe and Stine F. Pedersen (Feb. 10, 2020). "The Acidic Tumor Microenvironment as a Driver of Cancer". In: *Annual Review of Physiology* 82, pp. 103–126. ISSN: 1545-1585. DOI: 10.1146/annurev-physiol-021119-034627.

Bożyk, Aleksandra et al. (June 2022). "Tumor Microenvironment—A Short Review of Cellular and Interaction Diversity". In: *Biology* 11.6. Number: 6 Publisher: Multidisciplinary Digital Publishing Institute, p. 929. ISSN: 2079-7737. DOI: 10.3390/biology11060929.

Bull, Joshua A. et al. (Oct. 2020). "Combining multiple spatial statistics enhances the description of immune cell localisation within tumours". en. In: *Scientific Re-*

ports 10.1. Publisher: Nature Publishing Group, p. 18624. ISSN: 2045-2322. DOI: 10.1038/s41598-020-75180-9.

Canete, Nicolas P et al. (May 26, 2022). "spicyR: spatial analysis of in situ cytometry data in R". In: *Bioinformatics* 38.11, pp. 3099–3105. ISSN: 1367-4803. DOI: 10.1093/bioinformatics/btac268.

Cao, Zhipeng et al. (Jan. 2025). "Cancer-associated fibroblasts as therapeutic targets for cancer: advances, challenges, and future prospects". In: *Journal of Biomedical Science* 32.1, p. 7. ISSN: 1423-0127. DOI: 10.1186/s12929-024-01099-2.

Carpenter, Bob et al. (2017). "Stan: A probabilistic programming language". In: *Journal of statistical software* 76.1.

Carstens, Julienne L., Pedro Correa de Sampaio, et al. (Apr. 2017). "Spatial computation of intratumoral T cells correlates with survival of patients with pancreatic cancer". In: *Nature Communications* 8.1. Publisher: Nature Publishing Group, p. 15095. ISSN: 2041-1723. DOI: 10.1038/ncomms15095.

Carstens, Julienne L., Santhoshi N. Krishnan, et al. (Aug. 2024). "Spatial multiplexing and omics". en. In: *Nature Reviews Methods Primers* 4.1. Publisher: Nature Publishing Group, pp. 1–19. ISSN: 2662-8449. DOI: 10.1038/s43586-024-00330-6.

Chang, Ching-Yuan et al. (Mar. 2020). "Immune Checkpoint Inhibitors and Immune-Related Adverse Events in Patients With Advanced Melanoma: A Systematic Review and Network Meta-analysis". In: *JAMA Network Open* 3.3, e201611. ISSN: 2574-3805. DOI: 10.1001/jamanetworkopen.2020.1611.

Chen, Minjie et al. (Aug. 2024). "CAFs and T cells interplay: The emergence of a new arena in cancer combat". In: *Biomedicine & Pharmacotherapy* 177, p. 117045. ISSN: 0753-3322. DOI: 10.1016/j.biopha.2024.117045.

Chow, Andrew et al. (Dec. 2022). "Clinical implications of T cell exhaustion for cancer immunotherapy". In: *Nature Reviews Clinical Oncology* 19.12. Publisher: Nature Publishing Group, pp. 775–790. ISSN: 1759-4782. DOI: 10.1038/s41571-022-00689-z.

Clark, James S. et al. (2014). "More than the sum of the parts: forest climate response from joint species distribution models". In: *Ecological Applications* 24.5.

_eprint: https://esajournals.onlinelibrary.wiley.com/doi/pdf/10.1890/13-1015.1, pp. 990–999. ISSN: 1939-5582. DOI: 10.1890/13-1015.1.

Cranmer, Kyle, Johann Brehmer, and Gilles Louppe (Dec. 2020). "The frontier of simulation-based inference". In: *Proceedings of the National Academy of Sciences* 117.48. Publisher: Proceedings of the National Academy of Sciences, pp. 30055–30062. DOI: 10.1073/pnas.1912789117.

Creed, Jordan H et al. (Dec. 7, 2021). "spatialTIME and iTIME: R package and Shiny application for visualization and analysis of immunofluorescence data". In: *Bioinformatics* 37.23, pp. 4584–4586. ISSN: 1367-4803. DOI: 10.1093/bioinformatics/btab757.

Cribari-Neto, Francisco and Achim Zeileis (Apr. 2010). "Beta Regression in R". In: *Journal of Statistical Software* 34, pp. 1–24. ISSN: 1548-7660. DOI: 10.18637/jss.v034.i02.

Dagum, Leonardo and Ramesh Menon (1998). "OpenMP: an industry standard API for shared-memory programming". In: *Computational Science & Engineering, IEEE* 5.1, pp. 46–55.

Dam, Stephanie van, Matthijs J. D. Baars, and Yvonne Vercoulen (Jan. 2022). "Multiplex Tissue Imaging: Spatial Revelations in the Tumor Microenvironment". In: *Cancers* 14.13. Number: 13 Publisher: Multidisciplinary Digital Publishing Institute, p. 3170. ISSN: 2072-6694. DOI: 10.3390/cancers14133170.

Dang, Hao et al. (Apr. 2023). "Cancer-Associated Fibroblasts Are Key Determinants of Cancer Cell Invasion in the Earliest Stage of Colorectal Cancer". In: *Cellular and Molecular Gastroenterology and Hepatology* 16.1, pp. 107–131. ISSN: 2352-345X. DOI: 10.1016/j.jcmgh.2023.04.004.

Datta, Abhirup et al. (2016). "Hierarchical Nearest-Neighbor Gaussian Process Models for Large Geostatistical Datasets". eng. In: *Journal of the American Statistical Association* 111.514, pp. 800–812. ISSN: 0162-1459. DOI: 10.1080/01621459.2015.1044091.

De Vito, Roberta et al. (Mar. 2019). "Multi-study factor analysis". eng. In: *Biometrics* 75.1, pp. 337–346. ISSN: 1541-0420. DOI: 10.1111/biom.12974.

DeNardo, David G. and Brian Ruffell (June 2019). "Macrophages as regulators of tumor immunity and immunotherapy". In: *Nature reviews. Immunology* 19.6, pp. 369–382. ISSN: 1474-1733. DOI: 10.1038/s41577-019-0127-6.

