

SOCIAL CONTAGION IN SCIENCE

S. Venturini · *Satyaki Sikdar* · F. Rinaldi · F. Tudisco · S. Fortunato

NetSci 2023

July 12, 2023

SCIENTIFIC COLLABORATIONS

- ◆ Exposure to *new* tools and theories
- ◆ Facilitates the *diffusion* of ideas
- ◆ Driven by *homophily* and *selection*

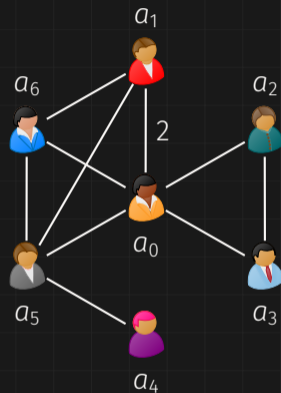
INTRODUCTION

SCIENTIFIC COLLABORATIONS

- ◆ Exposure to *new* tools and theories
- ◆ Facilitates the *diffusion* of ideas
- ◆ Driven by *homophily* and *selection*

COLLABORATION NETWORKS

- ◆ *Manifestations* of collaborations
- ◆ A *weighted* undirected network



Collaboration Network

TOPIC SWITCHES

TOPIC SWITCH

The act of a scholar a starting to work on a *new* topic t

TOPIC SWITCHES

TOPIC SWITCH

The act of a scholar a starting to work on a *new* topic t

MAIN OBJECTIVE

Study the *interplay* between collaborations and topic switches

TOPIC SWITCHES

TOPIC SWITCH

The act of a scholar *a* starting to work on a *new* topic *t*

MAIN OBJECTIVE

Study the *interplay* between collaborations and topic switches

CAUSE AND EFFECT

- ♦ Can only *measure* effects of collaborations on topic switches
- ♦ Cannot establish any *causal* relationship between the two

SETUP

Interaction Window (IW)

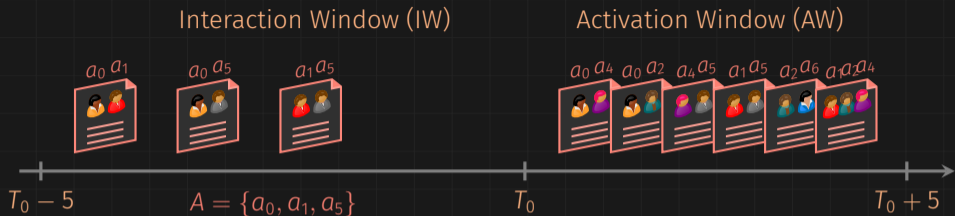
Activation Window (AW)



METHODOLOGY

- 1 Select a topic t , start year T_0 , and construct IW and AW

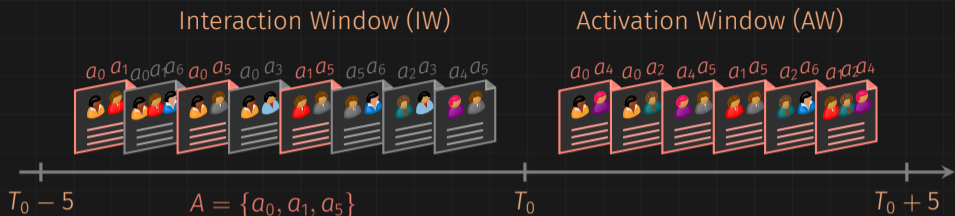
SETUP



METHODOLOGY

- 1 Select a topic t , start year T_0 , and construct IW and AW
- 2 Identify *active* authors A who publish on t during the IW $[T_0 - 5, T_0)$

