Temporal-topological properties of higher-order evolving networks
Supplementary Material
Alberto Ceria\textsuperscript{1,*} and Huijuan Wang\textsuperscript{1}

\textsuperscript{1}Faculty of Electrical Engineering, Mathematics, and Computer Science, Delft University of Technology, Mekelweg 4, 2628 CD, Delft, The Netherlands
\textsuperscript{*}A.Ceria@tudelft.nl

A General Statistics

Figure S1. Total number of events ($|\mathcal{E}_d|$) for each order $d$ in physical contact networks. Vertical axis is presented in logarithmic scale.
**Figure S2.** Total number of events ($| e_d |$) for each order $d$ in collaboration networks. Vertical and horizontal axes are presented in logarithmic scale.

**Figure S3.** Total number of hyperlinks ($| L_d |$) in the time aggregated higher-order network for each order $d$ for physical contact datasets. Vertical axis is presented in logarithmic scale.
Figure S4. Total number of hyperlinks ($|Z_d|$) for each order $d$ in collaboration networks. Vertical and horizontal axes are presented in logarithmic scale.
B Temporal-topological correlation of events

B.1 Correlation of temporal and topological distance of events

$\mu_d(\Delta t) = \frac{E[\eta(e,e')]}{E[\eta(e,e')]}$, between an order $d = 2$ event and an event of a different order, in each physical contact network and its corresponding three randomized null models $H_d^1$ (yellow), $H_d^2$ (green) and $H_d^3$ (red), which preserve or destroy specific properties of order $d = 3$ events.

$$\lim_{\Delta t \to \infty} E[\eta(e,e')]|T(e,e') < \Delta t, e \in \mathcal{E}_d, e' \in \mathcal{E}_d \setminus \mathcal{E}_d| = E[\eta(e,e')]| e \in \mathcal{E}_d, e' \in \mathcal{E}_d \setminus \mathcal{E}_d$$

for any $d$. The horizontal axes are presented in logarithmic scale. The dashed line in each figure corresponds to the linear fit (with slope $m$) of $\mu_d(\Delta t)$ as a function of $\log_{10}(\Delta t)$ in $H$, for the part that the curve has an increasing trend. For each dataset, the results of the three corresponding randomized models are obtained from 10 independent realizations.

Figure S5. The normalized average topological distance $\mu_d(\Delta t)$, between an order $d = 2$ event and an event of a different order, in each physical contact network and its corresponding three randomized null models $H_d^1$ (yellow), $H_d^2$ (green) and $H_d^3$ (red), which preserve or destroy specific properties of order $d = 3$ events. The dashed line in each figure corresponds to the linear fit (with slope $m$) of $\mu_d(\Delta t)$ as a function of $\log_{10}(\Delta t)$ in $H$, for the part that the curve has an increasing trend. For each dataset, the results of the three corresponding randomized models are obtained from 10 independent realizations.
Figure S6. The normalized average topological distance $\mu_d(\Delta t) = \frac{E[\eta(e,e')] T(e,e') < \Delta t, e \in \mathcal{E}_d, e' \in \mathcal{E} \setminus \mathcal{E}_d]}{E[\eta(e,e')] T(e,e') < \Delta t, e \in \mathcal{E}_d, e' \in \mathcal{E} \setminus \mathcal{E}_d}$, between an order $d = 2$ event and an event of a different order, in each collaboration network and its corresponding three randomized null models $\mathcal{H}_d^1$ (yellow), $\mathcal{H}_d^2$ (green) and $\mathcal{H}_d^3$ (red), which preserve or destroy specific properties of order $d = 2$ events. $\lim_{\Delta t \to \infty} E[\eta(e,e')] T(e,e') < \Delta t, e \in \mathcal{E}_d, e' \in \mathcal{E} \setminus \mathcal{E}_d] = E[\eta(e,e')] T(e,e') < \Delta t, e \in \mathcal{E}_d, e' \in \mathcal{E} \setminus \mathcal{E}_d]$ for any $d$. The horizontal axes are presented in logarithmic scale. The dashed line in each figure corresponds to the linear fit (with slope $m$) of $\mu_d(\Delta t)$ as a function of $\log_{10}(\Delta t)$ in $\mathcal{H}$, for the part that the curve has an increasing trend. For each dataset, the results of the three corresponding randomized models are obtained from 10 independent realizations.
Figure S7. The normalized average topological distance \( \mu_d(\Delta t) \) = \( \frac{E[\eta(e,e')|T(e,e')<\Delta t, e \in \mathcal{E}_d, e' \in \mathcal{E}\setminus\mathcal{E}_d]}{E[\eta(e,e')]|e \in \mathcal{E}_d, e' \in \mathcal{E}\setminus\mathcal{E}_d] \), between an order \( d = 4 \) event and an event of a different order, in each physical contact network and its corresponding three randomized null models \( \mathcal{H}_d \) (yellow), \( \mathcal{H}_d^1 \) (green) and \( \mathcal{H}_d^2 \) (red), which preserve or destroy specific properties of order \( d = 4 \) events. 