Diggle, Peter J. et al. (Nov. 2013). "Spatial and Spatio-Temporal Log-Gaussian Cox Processes: Extending the Geostatistical Paradigm". In: *Statistical Science* 28.4. arXiv:1312.6536 [stat]. ISSN: 0883-4237. DOI: 10.1214/13-STS441.

Ding, Zhen et al. (Feb. 2021). "iNOS Associates With Poor Survival in Melanoma: A Role for Nitric Oxide in the PI3K-AKT Pathway Stimulation and PTEN S-Nitrosylation". In: *Frontiers in Oncology* 11. Publisher: Frontiers. ISSN: 2234-943X. DOI: 10.3389/fonc.2021.631766.

Dolina, Joseph S. et al. (2021). "CD8+ T Cell Exhaustion in Cancer". eng. In: *Frontiers in Immunology* 12, p. 715234. ISSN: 1664-3224. DOI: 10.3389/fimmu.2021.715234.

Dummer, Reinhard et al. (Dec. 2020). "Rationale for Immune Checkpoint Inhibitors Plus Targeted Therapy in Metastatic Melanoma: A Review". In: *JAMA oncology* 6.12, pp. 1957–1966. ISSN: 2374-2445. DOI: 10.1001/jamaoncol.2020.4401.

Eddelbuettel, Dirk and Conrad Sanderson (Mar. 2014). "RcppArmadillo: Accelerating R with high-performance C++ linear algebra". In: *Computational Statistics and Data Analysis* 71, pp. 1054–1063.

Ekmekcioglu, Suhendan, Michael A. Davies, et al. (June 2016). "Inflammatory Marker Testing Identifies CD74 Expression in Melanoma Tumor Cells, and Its Expression Associates with Favorable Survival for Stage III Melanoma". In: *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research* 22.12, pp. 3016–3024. ISSN: 1557-3265. DOI: 10.1158/1078-0432.CCR-15-2226.

Ekmekcioglu, Suhendan, Julie A. Ellerhorst, et al. (Aug. 2006). "Tumor iNOS predicts poor survival for stage III melanoma patients". In: *International Journal of Cancer* 119.4, pp. 861–866. ISSN: 0020-7136. DOI: 10.1002/ijc.21767.

Eliason, Joel, Santhoshi Krishnan, et al. (Apr. 2025). "Characterizing spatial immune architecture in metastatic melanoma using high-dimensional multiplex imaging". English. In: *Frontiers in Immunology* 16. Publisher: Frontiers, p. 1560778. ISSN: 1664-3224. DOI: 10.3389/fimmu.2025.1560778.

Eliason, Joel and Arvind Rao (June 2024). "Investigating Ecological Interactions in the Tumor Microenvironment Using Joint Species Distribution Models for Point Patterns". en. In: *The New England Journal of Statistics in Data Science* 2.3. Publisher: New England Statistical Society, pp. 296–310. ISSN: 2693-7166. DOI: 10.51387/24-NEJSDS66.

Fedele, Vita and Davide Melisi (Apr. 2020). "Permissive State of EMT: The Role of Immune Cell Compartment". In: *Frontiers in Oncology* 10, p. 587. ISSN: 2234-943X. DOI: 10.3389/fonc.2020.00587.

Fife, Christopher et al. (Nov. 2024). "Natural killer cells are required for the recruitment of CD8+ T cells and the efficacy of immune checkpoint blockade in melanoma brain metastases". In: *Journal for ImmunoTherapy of Cancer* 12.11. Publisher: BMJ Specialist Journals Section: Clinical/translational cancer immunotherapy, e009522. ISSN: 2051-1426. DOI: 10.1136/jitc-2024-009522.

Flint, Ian et al. (Nov. 1, 2022). "The Saturated Pairwise Interaction Gibbs Point Process as a Joint Species Distribution Model". In: *Journal of the Royal Statistical Society Series C: Applied Statistics* 71.5, pp. 1721–1752. ISSN: 0035-9254. DOI: 10.1111/rssc.12596.

Fortin, Marie-Josée, Mark R.T. Dale, and Jay M. Ver Hoef (2016). "Spatial Analysis in Ecology". In: *Wiley StatsRef: Statistics Reference Online*. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/9781118445112.stat07766.pub2. John Wiley & Sons, Ltd, pp. 1–13. ISBN: 978-1-118-44511-2. DOI: 10.1002/9781118445112.stat07766.pub2.

Freeman, Patrick and Ainhoa Mielgo (Sept. 2020). "Cancer-Associated Fibroblast Mediated Inhibition of CD8+ Cytotoxic T Cell Accumulation in Tumours: Mechanisms and Therapeutic Opportunities". In: *Cancers* 12.9, p. 2687. ISSN: 2072-6694. DOI: 10.3390/cancers12092687.

Fu, Tong et al. (June 2021). "Spatial architecture of the immune microenvironment orchestrates tumor immunity and therapeutic response". en. In: *Journal of Hematology & Oncology* 14.1, p. 98. ISSN: 1756-8722. DOI: 10.1186/s13045-021-01103-4.

Fukuda, Yasunari et al. (July 2023). "Inhibition of Microsomal Prostaglandin E2 Synthase Reduces Collagen Deposition in Melanoma Tumors and May Improve Immunotherapy

Efficacy by Reducing T-cell Exhaustion". In: *Cancer Research Communications* 3.7, pp. 1397–1408. ISSN: 2767-9764. DOI: 10.1158/2767-9764.CRC-23-0210.

Gabry, Jonah et al. (2024). *cmdstanr: R Interface to 'CmdStan'*. R package version 0.8.1, https://discourse.mc-stan.org.

Galon, Jérôme et al. (Oct. 3, 2012). "Cancer classification using the Immunoscore: a worldwide task force". In: *Journal of Translational Medicine* 10.1, p. 205. ISSN: 1479-5876. DOI: 10.1186/1479-5876-10-205.

Genton, Marc G. and William Kleiber (2015). "Cross-Covariance Functions for Multivariate Geostatistics". In: *Statistical Science* 30.2, pp. 147–163.

Geyer, C. (1999). "Likelihood inference for spatial point processes". In: *Stochastic Geometry: Likelihood and Computation, chapter 3.* Ed. by O.E. Barndorff-Nielsen, W.S. Kendall, and M.N.M. Van Lieshout. Monographs on Statistics and Applied Probability 80. Chapman and Hall/CRC, pp. 79–140. ISBN: 978-0-203-73827-6.