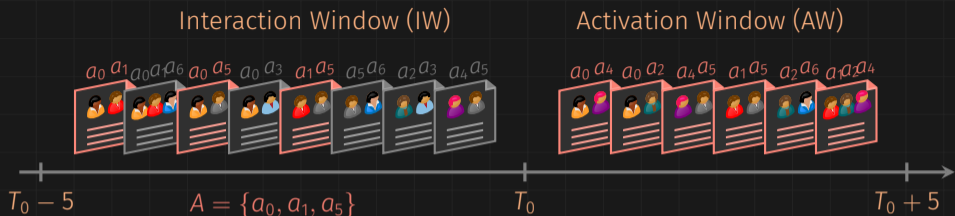
SETUP



METHODOLOGY

- 1 Select a topic t , start year T_0 , and construct IW and AW
- 2 Identify *active* authors A who publish on t during the IW $[T_0 - 5, T_0)$
- 3 Construct P : papers written by A during IW after becoming active

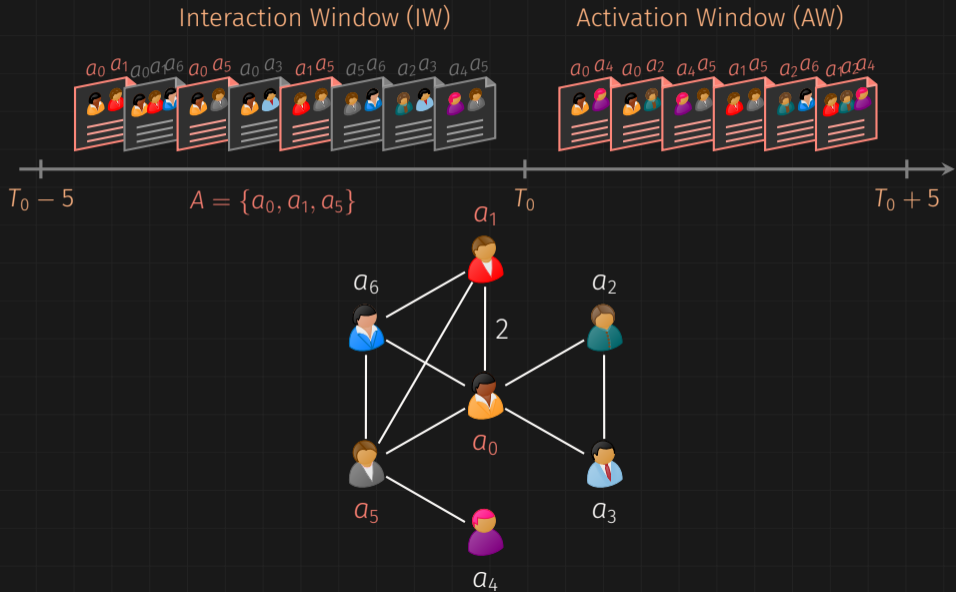
SETUP



METHODOLOGY

- 1 Select a topic t , start year T_0 , and construct IW and AW
- 2 Identify *active* authors A who publish on t during the IW $[T_0 - 5, T_0)$
- 3 Construct P : papers written by A during IW after becoming active
- 4 Build collaboration network G using P

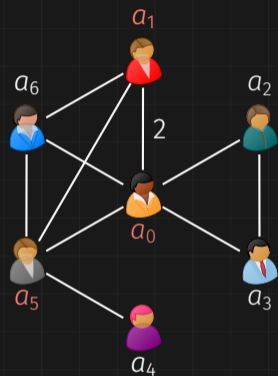
SETUP



EXPERIMENT I: MEMBERSHIP CLOSURE

CONTACTS WITH ACTIVE AUTHORS: k

Weighted degree in G wrt *active* neighbors

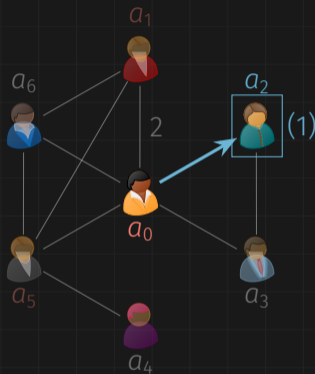


Collaboration Network G

EXPERIMENT I: MEMBERSHIP CLOSURE

CONTACTS WITH ACTIVE AUTHORS: k

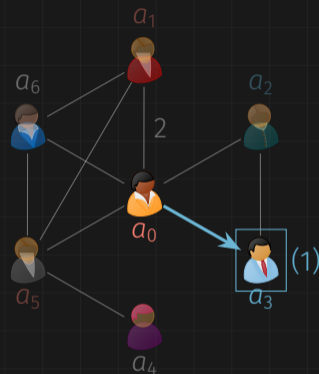
Weighted degree in G wrt *active* neighbors



