# Uncovering Causality in Change Point Regressions

Shakeel Gavioli-Akilagun  Piotr Fryzlewicz

London School of Economics
Department of Statistics

# Change Point Regression

- Modern renaissance in change point detection...
- In applications there is often **reason to believe change points are dynamically dependent or even causally linked**
- Since typically change points are held to be non-stochastic, such considerations are usually meaningless

# Change Point Regression

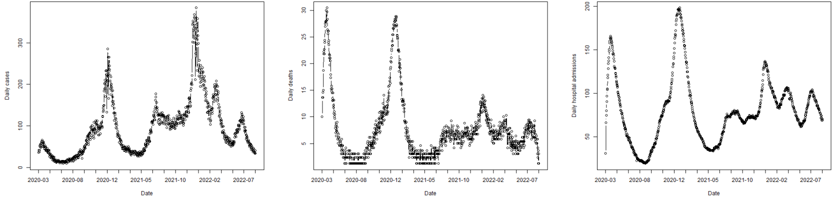
- Modern renaissance in change point detection...
- In applications there is often **reason to believe change points are dynamically dependent or even causally linked**
- Since typically change points are held to be non-stochastic, such considerations are usually meaningless

# Change Point Regression

- Modern renaissance in change point detection...
- In applications there is often **reason to believe change points are dynamically dependent or even causally linked**
- Since typically change points are held to be non-stochastic, such considerations are usually meaningless

# Motivating Problem 1: *changes in COVID-19 trajectories*

- Abundance of recent papers modelling COVID-19 trajectories as piecewise linear (Jiang et al. 2021, Anastasiou & Papanastasiou 2022, etc.)

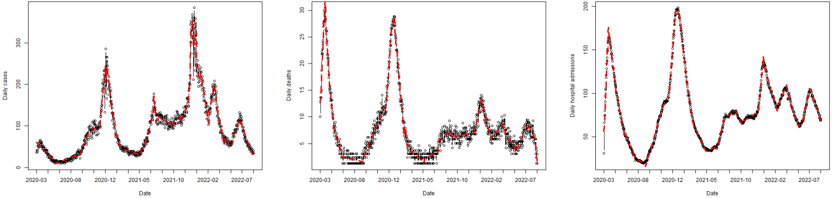



(a) Cases (b) Deaths (c) Hospitalizations

Figure 1: Anscombe transform of daily COVID-19 trajectories across all London boroughs

# Motivating Problem 1: *changes in COVID-19 trajectories*

- Abundance of recent papers modelling COVID-19 trajectories as piecewise linear (Jiang et al. 2021, Anastasiou & Papanastasiou 2022, etc.)




(a) Cases (b) Deaths (c) Hospitalizations

Figure 2: Anscombe transform of daily COVID-19 trajectories across all London boroughs along with piecewise linear trend (- - -) recovered using the NOT algorithm (Baranowski et al. 2019) equipped with seeded intervals (Kovács, Li & Bühlmann 2020)t

# Motivating Problem 1: *changes in COVID-19 trajectories*

- We may have reason to suspect that changes in case numbers "cause" changes in hospitalizations, deaths, etc. However looking at estimated change point locations **the relationship is not so clear**

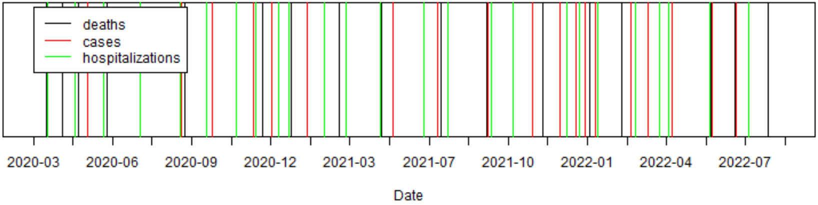
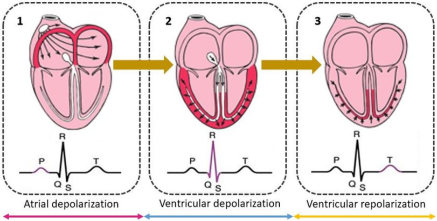



Figure 3: Locations of change points in Anscombe transformed series of daily COVID-19 cases, daily COVID-19 deaths, and daily hospitalizations recovered by the NOT algorithm

# Motivating Problem 2: *cardiac response to live music performance*

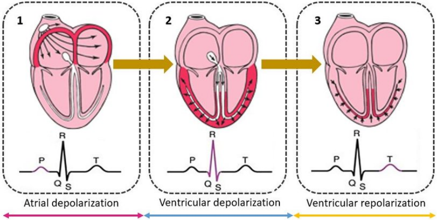
- Joint ongoing work with **Elaine Chew** and **Pier Lambiase** from KCL. Patients with biventricular pacemakers invited to listen to live piano concerts while cardiac activity monitored




- To track millisecond changes in the heart's action potential duration in response to the music, **each patient's pacemaker was re-programmed to remove heart rate variations**