Geyer, Charles J. (2011). "Introduction to Markov Chain Monte Carlo". In: *Handbook of Markov Chain Monte Carlo*. Num Pages: 46. Chapman and Hall/CRC. ISBN: 978-0-429-13850-8.

Giesen, Charlotte et al. (Apr. 2014). "Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry". en. In: *Nature Methods* 11.4. Publisher: Nature Publishing Group, pp. 417–422. ISSN: 1548-7105. DOI: 10.1038/nmeth.2869.

Gilkes, Daniele M., Gregg L. Semenza, and Denis Wirtz (June 2014). "Hypoxia and the extracellular matrix: drivers of tumour metastasis". en. In: *Nature Reviews Cancer* 14.6. Publisher: Nature Publishing Group, pp. 430–439. ISSN: 1474-1768. DOI: 10.1038/nrc3726.

Goltsev, Yury et al. (Aug. 9, 2018). "Deep Profiling of Mouse Splenic Architecture with CODEX Multiplexed Imaging". In: *Cell* 174.4, 968–981.e15. ISSN: 0092-8674. DOI: 10.1016/j.cell.2018.07.010.

Gong, Chang et al. (July 2021). "A Spatial Quantitative Systems Pharmacology Platform spQSP-IO for Simulations of Tumor–Immune Interactions and Effects of Checkpoint Inhibitor Immunotherapy". In: *Cancers* 13.15, p. 3751. ISSN: 2072-6694. DOI: 10.3390/cancers13153751.

Goulart, Michelle R et al. (Dec. 2021). "T cells in pancreatic cancer stroma". In: *World Journal of Gastroenterology* 27.46, pp. 7956–7968. ISSN: 1007-9327. DOI: 10.3748/wjg.v27.i46.7956.

Grabarnik, Pavel and Aila Särkkä (May 17, 2009). "Modelling the spatial structure of forest stands by multivariate point processes with hierarchical interactions". In: *Ecological Modelling* 220.9, pp. 1232–1240. ISSN: 0304-3800. DOI: 10.1016/j.ecolmodel.2009.02.021.

Grimm, Elizabeth A., Andrew G. Sikora, and Suhendan Ekmekcioglu (Oct. 2013). "Molecular pathways: inflammation-associated nitric-oxide production as a cancer-supporting redox mechanism and a potential therapeutic target". In: *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research* 19.20, pp. 5557–5563. ISSN: 1557-3265. DOI: 10.1158/1078-0432.CCR-12-1554.

Gunaydin, Gurcan (July 2021). "CAFs Interacting With TAMs in Tumor Microenvironment to Enhance Tumorigenesis and Immune Evasion". In: *Frontiers in Oncology* 11, p. 668349. ISSN: 2234-943X. DOI: 10.3389/fonc.2021.668349.

Guo, Shuaiqingying et al. (Sept. 2023). "Cancer-associated fibroblasts: Just on the opposite side of antitumour immunity?" In: *International Immunopharmacology* 122, p. 110601. ISSN: 1567-5769. DOI: 10.1016/j.intimp.2023.110601.

Hailemichael, Yared et al. (May 2022). "Interleukin-6 blockade abrogates immunotherapy toxicity and promotes tumor immunity". In: *Cancer Cell* 40.5, 509–523.e6. ISSN: 1878-3686. DOI: 10.1016/j.ccell.2022.04.004.

Hauck, Walter W. and Allan Donner (1977). "Wald's Test as Applied to Hypotheses in Logit Analysis". In: *Journal of the American Statistical Association* 72.360. Publisher: [American Statistical Association, Taylor & Francis, Ltd.], pp. 851–853. ISSN: 0162-1459. DOI: 10.2307/2286473.

Hendry, Shona A. et al. (Dec. 2016). "The Role of the Tumor Vasculature in the Host Immune Response: Implications for Therapeutic Strategies Targeting the Tumor Microenvironment". In: *Frontiers in Immunology* 7, p. 621. ISSN: 1664-3224. DOI: 10.3389/fimmu.2016.00621.

Hernandez, Sharia et al. (May 2022). "Diminished Immune Surveillance during Histologic Progression of Intraductal Papillary Mucinous Neoplasms Offers a Therapeutic Oppor-

tunity for Cancer Interception". In: *Clinical Cancer Research* 28.9, pp. 1938–1947. ISSN: 1078-0432. DOI: 10.1158/1078-0432.CCR-21-2585.

HöUgmander, Harri and Aila SäUrkkä (1999). "Multitype Spatial Point Patterns with Hierarchical Interactions". In: *Biometrics* 55.4. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.0006-341X.1999.01051.x, pp. 1051–1058. ISSN: 1541-0420. DOI: 10.1111/j.0006-341X.1999.01051.x.

Hu, Jian et al. (Nov. 2021). "SpaGCN: Integrating gene expression, spatial location and histology to identify spatial domains and spatially variable genes by graph convolutional network". en. In: *Nature Methods* 18.11. Publisher: Nature Publishing Group, pp. 1342–1351. ISSN: 1548-7105. DOI: 10.1038/s41592-021-01255-8.

Huang, Alexander C. and Roberta Zappasodi (May 2022). "A decade of checkpoint blockade immunotherapy in melanoma: understanding the molecular basis for immune sensitivity and resistance". In: *Nature Immunology* 23.5. Publisher: Nature Publishing Group, pp. 660–670. ISSN: 1529-2916. DOI: 10.1038/s41590-022-01141-1.

Illian, Janine B. and Ditte K. Hendrichsen (2010). "Gibbs point process models with mixed effects". In: *Environmetrics* 21.3. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/env.1008, pp. 341–353. ISSN: 1099-095X. DOI: 10.1002/env.1008.

Illian, Janine B., Sara Martino, et al. (2013). "Fitting complex ecological point process models with integrated nested Laplace approximation". In: *Methods in Ecology and Evolution* 4.4. _eprint: https://besjournals.onlinelibrary.wiley.com/doi/pdf/10.1111/2041-210x.12017, pp. 305–315. ISSN: 2041-210X. DOI: 10.1111/2041-210x.12017.

Illian, Janine B., Jesper Møller, and Rasmus P. Waagepetersen (Sept. 1, 2009). "Hierarchical spatial point process analysis for a plant community with high biodiversity". In: *Environmental and Ecological Statistics* 16.3, pp. 389–405. ISSN: 1573-3009. DOI: 10.1007/s10651-007-0070-8.