\[
\lim_{\Delta t \to \infty} E[\eta(e,e')|T(e,e')<\Delta t, e \in \mathcal{E}_d, e' \in \mathcal{E}\setminus\mathcal{E}_d] = E[\eta(e,e')]|e \in \mathcal{E}_d, e' \in \mathcal{E}\setminus\mathcal{E}_d \]

for any \( d \). The horizontal axes are presented in logarithmic scale. The dashed line in each figure corresponds to the linear fit (with slope \( m \)) of \( \mu_d(\Delta t) \) as a function of \( \log_{10}(\Delta t) \) in \( \mathcal{H}_d \), for the part that the curve has an increasing trend. For each dataset, the results of the three corresponding randomized models are obtained from 10 independent realizations.
Figure S8. The normalized average topological distance \( \mu_d(\Delta t) = \frac{\mathbb{E}[\eta(e,e')] | T(e,e') < \Delta t, e \in \mathcal{E}_d, e' \in \mathcal{E} \setminus \mathcal{E}_d]}{\mathbb{E}[\eta(e,e')] | e \in \mathcal{E}_d, e' \in \mathcal{E} \setminus \mathcal{E}_d} \), between an order \( d = 4 \) event and an event of a different order, in each collaboration network and its corresponding three randomized null models \( \mathcal{H}^1_d \) (yellow), \( \mathcal{H}^2_d \) (green) and \( \mathcal{H}^3_d \) (red), which preserve or destroy specific properties of order \( d = 3 \) events.