# Motivating Problem 2: *cardiac response to live music performance*

- Joint ongoing work with **Elaine Chew** and **Pier Lambiase** from KCL. Patients with biventricular pacemakers invited to listen to live piano concerts while cardiac activity monitored




- To track millisecond changes in the heart's action potential duration in response to the music, **each patient's pacemaker was re-programmed to remove heart rate variations**

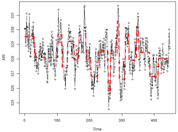
# Motivating Problem 2: *cardiac response to live music performance*

- **Goal:** we are interested in understanding the mechanism by which stress and strong emotions de-stabilises the heart's electrical pathways

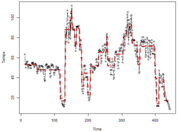

Figure 4: Beginning of the second "Presto con fuoco" in Chopin's Ballade No. 2

# Motivating Problem 2: *cardiac response to live music performance*

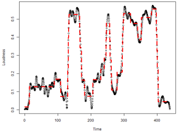
- **Goal:** we are interested in understanding the mechanism by which stress and strong emotions de-stabilises the heart's electrical pathways



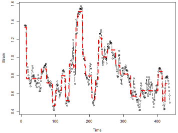
(a) ARI



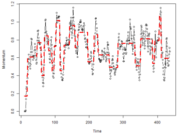
(b) tempo



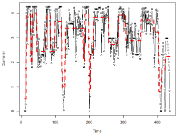
(c) loudness



(d) strain



(e) momentum



(f) diameter

# Motivating Problem 2: *cardiac response to live music performance*

Many others are involved with this project!

- **Study Leads:** Elaine Chew, Pier Lambiase, Peter Taggart
- **Patient Recruitment & Consent:** Hakkam Abbass, Peter Waddingham
- **Abbott support:** Louise Sheil, Jan Mangual
- **Cardiologists on-site:** P. Lambiase, P. Waddingham
- **Physiologists on-site:** Holly Daw, Daniel Meese, Genine Sambile
- **Audio Recording:** James Weaver
- **Valence-arousal App:** Courtney Reed
- **Tension Rating App:** Shamindra De Zylva
- **Interviewers:** Vanessa Pope, Sebastian Ruiz, Changhong Wang, Simin Yang
- **Cardiac Signal Processing:** E. Chew, P. Taggart, J. Mangual, Daniel Bedoya
- **Music Signal Processing:** E. Chew, Corentin Guichaoua, D. Bedoya

# Problem Statement

- We consider $d$ data streams $\{\mathbf{Y}_j\}$ from 'signal + noise' models where the signals are piecewise parametric (i.e. piecewise constant)

$$Y_{jt} = f_j\left(t/n\right) + \zeta_{jt} \quad t = 1, \ldots n$$

- Associated with homogeneous regions of $f_j(\cdot)$ are $N_j$ change point locations $\eta_{j1} < \eta_{j2} < \ldots$ or equivalently **a counting measure** $N_j(\cdot)$
- **Goal:** recover a graph $\mathcal{G} = G(\mathcal{V}, \mathcal{E})$ where $\mathcal{V} = \{1, \ldots, d\}$ and in some sense $(i,j) \notin \mathcal{E} \Rightarrow$ "*$N_i(\cdot)$ does not cause $N_j(\cdot)$*"

# Problem Statement

- We consider $d$ data streams $\{\mathbf{Y}_j\}$ from 'signal + noise' models where the signals are piecewise parametric (i.e. piecewise constant)

$$Y_{jt} = f_j\left(t/n\right) + \zeta_{jt} \quad t = 1, \ldots n$$

- Associated with homogeneous regions of $f_j\left(\cdot\right)$ are $N_j$ change point locations $\eta_{j1} < \eta_{j2} < \ldots$ or equivalently **a counting measure** $N_j\left(\cdot\right)$
- **Goal:** recover a graph $\mathcal{G} = G\left(\mathcal{V}, \mathcal{E}\right)$ where $\mathcal{V} = \{1, \ldots, d\}$ and in some sense $(i, j) \notin \mathcal{E} \Rightarrow$ "$N_i\left(\cdot\right)$ *does not cause* $N_j\left(\cdot\right)$"

# Problem Statement

- We consider $d$ data streams $\{\mathbf{Y}_j\}$ from 'signal + noise' models where the signals are piecewise parametric (i.e. piecewise constant)

$$Y_{jt} = f_j(t/n) + \zeta_{jt} \quad t = 1, \ldots n$$

- Associated with homogeneous regions of $f_j(\cdot)$ are $N_j$ change point locations $\eta_{j1} < \eta_{j2} < \ldots$ or equivalently **a counting measure** $N_j(\cdot)$
- **Goal:** recover a graph $\mathcal{G} = G(\mathcal{V}, \mathcal{E})$ where $\mathcal{V} = \{1, \ldots, d\}$ and in some sense $(i, j) \notin \mathcal{E} \Rightarrow$ "$N_i(\cdot)$ *does not cause* $N_j(\cdot)$"