Intel Corporation (2024). *Intel Math Kernel Library*. Version 2024-2, Software available at https://software.intel.com/content/www/us/en/develop/tools/math-kernel-library.html.

Jackson, Hartland W. et al. (Feb. 2020). "The single-cell pathology landscape of breast cancer". en. In: *Nature* 578.7796. Publisher: Nature Publishing Group, pp. 615–620. ISSN: 1476-4687. DOI: 10.1038/s41586-019-1876-x.

Jamouss, Kevin T. et al. (Feb. 2025). "Tumor immune microenvironment alterations associated with progression in human intraductal papillary mucinous neoplasms". eng. In: *The Journal of Pathology*. ISSN: 1096-9896. DOI: 10.1002/path.6402.

Jenkins, Liam et al. (Aug. 2022). "Cancer-Associated Fibroblasts Suppress CD8+ T-cell Infiltration and Confer Resistance to Immune-Checkpoint Blockade". In: *Cancer Research* 82.16, pp. 2904–2917. ISSN: 0008-5472. DOI: 10.1158/0008-5472.CAN-21-4141.

Jenkins, Russell W., David A. Barbie, and Keith T. Flaherty (Jan. 2018). "Mechanisms of resistance to immune checkpoint inhibitors". In: *British Journal of Cancer* 118.1. Publisher: Nature Publishing Group, pp. 9–16. ISSN: 1532-1827. DOI: 10.1038/bjc.2017.434.

Karasarides, Maria et al. (Apr. 2022). "Hallmarks of Resistance to Immune-Checkpoint Inhibitors". In: *Cancer Immunology Research* 10.4, pp. 372–383. ISSN: 2326-6074. DOI: 10.1158/2326-6066.CIR-20-0586.

Kassambara, Alboukadel et al. (Oct. 2024). *survminer: Drawing Survival Curves using 'ggplot2'*.

Katzfuss, Matthias and Joseph Guinness (Feb. 2021). "A General Framework for Vecchia Approximations of Gaussian Processes". In: *Statistical Science* 36.1. Publisher: Institute of Mathematical Statistics, pp. 124–141. ISSN: 0883-4237, 2168-8745. DOI: 10.1214/19-STS755.

King, Ruth et al. (Dec. 1, 2012). "A Bayesian Approach to Fitting Gibbs Processes with Temporal Random Effects". In: *Journal of Agricultural, Biological, and Environmental Statistics* 17.4, pp. 601–622. ISSN: 1537-2693. DOI: 10.1007/s13253-012-0111-0.

Kinker, Gabriela Sarti et al. (June 2021). "B Cell Orchestration of Anti-tumor Immune Responses: A Matter of Cell Localization and Communication". In: *Frontiers in Cell and Developmental Biology* 9, p. 678127. ISSN: 2296-634X. DOI: 10.3389/fcell.2021.678127.

Kitamura, Takanori (Sept. 2023). "Tumour-associated macrophages as a potential target to improve natural killer cell-based immunotherapies". In: *Essays in Biochemistry* 67.6, pp. 1003–1014. ISSN: 0071-1365. DOI: 10.1042/EBC20230002.

Knight, Andrew D. and Jason J. Luke (July 2024). "Beyond Immune Checkpoint Inhibitors: Emerging Targets in Melanoma Therapy". In: *Current Oncology Reports* 26.7, pp. 826–839. ISSN: 1534-6269. DOI: 10.1007/s11912-024-01551-4.

Krishnan, Santhoshi N., Sunjong Ji, et al. (Nov. 2024). "Proximogram—A multi-omics network-based framework to capture tissue heterogeneity integrating single-cell omics and spatial profiling". In: *Computers in Biology and Medicine* 182, p. 109082. ISSN: 0010-4825. DOI: 10.1016/j.compbiomed.2024.109082.

Krishnan, Santhoshi N., Shariq Mohammed, et al. (Mar. 2022). "GaWRDenMap: a quantitative framework to study the local variation in cell–cell interactions in pancreatic disease subtypes". en. In: *Scientific Reports* 12.1. Publisher: Nature Publishing Group, p. 3708. ISSN: 2045-2322. DOI: 10.1038/s41598-022-06602-z.

Kuburich, Nick A. et al. (Dec. 2024). "The Cancer Chimera: Impact of Vimentin and Cytokeratin Co-Expression in Hybrid Epithelial/Mesenchymal Cancer Cells on Tumor Plasticity and Metastasis". In: *Cancers* 16.24, p. 4158. ISSN: 2072-6694. DOI: 10.3390/cancers16244158.

Lanitis, Evripidis, Melita Irving, and George Coukos (Apr. 2015). "Targeting the tumor vasculature to enhance T cell activity". In: *Current Opinion in Immunology* 33, pp. 55–63. ISSN: 0952-7915. DOI: 10.1016/j.coi.2015.01.011.

Law, Richard et al. (2009). "Ecological information from spatial patterns of plants: insights from point process theory". In: *Journal of Ecology* 97.4. _eprint: https://besjournals.onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-2745.2009.01510.x, pp. 616–628. ISSN: 1365-2745. DOI: 10.1111/j.1365-2745.2009.01510.x.

Lazarus, Jenny et al. (Nov. 2018). "Spatial and phenotypic immune profiling of metastatic colon cancer". eng. In: *JCI insight* 3.22, e121932, 121932. ISSN: 2379-3708. DOI: 10.1172/jci.insight.121932.

Lee, Hansol et al. (Mar. 2021). "Targeting NK Cells to Enhance Melanoma Response to Immunotherapies". In: *Cancers* 13.6, p. 1363. ISSN: 2072-6694. DOI: 10.3390/cancers13061363.

Li, Xiaoying et al. (Oct. 7, 2021). "Crosstalk Between the Tumor Microenvironment and Cancer Cells: A Promising Predictive Biomarker for Immune Checkpoint Inhibitors". In: *Frontiers in Cell and Developmental Biology* 9, p. 738373. ISSN: 2296-634X. DOI: 10.3389/fcell.2021.738373.

Li, Yue, Long Zhao, and Xiao-Feng Li (Aug. 5, 2021). "Hypoxia and the Tumor Microenvironment". In: *Technology in Cancer Research & Treatment* 20, p. 15330338211036304. ISSN: 1533-0346. DOI: 10.1177/15330338211036304.