EXPERIMENT I: MEMBERSHIP CLOSURE

CONTACTS WITH ACTIVE AUTHORS: k

Weighted degree in G wrt *active* neighbors

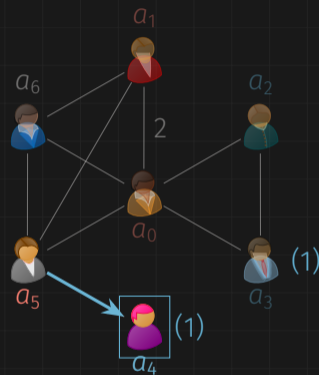


Collaboration Network G

EXPERIMENT I: MEMBERSHIP CLOSURE

CONTACTS WITH ACTIVE AUTHORS: k

Weighted degree in G wrt *active* neighbors

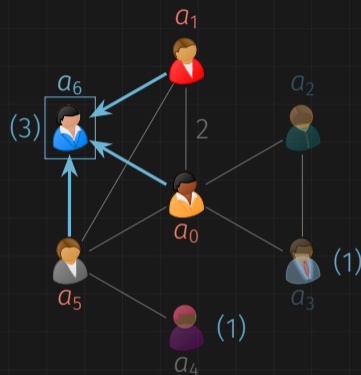


Collaboration Network G

EXPERIMENT I: MEMBERSHIP CLOSURE

CONTACTS WITH ACTIVE AUTHORS: k

Weighted degree in G wrt *active* neighbors



Collaboration Network G

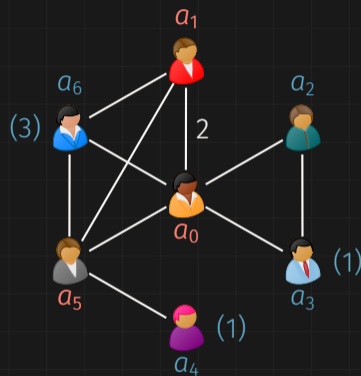
EXPERIMENT I: MEMBERSHIP CLOSURE

CONTACTS WITH ACTIVE AUTHORS: k

Weighted degree in G wrt *active* neighbors

MEMBERSHIP CLOSURE

Probability a becomes *active* in AW as a function of number of contacts, k



Collaboration Network G

EXPERIMENT I: MEMBERSHIP CLOSURE

CONTACTS WITH ACTIVE AUTHORS: k

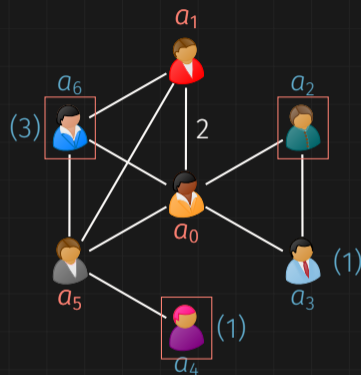
Weighted degree in G wrt *active* neighbors