# Change Points as Marked Point Processes

- Change points should be stochastic, therefore $\{N_j(\cdot)\}$ are random measures
- Fixing $j$ each sequence $\{\eta_{jk}\}$ is naturally associated with a **point process** and each sequence $\{(j, \eta_{jk})\}$ is associated with a **marked point process**
- The probabilistic structure of each $N_j(\cdot)$ is completely specified by the local intensity function (Daley et al. 2003) which is defined as

$$\lambda_j^*(t) = \lim_{s \downarrow 0} \frac{1}{s} \mathbb{E}\left\{N_j(t, t+s] \mid \sigma(\eta_{jk} < t)\right\}$$

# Change Points as Marked Point Processes

- Change points should be stochastic, therefore $\{N_j(\cdot)\}$ are random measures
- Fixing $j$ each sequence $\{\eta_{jk}\}$ is naturally associated with a **point process** and each sequence $\{(j, \eta_{jk})\}$ is associated with a **marked point process**
- The probabilistic structure of each $N_j(\cdot)$ is completely specified by the local intensity function (Daley et al. 2003) which is defined as

$$\lambda_j^*(t) = \lim_{s \downarrow 0} \frac{1}{s} \mathbb{E}\{N_j(t, t+s] \mid \sigma(\eta_{jk} < t)\}$$

# Change Points as Marked Point Processes

- Change points should be stochastic, therefore $\{N_j(\cdot)\}$ are random measures
- Fixing $j$ each sequence $\{\eta_{jk}\}$ is naturally associated with a **point process** and each sequence $\{(j, \eta_{jk})\}$ is associated with a **marked point process**
- The probabilistic structure of each $N_j(\cdot)$ is completely specified by the local intensity function (Daley et al. 2003) which is defined as

$$\lambda_j^*(t) = \lim_{s \downarrow 0} \frac{1}{s} \mathbb{E}\left\{N_j(t, t+s] \mid \sigma(\eta_{jk} < t)\right\}$$

# Local Independence for Marked Point Processes

## Local Independence (Didelez 2008)

Let $A, B, C$ be disjoint subsets of $\{1, \ldots, d\}$ and put $\mathcal{H}_t^D = \sigma\left(\eta_{jk} < t \mid j \in D\right)$. $\mathbf{N}_B\left(\cdot\right)$ is called locally independent of $\mathbf{N}_A\left(\cdot\right)$ given $\mathbf{N}_C\left(\cdot\right)$ if all $\mathcal{H}_t^{A\cup B\cup C}$-intensities $\lambda_j^*\left(\cdot\right)$ with $j \in B$ are $\mathcal{H}_t^{B\cup C}$ measurable, and we write $A \not\rightarrow B \mid C$. Else we speak of local dependence, and write $A \rightarrow B \mid C$.

- **Asymmetry:** $A \rightarrow B \mid C$ does not imply $B \rightarrow A \mid C$
- **Locality:** $A \not\rightarrow B \mid C$ does not imply $B \perp\!\!\!\perp A \mid C$
- **Causality:** if all relevant processes are observed one can speak of causality

# Local Independence for Marked Point Processes

## Local Independence (Didelez 2008)

Let $A, B, C$ be disjoint subsets of $\{1, \dots, d\}$ and put $\mathcal{H}_t^D = \sigma(\eta_{jk} < t \mid j \in D)$. $\mathbf{N}_B(\cdot)$ is called locally independent of $\mathbf{N}_A(\cdot)$ given $\mathbf{N}_C(\cdot)$ if all $\mathcal{H}_t^{A \cup B \cup C}$-intensities $\lambda_j^*(\cdot)$ with $j \in B$ are $\mathcal{H}_t^{B \cup C}$ measurable, and we write $A \not\rightarrow B \mid C$. Else we speak of local dependence, and write $A \rightarrow B \mid C$.

- **Asymmetry:** $A \rightarrow B \mid C$ does not imply $B \rightarrow A \mid C$
- **Locality:** $A \not\rightarrow B \mid C$ does not imply $B \perp\!\!\!\perp A \mid C$
- **Causality:** if all relevant processes are observed one can speak of causality

# Local Independence for Marked Point Processes

## Local Independence (Didelez 2008)

Let $A, B, C$ be disjoint subsets of $\{1, \ldots, d\}$ and put $\mathcal{H}_t^D = \sigma\left(\eta_{jk} < t \mid j \in D\right)$. $\mathbf{N}_B\left(\cdot\right)$ is called locally independent of $\mathbf{N}_A\left(\cdot\right)$ given $\mathbf{N}_C\left(\cdot\right)$ if all $\mathcal{H}_t^{A \cup B \cup C}$-intensities $\lambda_j^*\left(\cdot\right)$ with $j \in B$ are $\mathcal{H}_t^{B \cup C}$ measurable, and we write $A \not\rightarrow B \mid C$. Else we speak of local dependence, and write $A \rightarrow B \mid C$.