Lim, Su Yin et al. (Mar. 2023). "The molecular and functional landscape of resistance to immune checkpoint blockade in melanoma". In: *Nature Communications* 14.1. Publisher: Nature Publishing Group, p. 1516. ISSN: 2041-1723. DOI: 10.1038/s41467-023-36979-y.

Lipman, Yaron et al. (Sept. 2022). "Flow Matching for Generative Modeling". en. In.

Lock, Eric F., Jun Young Park, and Katherine A. Hoadley (Mar. 2022). "Bidimensional linked matrix factorization for pan-omics pan-cancer analysis". In: *The Annals of Applied Statistics* 16.1. Publisher: Institute of Mathematical Statistics, pp. 193–215. ISSN: 1932-6157, 1941-7330. DOI: 10.1214/21-AOAS1495.

Maley, Carlo C. et al. (Oct. 2017). "Classifying the evolutionary and ecological features of neoplasms". en. In: *Nature Reviews Cancer* 17.10. Publisher: Nature Publishing Group, pp. 605–619. ISSN: 1474-1768. DOI: 10.1038/nrc.2017.69.

Maoz, Asaf, Michael Dennis, and Joel K. Greenson (Aug. 2019). "The Crohn's-Like Lymphoid Reaction to Colorectal Cancer-Tertiary Lymphoid Structures With Immunologic and Potentially Therapeutic Relevance in Colorectal Cancer". In: *Frontiers in Immunology* 10, p. 1884. ISSN: 1664-3224. DOI: 10.3389/fimmu.2019.01884.

Mauri, Lorenzo, Niccolò Anceschi, and David B. Dunson (Feb. 2025). *Spectral decomposition-assisted multi-study factor analysis*. arXiv:2502.14600 [stat]. DOI: 10.48550/arXiv.2502.14600.

McKenzie, Brienne and Salvatore Valitutti (Mar. 2023). "Resisting T cell attack: tumor-cell-intrinsic defense and reparation mechanisms". In: *Trends in Cancer* 9.3, pp. 198–211. ISSN: 2405-8025. DOI: 10.1016/j.trecan.2022.12.003.

McKinley, Eliot T et al. (2022). "MIRIAM: A machine and deep learning single-cell segmentation and quantification pipeline for multi-dimensional tissue images". In: *Cytometry Part A* 101.6, pp. 521–528. DOI: 10.1002/cyto.a.24541.

Mhaidly, Rana and Fatima Mechta-Grigoriou (July 2021). "Role of cancer-associated fibroblast subpopulations in immune infiltration, as a new means of treatment in cancer". In: *Immunological Reviews* 302.1, pp. 259–272. ISSN: 0105-2896. DOI: 10.1111/imr.12978.

Milosevic, Vladan and Arne Östman (May 2024). "Interactions between cancer-associated fibroblasts and T-cells: functional crosstalk with targeting and biomarker potential". In: *Upsala Journal of Medical Sciences* 129, 10.48101/ujms.v126.10710. ISSN: 0300-9734. DOI: 10.48101/ujms.v129.10710.

Moller, Jesper and Rasmus Plenge Waagepetersen (Sept. 2003). *Statistical Inference and Simulation for Spatial Point Processes*. New York: Chapman and Hall/CRC. ISBN: 978-0-203-49693-0. DOI: 10.1201/9780203496930.

Mota Reyes, Carmen et al. (Sept. 2022). "Regulatory T Cells in Pancreatic Cancer: Of Mice and Men". In: *Cancers* 14.19, p. 4582. ISSN: 2072-6694. DOI: 10.3390/cancers14194582.

Neal, Radford M. (June 2003). "Slice sampling". In: *The Annals of Statistics* 31.3. Publisher: Institute of Mathematical Statistics, pp. 705–767. ISSN: 0090-5364, 2168-8966. DOI: 10.1214/aos/1056562461.

Ni, Yanghong et al. (May 20, 2021). "The Role of Tumor-Stroma Interactions in Drug Resistance Within Tumor Microenvironment". In: *Frontiers in Cell and Developmental Biology* 9, p. 637675. ISSN: 2296-634X. DOI: 10.3389/fcell.2021.637675.

Noe, Jordan T. and Robert A. Mitchell (Nov. 2020). "MIF-Dependent Control of Tumor Immunity". In: *Frontiers in Immunology* 11. Publisher: Frontiers. ISSN: 1664-3224. DOI: 10.3389/fimmu.2020.609948.

Nowosad, Ada, Jean-Christophe Marine, and Panagiotis Karras (Nov. 2023). "Perivascular niches: critical hubs in cancer evolution". In: *Trends in Cancer* 9.11, pp. 897–910. ISSN: 2405-8033. DOI: 10.1016/j.trecan.2023.06.010.

Noy, Roy and Jeffrey W. Pollard (July 2014). "Tumor-Associated Macrophages: From Mechanisms to Therapy". In: *Immunity* 41.1, pp. 49–61. ISSN: 10747613. DOI: 10.1016/j.immuni.2014.06.010.

Ogata, Dai et al. (Dec. 2020). "The Expression of CD74-Regulated Inflammatory Markers in Stage IV Melanoma: Risk of CNS Metastasis and Patient Survival". In: *Cancers* 12.12, p. 3754. ISSN: 2072-6694. DOI: 10.3390/cancers12123754.

Ovaskainen, Otso et al. (2016). "Uncovering hidden spatial structure in species communities with spatially explicit joint species distribution models". In: *Methods in Ecology and Evolution* 7.4. _eprint: https://besjournals.onlinelibrary.wiley.com/doi/pdf/10.1111/2041-210X.12502, pp. 428–436. ISSN: 2041-210X. DOI: 10.1111/2041-210X.12502.

Palla, Giovanni et al. (Apr. 2021). *Squidpy: a scalable framework for spatial single cell analysis*. en. Tech. rep. Company: Cold Spring Harbor Laboratory Distributor: Cold Spring Harbor Laboratory Label: Cold Spring Harbor Laboratory Section: New Results, p. 2021.02.19.431994. DOI: 10.1101/2021.02.19.431994.

Peruzzi, Michele (Dec. 2024). *Inside-out cross-covariance for spatial multivariate data*. arXiv:2412.12407 [stat]. DOI: 10.48550/arXiv.2412.12407.