\[ \lim_{\Delta t \to \infty} \mathbb{E}[\eta(e,e')] | T(e,e') < \Delta t, e \in \mathcal{E}_d, e' \in \mathcal{E} \setminus \mathcal{E}_d] = \mathbb{E}[\eta(e,e')] | e \in \mathcal{E}_d, e' \in \mathcal{E} \setminus \mathcal{E}_d] \] for any \( d \). The horizontal axes are presented in logarithmic scale. The dashed line in each figure corresponds to the linear fit (with slope \( m \)) of \( \mu_d(\Delta t) \) as a function of \( \log_{10}(\Delta t) \) in \( \mathcal{H} \), for the part that the curve has an increasing trend. For each dataset, the results of the three corresponding randomized models are obtained from 10 independent realizations.
Figure S9. The normalized average topological distance $\nu_d(\Delta t) = \frac{E[\eta(e,e')|T(e,e')<\Delta t, e, e' \in \mathcal{E}_d]}{E[\eta(e,e')| e, e' \in \mathcal{E}_d]}$, between two order $d = 2$ events, in each physical contact network and its corresponding three randomized null models $\mathcal{H}^1_d$ (yellow), $\mathcal{H}^2_d$ (green) and $\mathcal{H}^3_d$ (red), which preserve or destroy specific properties of order $d = 2$ events. $\lim_{\Delta t \to \infty} E[\eta(e,e')|T(e,e')<\Delta t, e, e' \in \mathcal{E}_d] = E[\eta(e,e')| e, e' \in \mathcal{E}_d]$ for any $d$. The horizontal axes are presented in logarithmic scale. The dashed line in each figure corresponds to the linear fit (with slope $m$) of $\nu_d(\Delta t)$ as a function of $\log_{10}(\Delta t)$ in $\mathcal{H}$, for the part that the curve has an increasing trend. For each dataset, the results of the three corresponding randomized models are obtained from 10 independent realizations.
Figure S10. The normalized average topological distance \( \nu_d(\Delta t) = \frac{E[\eta(e, e')|\mathcal{T}(e, e') < \Delta t, e, e' \in \mathcal{E}_d]}{E[\eta(e, e')|e, e' \in \mathcal{E}_d]} \), between two order \( d = 2 \) events, in each collaboration network and its corresponding three randomized null models \( \mathcal{H}_d^1 \) (yellow), \( \mathcal{H}_d^2 \) (green) and \( \mathcal{H}_d^3 \) (red), which preserve or destroy specific properties of order \( d = 2 \) events. \( \lim_{\Delta t \to \infty} E[\eta(e, e')|\mathcal{T}(e, e') < \Delta t, e, e' \in \mathcal{E}_d] = E[\eta(e, e')|e, e' \in \mathcal{E}_d] \) for any \( d \). The horizontal axes are presented in logarithmic scale. The dashed line in each figure corresponds to the linear fit (with slope \( m \)) of \( \nu_d(\Delta t) \) as a function of \( \log_{10}(\Delta t) \) in \( \mathcal{H} \), for the part that the curve has an increasing trend. For each dataset, the results of the three corresponding randomized models are obtained from 10 independent realizations.
Figure S11. The normalized average topological distance \( \nu_d(\Delta t) = \frac{E[\eta(e,e')]|T(e,e')<\Delta t, e, e' \in E_d]}{E[\eta(e,e')]|e, e' \in E_d]} \), between two order \( d = 4 \) events, in each physical contact network and its corresponding three randomized null models \( \mathcal{H}_d^1 \) (yellow), \( \mathcal{H}_d^2 \) (green) and \( \mathcal{H}_d^3 \) (red), which preserve or destroy specific properties of order \( d = 4 \) events.

\[ \lim_{\Delta t \to \infty} E[\eta(e,e')|T(e,e')<\Delta t, e, e' \in E_d] = E[\eta(e,e')]|e, e' \in E_d] \] for any \( d \). The horizontal axes are presented in logarithmic scale. The dashed line in each figure corresponds to the linear fit (with slope \( m \)) of \( \nu_d(\Delta t) \) as a function of \( \log_{10}(\Delta t) \) in \( \mathcal{H}_d \), for the part that the curve has an increasing trend. For each dataset, the results of the three corresponding randomized models are obtained from 10 independent realizations.
Figure S12. The normalized average topological distance \( \nu_d(\Delta t) = \frac{E[\eta(e,e')|T(e,e')<\Delta t]}{E[\eta(e,e')]} \), between two order \( d = 4 \) events, in each collaboration network and its corresponding three randomized null models \( H_1^d \) (yellow), \( H_2^d \) (green) and \( H_3^d \) (red), which preserve or destroy specific properties of order \( d = 4 \) events. \( \lim_{\Delta t \to \infty} E[\eta(e,e')|T(e,e')<\Delta t] = E[\eta(e,e')] \) for any \( d \). The horizontal axes are presented in logarithmic scale. The dashed line in each figure corresponds to the linear fit (with slope \( m \)) of \( \nu_d(\Delta t) \) as a function of \( \log_{10}(\Delta t) \) in \( H_1^d \), for the part that the curve has an increasing trend. For each dataset, the results of the three corresponding randomized models are obtained from 10 independent realizations.
B.2 Topological correlation of events