- **Asymmetry:** $A \rightarrow B \mid C$ does not imply $B \rightarrow A \mid C$
- **Locality:** $A \not\rightarrow B \mid C$ does not imply $B \perp\!\!\!\perp A \mid C$
- **Causality:** if all relevant processes are observed one can speak of causality

# Local Independence for Marked Point Processes

## Local Independence (Didelez 2008)

Let $A, B, C$ be disjoint subsets of $\{1, \ldots, d\}$ and put $\mathcal{H}_t^D = \sigma\left(\eta_{jk} < t \mid j \in D\right)$. $\mathbf{N}_B\left(\cdot\right)$ is called locally independent of $\mathbf{N}_A\left(\cdot\right)$ given $\mathbf{N}_C\left(\cdot\right)$ if all $\mathcal{H}_t^{A\cup B\cup C}$-intensities $\lambda_j^*\left(\cdot\right)$ with $j \in B$ are $\mathcal{H}_t^{B\cup C}$ measurable, and we write $A \not\rightarrow B \mid C$. Else we speak of local dependence, and write $A \rightarrow B \mid C$.

- **Asymmetry:** $A \rightarrow B \mid C$ does not imply $B \rightarrow A \mid C$
- **Locality:** $A \not\rightarrow B \mid C$ does not imply $B \perp\!\!\!\perp A \mid C$
- **Causality:** if all relevant processes are observed one can speak of causality

# Change Points as Hawkes Processes

- **Point process sample paths should be multi-scale.** A natural candidate is the Hawkes process (Hawkes 1971*a*, Hawkes 1971*b*), which can be thought of as the point process analogue of an auto-regression

$$\lambda_j^*(t) = \mu_j + \sum_{i=1}^{d} \int_{-\infty}^{t} g_{ij}(t-u)\,\mathrm{d}N_i(u)$$

- Generally we require each $g_{ij} \geqslant 0$ else $\lambda_j^*(\cdot)$ may be negative
- Here, local independence is naturally understood in terms of the kernels $\{g_{ij}\}$

$$A \not\rightarrow j \Leftrightarrow \left\{ i \in A \mid \int g_{ij} > 0 \right\} = \emptyset$$

- Following Didelez (2008) we aim to recover $\mathcal{E} = \left\{(i,j) \mid \int g_{ij} > 0\right\}$

# Change Points as Hawkes Processes

- **Point process sample paths should be multi-scale.** A natural candidate is the Hawkes process (Hawkes 1971*a*, Hawkes 1971*b*), which can be thought of as the point process analogue of an auto-regression

$$\lambda_j^*(t) = \mu_j + \sum_{i=1}^{d} \int_{-\infty}^{t} g_{ij}(t-u)\,\mathrm{d}N_i(u)$$

- Generally we require each $g_{ij} \geqslant 0$ else $\lambda_j^*(\cdot)$ may be negative
- Here, local independence is naturally understood in terms of the kernels $\{g_{ij}\}$

$$A \not\to j \Leftrightarrow \left\{ i \in A \mid \int g_{ij} > 0 \right\} = \emptyset$$

- Following Didelez (2008) we aim to recover $\mathcal{E} = \left\{ (i,j) \mid \int g_{ij} > 0 \right\}$

# Change Points as Hawkes Processes

- **Point process sample paths should be multi-scale.** A natural candidate is the Hawkes process (Hawkes 1971*a*, Hawkes 1971*b*), which can be thought of as the point process analogue of an auto-regression

$$\lambda_j^*(t) = \mu_j + \sum_{i=1}^{d} \int_{-\infty}^{t} g_{ij}(t-u)\,\mathrm{d}N_i(u)$$

- Generally we require each $g_{ij} \geqslant 0$ else $\lambda_j^*(\cdot)$ may be negative
- Here, local independence is naturally understood in terms of the kernels $\{g_{ij}\}$

$$A \nrightarrow j \Leftrightarrow \left\{ i \in A \mid \int g_{ij} > 0 \right\} = \emptyset$$

- Following Didelez (2008) we aim to recover $\mathcal{E} = \left\{(i,j) \mid \int g_{ij} > 0\right\}$

# Change Points as Hawkes Processes

- **Point process sample paths should be multi-scale.** A natural candidate is the Hawkes process (Hawkes 1971*a*, Hawkes 1971*b*), which can be thought of as the point process analogue of an auto-regression

$$\lambda_j^*(t) = \mu_j + \sum_{i=1}^{d} \int_{-\infty}^{t} g_{ij}(t-u)\,\mathrm{d}N_i(u)$$

- Generally we require each $g_{ij} \geqslant 0$ else $\lambda_j^*(\cdot)$ may be negative
- Here, local independence is naturally understood in terms of the kernels $\{g_{ij}\}$

$$A \not\rightarrow j \Leftrightarrow \left\{ i \in A \mid \int g_{ij} > 0 \right\} = \emptyset$$

- Following Didelez (2008) we aim to recover $\mathcal{E} = \left\{ (i,j) \mid \int g_{ij} > 0 \right\}$