Peruzzi, Michele, Sudipto Banerjee, and Andrew O. Finley (Apr. 2022). "Highly Scalable Bayesian Geostatistical Modeling via Meshed Gaussian Processes on Partitioned Domains". In: *Journal of the American Statistical Association* 117.538. Publisher: ASA Website _eprint: https://doi.org/10.1080/01621459.2020.1833889, pp. 969–982. ISSN: 0162-1459. DOI: 10.1080/01621459.2020.1833889.

Peruzzi, Michele and David B. Dunson (2024). "Spatial meshing for general Bayesian multivariate models". In: *Journal of Machine Learning Research* 25.87, pp. 1–49. ISSN: 1533-7928.

Petrova, Varvara et al. (Jan. 24, 2018). "The hypoxic tumour microenvironment". In: *Oncogenesis* 7.1. Number: 1 Publisher: Nature Publishing Group, pp. 1–13. ISSN: 2157-9024. DOI: 10.1038/s41389-017-0011-9.

Piatakova, Adrianna et al. (June 2021). "Distinct Responsiveness of Tumor-Associated Macrophages to Immunotherapy of Tumors with Different Mechanisms of Major Histocompatibility Complex Class I Downregulation". In: *Cancers* 13.12, p. 3057. ISSN: 2072-6694. DOI: 10.3390/cancers13123057.

Pienta, Kenneth J et al. (Dec. 2008). "Ecological Therapy for Cancer: Defining Tumors Using an Ecosystem Paradigm Suggests New Opportunities for Novel Cancer Treatments". In: *Translational Oncology* 1.4, pp. 158–164. ISSN: 1936-5233.

Poh, Ashleigh R. and Matthias Ernst (June 2021). "Tumor-Associated Macrophages in Pancreatic Ductal Adenocarcinoma: Therapeutic Opportunities and Clinical Challenges". In: *Cancers* 13.12, p. 2860. ISSN: 2072-6694. DOI: 10.3390/cancers13122860.

Radev, Stefan T., Ulf K. Mertens, et al. (Apr. 2022). "BayesFlow: Learning Complex Stochastic Models With Invertible Neural Networks". In: *IEEE Transactions on Neural Networks and Learning Systems* 33.4, pp. 1452–1466. ISSN: 2162-2388. DOI: 10.1109/TNNLS.2020.3042395.

Radev, Stefan T., Marvin Schmitt, et al. (Sept. 2023). "BayesFlow: Amortized Bayesian Workflows With Neural Networks". en. In: *Journal of Open Source Software* 8.89, p. 5702. ISSN: 2475-9066. DOI: 10.21105/joss.05702.

Rajala, T., D. J. Murrell, and S. C. Olhede (Nov. 1, 2018). "Detecting Multivariate Interactions in Spatial Point Patterns with Gibbs Models and Variable Selection". In: *Journal of the Royal Statistical Society Series C: Applied Statistics* 67.5, pp. 1237–1273. ISSN: 0035-9254, 1467-9876. DOI: 10.1111/rssc.12281.

Raskov, Hans et al. (Jan. 2021). "Cytotoxic CD8+ T cells in cancer and cancer immunotherapy". In: *British Journal of Cancer* 124.2. Number: 2 Publisher: Nature Publishing Group, pp. 359–367. ISSN: 1532-1827. DOI: 10.1038/s41416-020-01048-4.

Renner, Ian W. et al. (2015). "Point process models for presence-only analysis". In: *Methods in Ecology and Evolution* 6.4. _eprint: https://besjournals.onlinelibrary.wiley.com/doi/pdf/10.1111/2041-210X.12352, pp. 366–379. ISSN: 2041-210X. DOI: 10.1111/2041-210X.12352.

Reynolds, Brent A., Monika W. Oli, and Madan K. Oli (2020). "Eco-oncology: Applying ecological principles to understand and manage cancer". In: *Ecology and Evolution* 10.16.

_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/ece3.6590, pp. 8538–8553. ISSN: 2045-7758. DOI: 10.1002/ece3.6590.

Säilynoja, Teemu, Paul-Christian Bürkner, and Aki Vehtari (Apr. 2022). "Graphical Test for Discrete Uniformity and its Applications in Goodness of Fit Evaluation and Multiple Sample Comparison". In: *Statistics and Computing* 32.2. arXiv:2103.10522 [stat], p. 32. ISSN: 0960-3174, 1573-1375. DOI: 10.1007/s11222-022-10090-6.

Sanderson, Conrad and Ryan Curtin (2016). "Armadillo: a template-based C++ library for linear algebra". In: *Journal of Open Source Software* 1, p. 26.

Sanyal, Nilotpal and Marco A. R. Ferreira (Nov. 2012). "Bayesian hierarchical multi-subject multiscale analysis of functional MRI data". In: *NeuroImage* 63.3, pp. 1519–1531. ISSN: 1053-8119. DOI: 10.1016/j.neuroimage.2012.08.041.

Schad, Daniel J., Michael Betancourt, and Shravan Vasishth (Feb. 2020). *Toward a principled Bayesian workflow in cognitive science*. arXiv:1904.12765 [stat]. DOI: 10.48550/arXiv.1904.12765.

Schapiro, Denis, Hartland W. Jackson, et al. (Sept. 2017). "histoCAT: analysis of cell phenotypes and interactions in multiplex image cytometry data". en. In: *Nature Methods* 14.9. Bandiera_abtest: a Cg_type: Nature Research Journals Number: 9 Primary_atype: Research Publisher: Nature Publishing Group Subject_term: Imaging;Software Subject_term_id: imaging;software, pp. 873–876. ISSN: 1548-7105. DOI: 10.1038/nmeth.4391.

Schapiro, Denis, Artem Sokolov, et al. (2022). "MCMICRO: a scalable, modular image-processing pipeline for multiplexed tissue imaging". In: *Nature methods* 19.3, pp. 311–315. DOI: 10.1038/s41592-021-01308-y.

Scheffé, Henry (1970). "Multiple Testing Versus Multiple Estimation. Improper Confidence Sets. Estimation of Directions and Ratios". In: *The Annals of Mathematical Statistics* 41.1. Publisher: Institute of Mathematical Statistics, pp. 1–29. ISSN: 0003-4851.