MEMBERSHIP CLOSURE

Probability a becomes *active* in AW as a function of number of contacts, k

TARGET ACTIVATION PROBABILITY: $C^T(k)$

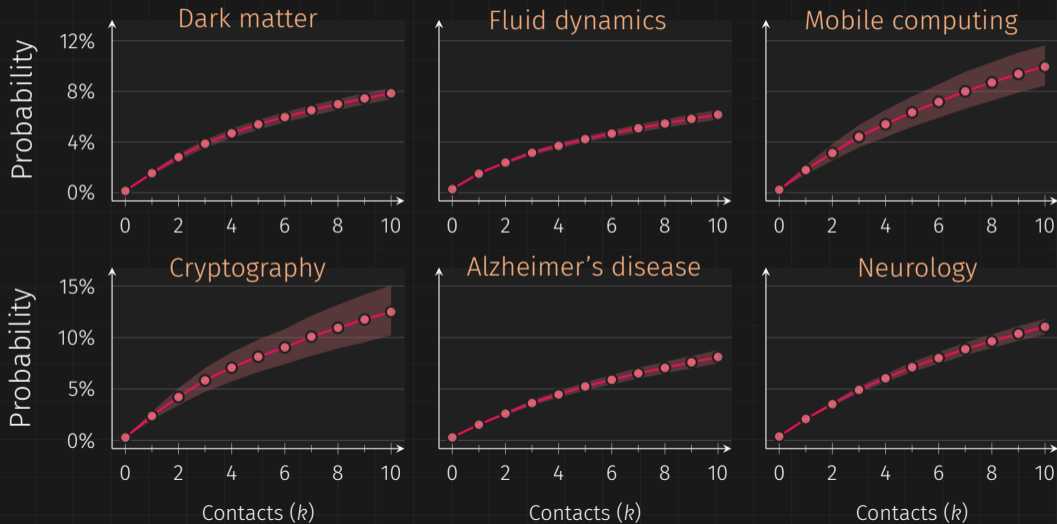
- ◆ Fraction of *inactive* authors who become *active* in AW with $\geq k$ contacts in IW
- ◆ $C^T(3) = \frac{1}{1} = 100\%$, $C^T(1) = \frac{3}{4} = 75\%$



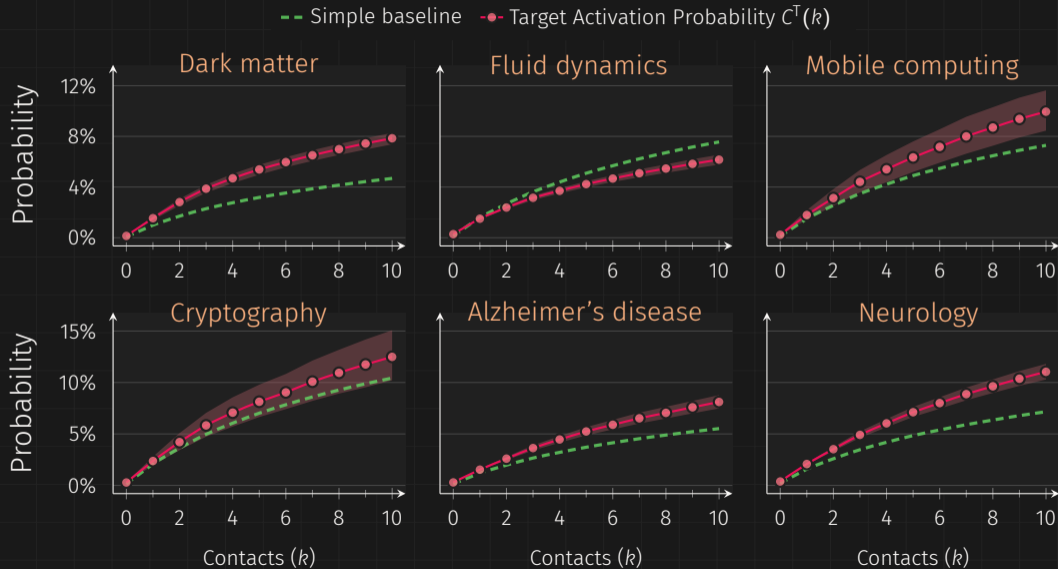
Collaboration Network G

EXPERIMENT I: TARGET ACTIVATION PROBABILITY

Simple baseline — Target Activation Probability $C^T(k)$



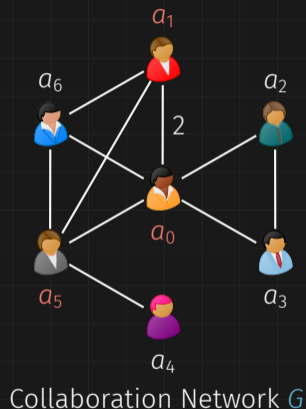
EXPERIMENT I: TARGET ACTIVATION PROBABILITY