# Recovering Hawkes Kernels

- Kirchner (2017) and Eichler et al. (2017) independently proposed the same simple non-parametric method for estimating Hawkes kernels
- Putting $Z_{jt}^h = N_j[(t-1)h, th)$ for some $h \downarrow 0$ and $\omega \uparrow \infty$ it holds that

$$\begin{aligned}
\mathbb{E}\left\{Z_{jt}^h \mid \sigma\left(\eta_{jk} < th\right)\right\} &= \mathbb{P}\left\{N_j(th,(t+1)h] = 1 \mid \sigma\left(\eta_{jk} < th\right)\right\} + o\left(h\right) \\
&= h\mu_j + \sum_{i=1}^{d} \int_{-\infty}^{t} g_{ij}\left(t-u\right) \mathrm{d}N_i\left(u\right) + o\left(h\right) \\
&\approx h\mu_j + \sum_{i=1}^{d} \sum_{\ell=1}^{\omega} h g_{ij}\left(\ell h\right)
\end{aligned}$$

- Then, the kernels $\{g_{ij}\}$ can be estimated at discrete points using least squares

# Recovering Hawkes Kernels

- Kirchner (2017) and Eichler et al. (2017) independently proposed the same simple non-parametric method for estimating Hawkes kernels
- Putting $Z_{jt}^h = N_j[(t-1)h, th)$ for some $h \downarrow 0$ and $\omega \uparrow \infty$ it holds that

$$\begin{aligned}
\mathbb{E}\left\{Z_{jt}^h \mid \sigma\left(\eta_{jk} < th\right)\right\} &= \mathbb{P}\left\{N_j(th, (t+1)h] = 1 \mid \sigma\left(\eta_{jk} < th\right)\right\} + o\left(h\right) \\
&= h\mu_j + \sum_{i=1}^{d} \int_{-\infty}^{t} g_{ij}\left(t-u\right) \mathrm{d}N_i\left(u\right) + o\left(h\right) \\
&\approx h\mu_j + \sum_{i=1}^{d} \sum_{\ell=1}^{\omega} hg_{ij}\left(\ell h\right)
\end{aligned}$$

- Then, the kernels $\{g_{ij}\}$ can be estimated at discrete points using least squares

# Recovering Hawkes Kernels

- Kirchner (2017) and Eichler et al. (2017) independently proposed the same simple non-parametric method for estimating Hawkes kernels
- Putting $Z_{jt}^h = N_j[(t-1)h, th)$ for some $h \downarrow 0$ and $\omega \uparrow \infty$ it holds that

$$\begin{aligned}
\mathbb{E}\left\{Z_{jt}^h \mid \sigma\left(\eta_{jk} < th\right)\right\} &= \mathbb{P}\left\{N_j(th,(t+1)h] = 1 \mid \sigma\left(\eta_{jk} < th\right)\right\} + o\left(h\right) \\
&= h\mu_j + \sum_{i=1}^{d} \int_{-\infty}^{t} g_{ij}\left(t-u\right) \mathrm{d}N_i\left(u\right) + o\left(h\right) \\
&\approx h\mu_j + \sum_{i=1}^{d} \sum_{\ell=1}^{\omega} h g_{ij}\left(\ell h\right)
\end{aligned}$$

- Then, the kernels $\{g_{ij}\}$ can be estimated at discrete points using least squares

# Graph Estimation using Conditional Least Squares

- **STEP 1:** Choose bin width $h$, auto-regression order $\omega$, and thresholds $\{\lambda_{ij}\}$
- **STEP 2:** From $\{\mathbf{Y}_j\}$ estimate change points $\{\widehat{\eta}_{jk}\}$ via some base method and put $\widehat{Z}_{jt}^{h} := \#\{\widehat{\eta}_{jk} \mid (t-t)h \leqslant \widehat{\eta}_{jk} < th\}$ and
- **STEP 3:** estimate $\{\widehat{g_{ij}}(\ell h)\}$ from $\left\{\widehat{Z}_{jt}^{h}\right\}$ using least squares
- **STEP 4:** put $\widehat{\mathcal{E}} = \{(i,j) \mid h\sum_{\ell=1}^{\omega} \widehat{g_{ij}}(\ell h) > \lambda_{ij}\}$

*Remark:* $\{\lambda_{ij}\}$ *can be chosen to control some desirable quantity respective the population edge set* $\mathcal{E}$ *such as FDR, FWE, etc.*

# Graph Estimation using Conditional Least Squares

- **STEP 1:** Choose bin width $h$, auto-regression order $\omega$, and thresholds $\{\lambda_{ij}\}$
- **STEP 2:** From $\{\mathbf{Y}_j\}$ estimate change points $\{\widehat{\eta}_{jk}\}$ via some base method and put $\widehat{Z}_{jt}^h := \#\{\widehat{\eta}_{jk} \mid (t-t)h \leqslant \widehat{\eta}_{jk} < th\}$ and
- **STEP 3:** estimate $\{\widehat{g_{ij}}(\ell h)\}$ from $\left\{\widehat{Z}_{jt}^h\right\}$ using least squares
- **STEP 4:** put $\widehat{\mathcal{E}} = \{(i,j) \mid h\sum_{\ell=1}^{\omega} \widehat{g_{ij}}(\ell h) > \lambda_{ij}\}$