Schürch, Christian M. et al. (Sept. 3, 2020). "Coordinated Cellular Neighborhoods Orchestrate Antitumoral Immunity at the Colorectal Cancer Invasive Front". In: *Cell* 182.5, 1341–1359.e19. ISSN: 1097-4172. DOI: 10.1016/j.cell.2020.07.005.

Sheng, Wenjie et al. (Feb. 2023). "Multiplex Immunofluorescence: A Powerful Tool in Cancer Immunotherapy". In: *International Journal of Molecular Sciences* 24.4, p. 3086. ISSN: 1422-0067. DOI: 10.3390/ijms24043086.

Sikora, Andrew G. et al. (Mar. 2010). "Targeted inhibition of inducible nitric oxide synthase inhibits growth of human melanoma in vivo and synergizes with chemotherapy". In: *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research* 16.6, pp. 1834–1844. ISSN: 1557-3265. DOI: 10.1158/1078-0432.CCR-09-3123.

Somarelli, Jason A. (Apr. 28, 2021). "The Hallmarks of Cancer as Ecologically Driven Phenotypes". In: *Frontiers in Ecology and Evolution* 9, p. 661583. ISSN: 2296-701X. DOI: 10.3389/fevo.2021.661583.

Spiekerman, C. F. and D. Y. Lin (1998). "Marginal Regression Models for Multivariate Failure Time Data". In: *Journal of the American Statistical Association* 93.443. Publisher: [American Statistical Association, Taylor & Francis, Ltd.], pp. 1164–1175. ISSN: 0162-1459. DOI: 10.2307/2669859.

Stan Development Team (2024). *Stan Reference Manual*. version 2.36.0.

— (2025). *Stan User's Guide - Reparameterization*.

Stoyan, Dietrich and Antti Penttinen (2000). "Recent Applications of Point Process Methods in Forestry Statistics". In: *Statistical Science* 15.1. Publisher: Institute of Mathematical Statistics, pp. 61–78. ISSN: 0883-4237.

Su, Yahui et al. (Dec. 2024). "The cross-talk between B cells and macrophages". In: *International Immunopharmacology* 143, p. 113463. ISSN: 1567-5769. DOI: 10.1016/j.intimp.2024.113463.

Summers, Huw D., John W. Wills, and Paul Rees (Nov. 2022). "Spatial statistics is a comprehensive tool for quantifying cell neighbor relationships and biological processes via tissue image analysis". In: *Cell Reports Methods* 2.11, p. 100348. ISSN: 2667-2375. DOI: 10.1016/j.crmeth.2022.100348.

Sun, Yu (Sept. 28, 2016). "Tumor microenvironment and cancer therapy resistance". In: *Cancer Letters* 380.1, pp. 205–215. ISSN: 0304-3835. DOI: 10.1016/j.canlet.2015.07.044.

Talts, Sean et al. (Oct. 2020). *Validating Bayesian Inference Algorithms with Simulation-Based Calibration*. arXiv:1804.06788 [stat]. DOI: 10.48550/arXiv.1804.06788.

Tan, Keely and Matthew J. Naylor (May 13, 2022). "Tumour Microenvironment-Immune Cell Interactions Influencing Breast Cancer Heterogeneity and Disease Progression". In: *Frontiers in Oncology* 12, p. 876451. ISSN: 2234-943X. DOI: 10.3389/fonc.2022.876451.

Tanese, Keiji, Yuuri Hashimoto, et al. (Nov. 2015). "Cell Surface CD74-MIF Interactions Drive Melanoma Survival in Response to Interferon-γ". In: *The Journal of Investigative Dermatology* 135.11, p. 2901. ISSN: 1523-1747. DOI: 10.1038/jid.2015.259.

Tanese, Keiji and Dai Ogata (2024). "The role of macrophage migration inhibitory factor family and CD74 in the pathogenesis of melanoma". In: *Experimental Dermatology* 33.6. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/exd.15122, e15122. ISSN: 1600-0625. DOI: 10.1111/exd.15122.

Tanevski, Jovan et al. (Apr. 14, 2022). "Explainable multiview framework for dissecting spatial relationships from highly multiplexed data". In: *Genome Biology* 23.1, p. 97. ISSN: 1474-760X. DOI: 10.1186/s13059-022-02663-5.

Therneau, Terry M. et al. (Dec. 2024). *survival: Survival Analysis*.

Thomas, Frédéric et al. (2020). "The interface between ecology, evolution, and cancer: More than ever a relevant research direction for both oncologists and ecologists". In: *Evolutionary Applications* 13.7. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/eva.13031, pp. 1545–1549. ISSN: 1752-4571. DOI: 10.1111/eva.13031.

Tikhonov, Gleb, Nerea Abrego, et al. (2017). "Using joint species distribution models for evaluating how species-to-species associations depend on the environmental context". In: *Methods in Ecology and Evolution* 8.4. _eprint: https://besjournals.onlinelibrary.wiley.com/doi/pdf/10.1111/2041-210X.12723, pp. 443–452. ISSN: 2041-210X. DOI: 10.1111/2041-210X.12723.

Tikhonov, Gleb, Øystein H. Opedal, et al. (2020). "Joint species distribution modelling with the r-package Hmsc". In: *Methods in Ecology and Evolution* 11.3. _eprint: https://besjournals.onlinelibrary.wiley.com/doi/pdf/10.1111/2041-210X.13345, pp. 442–447. ISSN: 2041-210X. DOI: 10.1111/2041-210X.13345.

Tsang, Ashley P. et al. (Nov. 2024). "Assessing the Tumor Immune Landscape Across Multiple Spatial Scales to Differentiate Immunotherapy Response in Metastatic Non-Small Cell Lung Cancer". In: *Laboratory Investigation* 104.11, p. 102148. ISSN: 0023-6837. DOI: 10.1016/j.labinv.2024.102148.

Vecchia, A. V. (Jan. 1988). "Estimation and Model Identification for Continuous Spatial Processes". In: *Journal of the Royal Statistical Society: Series B (Methodological)* 50.2, pp. 297–312. ISSN: 0035-9246. DOI: 10.1111/j.2517-6161.1988.tb01729.x.

Vehtari, Aki et al. (June 2021). "Rank-Normalization, Folding, and Localization: An Improved R^ for Assessing Convergence of MCMC (with Discussion)". In: *Bayesian Analysis* 16.2. Publisher: International Society for Bayesian Analysis, pp. 667–718. ISSN: 1936-0975, 1931-6690. DOI: 10.1214/20-BA1221.