EXPERIMENT II: SOURCE ACTIVATIONS

SOURCE ACTIVATION PROBABILITY P^S

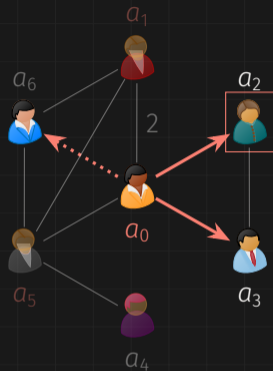
Fraction of a 's *exclusive inactive* coauthors who become *active* in AW



EXPERIMENT II: SOURCE ACTIVATIONS

SOURCE ACTIVATION PROBABILITY P^s

Fraction of a 's *exclusive inactive* coauthors who become *active* in AW



Collaboration Network G

$$P^s(a_0) = \frac{1}{2} = 50\%$$

$$P^s(a_1) = \text{NaN}, P^s(a_5) = 100\%$$

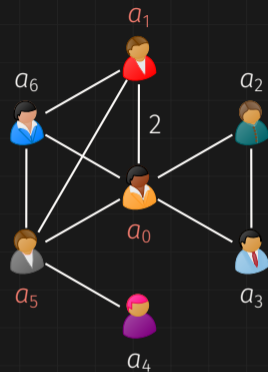
EXPERIMENT II: SOURCE ACTIVATIONS

SOURCE ACTIVATION PROBABILITY P^S

Fraction of a 's *exclusive inactive* coauthors who become *active* in AW

CHAPERONING PROPENSITY P^C

Fraction of a 's *exclusive inactive* coauthors who become *active and* write their first paper on t with a in AW