*Remark:* $\{\lambda_{ij}\}$ *can be chosen to control some desirable quantity respective the population edge set* $\mathcal{E}$ *such as FDR, FWE, etc.*

# Graph Estimation using Conditional Least Squares

- **STEP 1:** Choose bin width $h$, auto-regression order $\omega$, and thresholds $\{\lambda_{ij}\}$
- **STEP 2:** From $\{\mathbf{Y}_j\}$ estimate change points $\{\widehat{\eta}_{jk}\}$ via some base method and put $\widehat{Z}_{jt}^h := \#\{\widehat{\eta}_{jk} \mid (t-t)h \leqslant \widehat{\eta}_{jk} < th\}$ and
- **STEP 3:** estimate $\{\widehat{g_{ij}}(\ell h)\}$ from $\left\{\widehat{Z}_{jt}^h\right\}$ using least squares
- **STEP 4:** put $\widehat{\mathcal{E}} = \{(i,j) \mid h\sum_{\ell=1}^{\omega} \widehat{g_{ij}}(\ell h) > \lambda_{ij}\}$

*Remark: $\{\lambda_{ij}\}$ can be chosen to control some desirable quantity respective the population edge set $\mathcal{E}$ such as FDR, FWE, etc.*

# Graph Estimation using Conditional Least Squares

- **STEP 1:** Choose bin width $h$, auto-regression order $\omega$, and thresholds $\{\lambda_{ij}\}$
- **STEP 2:** From $\{\mathbf{Y}_j\}$ estimate change points $\{\widehat{\eta}_{jk}\}$ via some base method and put $\widehat{Z}_{jt}^{h} := \#\{\widehat{\eta}_{jk} \mid (t-t)h \leqslant \widehat{\eta}_{jk} < th\}$ and
- **STEP 3:** estimate $\{\widehat{g_{ij}}(\ell h)\}$ from $\left\{\widehat{Z}_{jt}^{h}\right\}$ using least squares
- **STEP 4:** put $\widehat{\mathcal{E}} = \{(i,j) \mid h\sum_{\ell=1}^{\omega} \widehat{g_{ij}}(\ell h) > \lambda_{ij}\}$

*Remark: $\{\lambda_{ij}\}$ can be chosen to control some desirable quantity respective the population edge set $\mathcal{E}$ such as FDR, FWE, etc.*

# Graph Estimation using Conditional Least Squares

- **STEP 1:** Choose bin width $h$, auto-regression order $\omega$, and thresholds $\{\lambda_{ij}\}$
- **STEP 2:** From $\{\mathbf{Y}_j\}$ estimate change points $\{\widehat{\eta}_{jk}\}$ via some base method and put $\widehat{Z}^h_{jt} := \#\{\widehat{\eta}_{jk} \mid (t-t)h \leqslant \widehat{\eta}_{jk} < th\}$ and
- **STEP 3:** estimate $\{\widehat{g_{ij}}(\ell h)\}$ from $\left\{\widehat{Z}^h_{jt}\right\}$ using least squares
- **STEP 4:** put $\widehat{\mathcal{E}} = \{(i,j) \mid h\sum_{\ell=1}^{\omega} \widehat{g_{ij}}(\ell h) > \lambda_{ij}\}$

*Remark:* $\{\lambda_{ij}\}$ *can be chosen to control some desirable quantity respective the population edge set* $\mathcal{E}$ *such as FDR, FWE, etc.*

# Simulation Study

We simulate 100 sample paths of length $n \in \{500, 1000\}$ from piecewise-constant mean model driven by a Hawkes processes with exponential kernels

- Base-line intensities: $\mu_j = n^{-7/10}$
- Kernels when non-zero: $g_{ij}(u) = \frac{0.6}{n^{3/10}} \exp\left(-\frac{0.8}{n^{3/10}} u\right)$
- Jump magnitudes: $|\Delta_{jk}| \sim \mathcal{U}\left[\frac{C_\Delta}{2}\sqrt{\frac{\log(n)}{\delta_{jk}}}, C_\Delta\sqrt{\frac{\log(n)}{\delta_{jk}}}\right]$
- Signal strengths: $C_\Delta \in \{2, 5, 7\}$
- Contaminating noise: $\zeta_{jt} \sim \mathcal{N}(0, 1)$

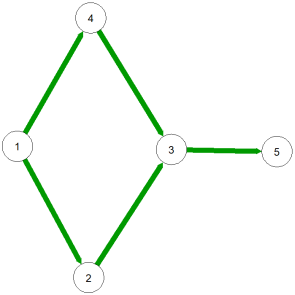



Figure 6: Local independence graph of process generating change points, with edges weighted by branching coefficients $\int g_{ij}$