Velázquez, Eduardo et al. (2016). "An evaluation of the state of spatial point pattern analysis in ecology". In: *Ecography* 39.11. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/ecog.01579, pp. 1042–1055. ISSN: 1600-0587. DOI: 10.1111/ecog.01579.

Virtanen, Pauli et al. (2020). "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python". In: *Nature Methods* 17, pp. 261–272. DOI: 10.1038/s41592-019-0686-2.

Visser, Karin E. de and Johanna A. Joyce (Mar. 13, 2023). "The evolving tumor microenvironment: From cancer initiation to metastatic outgrowth". In: *Cancer Cell* 41.3, pp. 374–403. ISSN: 1878-3686. DOI: 10.1016/j.ccell.2023.02.016.

Vito, Roberta De et al. (Dec. 2021). "Bayesian multistudy factor analysis for high-throughput biological data". In: *The Annals of Applied Statistics* 15.4. Publisher: Institute of Mathematical Statistics, pp. 1723–1741. ISSN: 1932-6157, 1941-7330. DOI: 10.1214/21-AOAS1456.

Vu, Thao et al. (June 15, 2022). "SPF: A spatial and functional data analytic approach to cell imaging data". In: *PLOS Computational Biology* 18.6. Publisher: Public Library of Science, e1009486. ISSN: 1553-7358. DOI: 10.1371/journal.pcbi.1009486.

Wang, Jiuzhou and Eric F. Lock (Jan. 2024). "Multiple augmented reduced rank regression for pan-cancer analysis". eng. In: *Biometrics* 80.1, ujad002. ISSN: 1541-0420. DOI: 10.1093/biomtc/ujad002.

Warchol, Simon et al. (Jan. 2023). "Visinity: Visual Spatial Neighborhood Analysis for Multiplexed Tissue Imaging Data". In: *IEEE transactions on visualization and computer graphics* 29.1, pp. 106–116. ISSN: 1077-2626. DOI: 10.1109/TVCG.2022.3209378.

Watanabe, Sumio (Dec. 2010). "Asymptotic Equivalence of Bayes Cross Validation and Widely Applicable Information Criterion in Singular Learning Theory". In: *J. Mach. Learn. Res.* 11, pp. 3571–3594. ISSN: 1532-4435.

Wen, Qingsong et al. (Aug. 2023). "Transformers in time series: a survey". In: *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence*. IJCAI '23. Macao, P.R.China, pp. 6778–6786. ISBN: 978-1-956792-03-4. DOI: 10.24963/ijcai.2023/759.

Wickham, Hadley et al. (2019). "Welcome to the tidyverse". In: *Journal of Open Source Software* 4.43, p. 1686. DOI: 10.21105/joss.01686.

Wilkinson, David P. et al. (2021). "Defining and evaluating predictions of joint species distribution models". In: *Methods in Ecology and Evolution* 12.3. _eprint: https://besjournals.onlinelibrary.wiley.com/doi/pdf/10.1111/2041-210X.13518, pp. 394–404. ISSN: 2041-210X. DOI: 10.1111/2041-210X.13518.

Wortman, Juliana C. et al. (July 2021). "Spatial distribution of B cells and lymphocyte clusters as a predictor of triple-negative breast cancer outcome". en. In: *npj Breast Cancer* 7.1. Publisher: Nature Publishing Group, pp. 1–13. ISSN: 2374-4677. DOI: 10.1038/s41523-021-00291-z.

Wrobel, Julia, Coleman Harris, and Simon Vandekar (2023). "Statistical Analysis of Multiplex Immunofluorescence and Immunohistochemistry Imaging Data". en. In: *Statistical Genomics*. Ed. by Brooke Fridley and Xuefeng Wang. New York, NY: Springer US, pp. 141–168. ISBN: 978-1-0716-2986-4. DOI: 10.1007/978-1-0716-2986-4_8.

Wu, Ling and Xiang H.-F. Zhang (Dec. 2020). "Tumor-Associated Neutrophils and Macrophages—Heterogenous but Not Chaotic". In: *Frontiers in Immunology* 11, p. 553967. ISSN: 1664-3224. DOI: 10.3389/fimmu.2020.553967.

Wu, Zhenqin et al. (Dec. 2022). "Graph deep learning for the characterization of tumour microenvironments from spatial protein profiles in tissue specimens". en. In: *Nature*

Biomedical Engineering 6.12. Publisher: Nature Publishing Group, pp. 1435–1448. ISSN: 2157-846X. DOI: 10.1038/s41551-022-00951-w.

Xie, Yuqi et al. (Apr. 2025). "Epithelial-mesenchymal transition orchestrates tumor microenvironment: current perceptions and challenges". In: *Journal of Translational Medicine* 23, p. 386. ISSN: 1479-5876. DOI: 10.1186/s12967-025-06422-5.

Xu, Lei et al. (Dec. 2009). "Modeling Inter-Subject Variability in fMRI Activation Location: A Bayesian Hierarchical Spatial Model". In: *Biometrics* 65.4, pp. 1041–1051. ISSN: 0006-341X. DOI: 10.1111/j.1541-0420.2008.01190.x.

Yuan, Yinyin (Aug. 2016). "Spatial Heterogeneity in the Tumor Microenvironment". In: *Cold Spring Harbor Perspectives in Medicine* 6.8, a026583. ISSN: 2157-1422. DOI: 10.1101/cshperspect.a026583.

Yuan, Zhiyuan et al. (Nov. 28, 2022). "SOTIP is a versatile method for microenvironment modeling with spatial omics data". In: *Nature Communications* 13.1. Number: 1 Publisher: Nature Publishing Group, p. 7330. ISSN: 2041-1723. DOI: 10.1038/s41467-022-34867-5.

Zhang, Xianyi (2020). *An Optimized BLAS Library Based on GotoBLAS2.*

Zhou, Yongqiang et al. (Feb. 2023). "NK cells are never alone: crosstalk and communication in tumour microenvironments". In: *Molecular Cancer* 22.1, p. 34. ISSN: 1476-4598. DOI: 10.1186/s12943-023-01737-7.

Zhu, Yichen et al. (Dec. 2024). "Radial neighbours for provably accurate scalable approximations of Gaussian processes". In: *Biometrika* 111.4, pp. 1151–1167. ISSN: 1464-3510. DOI: 10.1093/biomet/asae029.