Figure S13. The $d$-strength $s_d(v)$ versus the $d$-degree $k_d(v)$ of a node $v$ in the randomized reference model $H^2_d$ obtained from each real-world physical contact network, when $d$ is equal to 2 (blue dashed line), 3 (red dashed line) and 4 (green dashed line). The vertical axis is normalized by the average number $\omega_d$ of activations of a hyperlink of order $d$. The black dashed line represents the reference case $s_d(v) = \omega_d \times k_d(v)$. The error bar correspond to the standard deviation, centered in the mean value of 10 independent realizations of randomized reference model $H^2_d$. In total 30 linear bins are split for horizontal axis.
Figure S14. The $d$-strength $s_d(v)$ versus the the $d$-degree $k_d(v)$ of a node $v$ of the randomized reference model $H^2_d$ obtained from each real-world collaboration network, when $d$ is equal to 2 (blue dashed line), 3 (red dashed line) and 4 (green dashed line). The vertical axis is normalized by the average number $\omega_d$ of activations of a hyperlink of order $d$. The black dashed line represents the reference case $s_d(v) = \omega_d * k_d(v)$. The errorbar correspond to the standard deviation, centered in the mean value of 10 independent realizations of randomized reference model $H^2_d$. In total 30 linear bins are split for horizontal axis.
B.3 Temporal correlation of events at a local ego network

\[ \Delta t = 120 \text{ s} \]

**Figure S15.** Probability distribution \( \Pr(\mathcal{H}_s^* = s) \) of the size \( \mathcal{H}_s^* \) of trains (obtained from the activity series of egonetworks centered at each order 3 hyperlink), where a center link is activated at least once, in each physical contact network \( \mathcal{H} \) (blue) and its three randomized reference models \( \mathcal{H}_1^3 \) (yellow), \( \mathcal{H}_2^3 \) (green) and \( \mathcal{H}_3^3 \) (red). To identify the trains, we consider \( \Delta t = 120x \). For each network, the average size of the trains is reported. The maximum average size among network \( \mathcal{H} \), \( \mathcal{H}_1^3 \), \( \mathcal{H}_2^3 \) and \( \mathcal{H}_3^3 \) is in bold. The horizontal and vertical axes are presented in logarithmic scale.
\Delta t = 120 \text{ d}

Figure S16. Probability distribution $\Pr[S^* = s]$ of the size $S^*$ of trains (obtained from the activity series of egonetworks centered at each order 3 hyperlink), where a center link is activated at least once, in each collaboration network $\mathcal{H}$ (blue) and its three randomized reference models $\mathcal{H}^1_3$ (yellow), $\mathcal{H}^2_3$ (green) and $\mathcal{H}^3_3$ (red). To identify the trains, we consider $\Delta t = 120 d$. For each network, the average size of the trains is reported. The maximum average size among network $\mathcal{H}$, $\mathcal{H}^1_3$, $\mathcal{H}^2_3$ and $\mathcal{H}^3_3$ is in bold. The horizontal and vertical axes are presented in logarithmic scale.
Figure S17. Probability distribution $Pr[S^*_4 = s]$ of the size $S^*_4$ of trains (obtained from the activity series of egonetworks centered at each order 4 hyperlink), where a center link is activated at least once, in each physical contact network $H$ (blue) and its three randomized reference models $H^1_4$ (yellow), $H^2_4$ (green) and $H^3_4$ (red). To identify the trains, we consider $\Delta t = 60s$. For each network, the average size of the trains is reported. The maximum average size among network $H$, $H^1_4$, $H^2_4$ and $H^3_4$ is in bold. The horizontal and vertical axes are presented in logarithmic scale.
Figure S18. Probability distribution $\Pr[S^* = s]$ of the size $S^*_4$ of trains (obtained from the activity series of egonetworks centered at each order 4 hyperlink), where a center link is activated at least once, in each collaboration network $H$ (blue) and its three randomized reference models $H_1^4$ (yellow), $H_2^4$ (green) and $H_3^4$ (red). To identify the trains, we consider $\Delta t = 60d$. For each network, the average size of the trains is reported. The maximum average size among network $H$, $H_1^4$, $H_2^4$ and $H_3^4$ is in bold. The horizontal and vertical axes are presented in logarithmic scale.
Figure S19. Probability distribution $Pr[\mathcal{S}_4^* = s]$ of the size $\mathcal{S}_4^*$ of trains (obtained from the activity series of egonetworks centered at each order 4 hyperlink), where a center link is activated at least once, in each physical contact network $\mathcal{H}$ (blue) and its three randomized reference models $\mathcal{H}_1^4$ (yellow), $\mathcal{H}_2^4$ (green) and $\mathcal{H}_3^4$ (red). To identify the trains, we consider $\Delta t = 120$ s. For each network, the average size of the trains is reported. The maximum average size among network $\mathcal{H}$, $\mathcal{H}_1^4$, $\mathcal{H}_2^4$ and $\mathcal{H}_3^4$ is in bold. The horizontal and vertical axes are presented in logarithmic scale.
\[ Pr(S^* = s) \]