# Simulation Study

We simulate 100 sample paths of length $n \in \{500, 1000\}$ from piecewise-constant mean model driven by a Hawkes processes with exponential kernels

- Base-line intensities: $\mu_j = n^{-7/10}$
- Kernels when non-zero: $g_{ij}(u) = \frac{0.6}{n^{3/10}} \exp\left(-\frac{0.8}{n^{3/10}} u\right)$
- Jump magnitudes: $|\Delta_{jk}| \sim \mathcal{U}\left[\frac{C_\Delta}{2}\sqrt{\frac{\log(n)}{\delta_{jk}}}, C_\Delta\sqrt{\frac{\log(n)}{\delta_{jk}}}\right]$
- Signal strengths: $C_\Delta \in \{2, 5, 7\}$
- Contaminating noise: $\zeta_{jt} \sim \mathcal{N}(0, 1)$

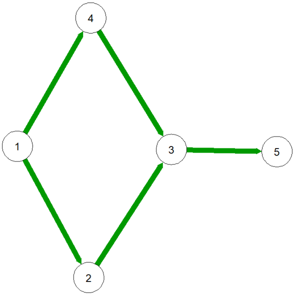



Figure 6: Local independence graph of process generating change points, with edges weighted by branching coefficients $\int g_{ij}$

# Simulation Study

We simulate 100 sample paths of length $n \in \{500, 1000\}$ from piecewise-constant mean model driven by a Hawkes processes with exponential kernels

- Base-line intensities:
  $\mu_j = n^{-7/10}$
- Kernels when non-zero:
  $g_{ij}(u) = \frac{0.6}{n^{3/10}} \exp\left(-\frac{0.8}{n^{3/10}} u\right)$
- Jump magnitudes:
  $|\Delta_{jk}| \sim \mathcal{U}\left[\frac{C_\Delta}{2}\sqrt{\frac{\log(n)}{\delta_{jk}}}, C_\Delta\sqrt{\frac{\log(n)}{\delta_{jk}}}\right]$
- Signal strengths:
  $C_\Delta \in \{2, 5, 7\}$
- Contaminating noise:
  $\zeta_{jt} \sim \mathcal{N}(0, 1)$

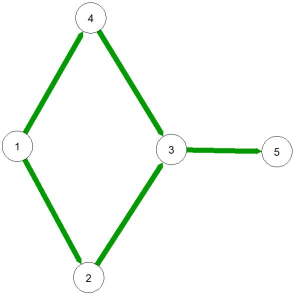



Figure 6: Local independence graph of process generating change points, with edges weighted by branching coefficients $\int g_{ij}$

# Simulation Study

We simulate 100 sample paths of length $n \in \{500, 1000\}$ from piecewise-constant mean model driven by a Hawkes processes with exponential kernels

- Base-line intensities: $\mu_j = n^{-7/10}$
- Kernels when non-zero: $g_{ij}(u) = \frac{0.6}{n^{3/10}} \exp\left(-\frac{0.8}{n^{3/10}} u\right)$
- Jump magnitudes: $|\Delta_{jk}| \sim \mathcal{U}\left[\frac{C_\Delta}{2}\sqrt{\frac{\log(n)}{\delta_{jk}}}, C_\Delta\sqrt{\frac{\log(n)}{\delta_{jk}}}\right]$
- Signal strengths: $C_\Delta \in \{2, 5, 7\}$
- Contaminating noise: $\zeta_{jt} \sim \mathcal{N}(0, 1)$

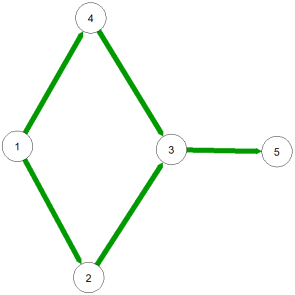



Figure 6: Local independence graph of process generating change points, with edges weighted by branching coefficients $\int g_{ij}$

# Simulation Study

We simulate 100 sample paths of length $n \in \{500, 1000\}$ from piecewise-constant mean model driven by a Hawkes processes with exponential kernels

- Base-line intensities:
  $\mu_j = n^{-7/10}$
- Kernels when non-zero:
  $g_{ij}(u) = \frac{0.6}{n^{3/10}} \exp\left(-\frac{0.8}{n^{3/10}} u\right)$
- Jump magnitudes:
  $|\Delta_{jk}| \sim \mathcal{U}\left[\frac{C_\Delta}{2}\sqrt{\frac{\log(n)}{\delta_{jk}}}, C_\Delta\sqrt{\frac{\log(n)}{\delta_{jk}}}\right]$
- Signal strengths:
  $C_\Delta \in \{2, 5, 7\}$
- Contaminating noise:
  $\zeta_{jt} \sim \mathcal{N}(0, 1)$

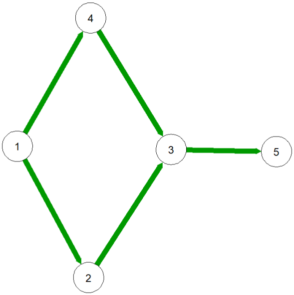
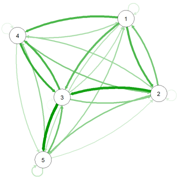


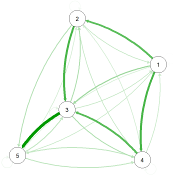

Figure 6: Local independence graph of process generating change points, with edges weighted by branching coefficients $\int g_{ij}$

# Simulation Study

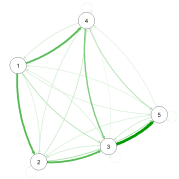
- We apply the our procedure with bin width $h = 1$, auto-regression order $\omega = \sqrt{n}$, and $\{\lambda_{ij}\}$ chosen so edge inclusion constitutes a test of size $\alpha = 0.1$



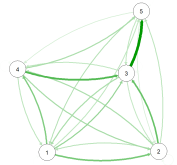
(a) $n = 500$ $C_\Delta = 2$



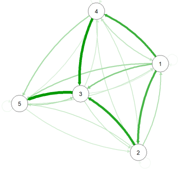
(b) $n = 500$ $C_\Delta = 5$



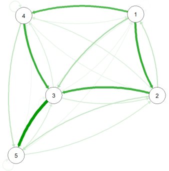
(c) $n = 500$ $C_\Delta = 7$



(d) $n = 1,000$ $C_\Delta = 2$



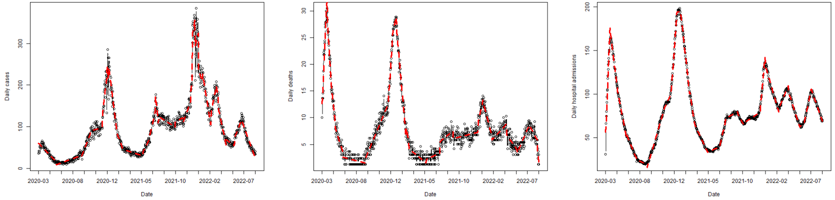
(e) $n = 1,000$ $C_\Delta = 5$



(f) $n = 1,000$ $C_\Delta = 7$

# Application 1: *changes in COVID-19 trajectories*

- We may have reason to suspect that changes in case numbers "cause" changes in hospitalizations, deaths, etc.




(a) Cases (b) Deaths (c) Hospitalizations

Figure 8: Anscombe transform of daily COVID-19 trajectories across all London boroughs along with piecewise linear trend (- - -) recovered using the NOT algorithm (Baranowski et al. 2019) equipped with seeded intervals (Kovács, Li & Bühlmann 2020)

# Application 1: *changes in COVID-19 trajectories*

- We apply our procedure with bin width $h = 1$, auto-regression order $\omega = 20$, and $\{\lambda_{ij}\}$ chosen so edge inclusion constitutes a test of size $\alpha = 0.1$

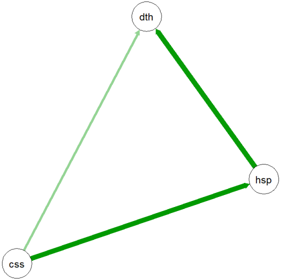
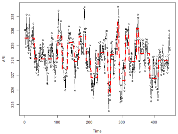


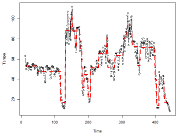

Figure 9: Estimated local independence graph with edges weighted by estimates branching coefficients $\int g_{ij}$

# Application 2: *cardiac response to live music performance*

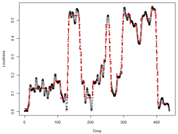
- **Goal:** we are interested in understanding the mechanism by which stress and strong emotions de-stabilises the heart's electrical pathways



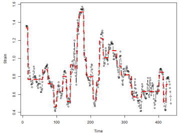

(a) ARI



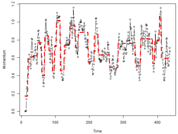

(b) tempo



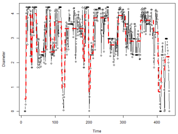

(c) loudness




(d) strain




(e) momentum




(f) diameter

# Application 2: *cardiac response to live music performance*

- We apply our procedure with bin width $h = 0.5$, auto-regression order $\omega = 20$, and $\{\lambda_{ij}\}$ chosen so edge inclusion constitutes a test of size $\alpha = 0.1$

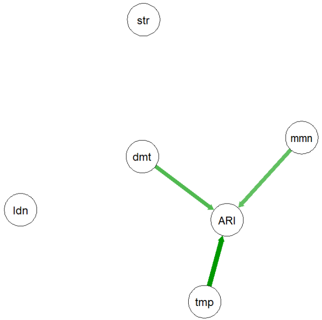



Figure 11: Estimated local independence graph with edges weighted by estimates branching coefficients $\int g_{ij}$

Thank you!