\[ \Delta t = 120 \text{ d} \]

**Figure S20.** Probability distribution \( Pr(S^* = s) \) of the size \( S^* \) of trains (obtained from the activity series of egonetworks centered at each order 4 hyperlink), where a center link is activated at least once, in each collaboration network \( H \) (blue) and its three randomized reference models \( H_1 \) (yellow), \( H_2 \) (green) and \( H_3 \) (red). To identify the trains, we consider \( \Delta t = 120d \). For each network, the average size of the trains is reported. The maximum average size among network \( H \), \( H_1 \), \( H_2 \) and \( H_3 \) is in bold. The horizontal and vertical axes are presented in logarithmic scale.
C Incomplete higher-order events

Figure S21. Total number of events ($|\mathcal{E}_d|$) in original network $\mathcal{H}$ and $\mathcal{H}_{\text{miss}}$ for each order $d$ in physical contact networks. Vertical axis is presented in logarithmic scale.
Figure S22. The normalized average topological distance $\mu_d(\Delta t) = \frac{E[\eta(e,e')] | \mathcal{F}(e,e') < \Delta t, e \in \delta_d, e' \in \delta \setminus \delta_d]}{E[\eta(e,e') | e \in \delta_d, e' \in \delta \setminus \delta_d]}$, between an order $d = 3$ event and an event of a different order, in each physical contact network $\mathcal{H}_{miss}$ and its corresponding three randomized null models $\mathcal{H}^1_{d, miss}$ (yellow), $\mathcal{H}^2_{d, miss}$ (green) and $\mathcal{H}^3_{d, miss}$ (red), which preserve or destroy specific properties of order $d = 3$ events. $\lim_{\Delta t \to \infty} E[\eta(e,e')] | \mathcal{F}(e,e') < \Delta t, e \in \delta_d, e' \in \delta \setminus \delta_d] = E[\eta(e,e')] | e \in \delta_d, e' \in \delta \setminus \delta_d]$ for any $d$. The horizontal axes are presented in logarithmic scale. The dashed line in each figure corresponds to the linear fit (with slope $m$) of $\mu_d(\Delta t)$ as a function of $\log_{10}(\Delta t)$ in $\mathcal{H}$, for the part that the curve has an increasing trend. For each dataset, the results of the three corresponding randomized models are obtained from 10 independent realizations.
The normalized average topological distance \( \mu_d(\Delta t) \) between an order \( d = 4 \) event and an event of a different order, in each physical contact network \( H_{\text{miss}} \) and its corresponding three randomized null models \( H_{\text{miss}}^1 \) (yellow), \( H_{\text{d,miss}}^2 \) (green) and \( H_{\text{d,miss}}^3 \) (red), which preserve or destroy specific properties of order \( d = 3 \) events.

\[
\lim_{\Delta t \to \infty} E[\eta(e,e') | \mathcal{F}(e,e') < \Delta t, e \in \mathcal{E}_d, e' \in \mathcal{E} \setminus \mathcal{E}_d] = E[\eta(e,e') | e \in \mathcal{E}_d, e' \in \mathcal{E} \setminus \mathcal{E}_d] \quad \text{for any } d.
\]

The horizontal axes are presented in logarithmic scale. The dashed line in each figure correspond to the linear fit (with slope \( m \)) of \( \mu_d(\Delta t) \) as a function of \( \log_{10}(\Delta t) \) in \( H \), for the part that the curve has an increasing trend. For each dataset, the results of the three corresponding randomized models are obtained from 10 independent realizations.

Figure S23.