Collaboration Network G

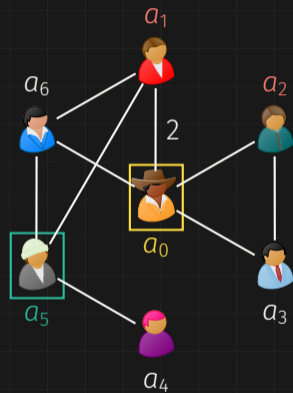
$$P^S(a_0) = \frac{1}{2} = 50\%$$

$$P^S(a_1) = \text{NaN}, P^S(a_5) = 100\%$$

EXPERIMENT II: CUMULATIVE SOURCE ACTIVATIONS

CUMULATIVE PROBABILITIES

$C^s(f)$: fraction of eligible *active* authors with $P^s \geq f$

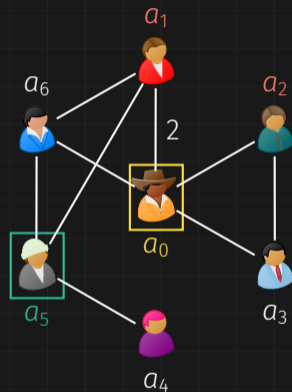


EXPERIMENT II: CUMULATIVE SOURCE ACTIVATIONS

CUMULATIVE PROBABILITIES

$C^s(f)$: fraction of eligible *active* authors with $P^s \geq f$

$C^c(f)$: fraction of eligible *active* authors with $P^c \geq f$



EXPERIMENT II: CUMULATIVE SOURCE ACTIVATIONS

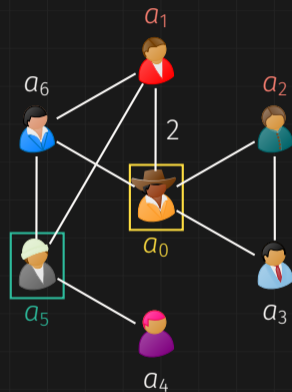
CUMULATIVE PROBABILITIES

$C^s(f)$: fraction of eligible *active* authors with $P^s \geq f$

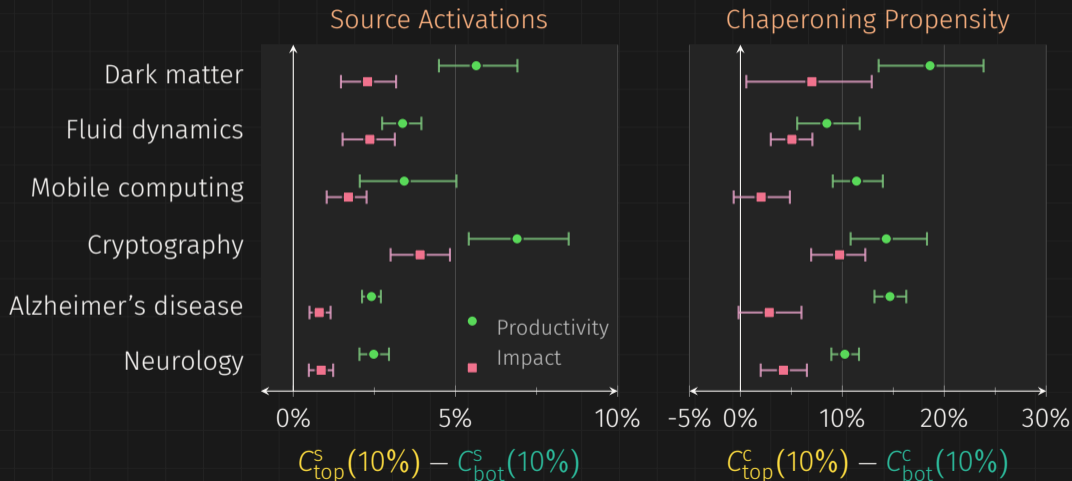
$C^c(f)$: fraction of eligible *active* authors with $P^c \geq f$

PROMINENT AUTHORS

- ◆ Mark **top 10%** and **bottom 10%** *active* authors A
 - Productivity
 - Impact
- ◆ Compute differences $C_{\text{top}}^s(f) - C_{\text{bot}}^s(f)$ and $C_{\text{top}}^c(f) - C_{\text{bot}}^c(f)$



EXPERIMENT II: CUMULATIVE PROBABILITIES



EXPERIMENT II: DILUTION EFFECT

STRENGTH OF CONNECTIONS

- ◆ Link between source activation probability and *team sizes*

EXPERIMENT II: DILUTION EFFECT

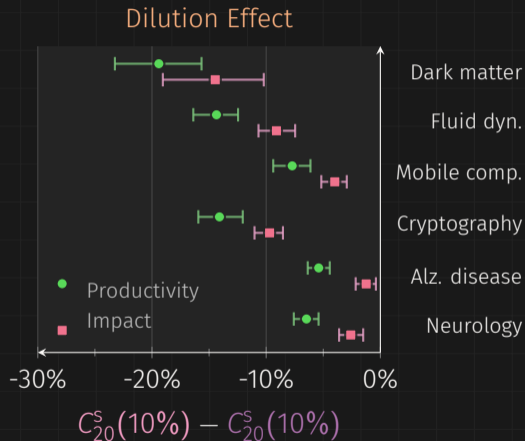
STRENGTH OF CONNECTIONS

- ◆ Link between source activation probability and *team sizes*
- ◆ Difference between *top 20%* and *bottom 20%* of top 10% *active* authors

EXPERIMENT II: DILUTION EFFECT

STRENGTH OF CONNECTIONS

- ◆ Link between source activation probability and *team sizes*
- ◆ Difference between *top 20%* and *bottom 20%* of top 10% *active* authors



CONCLUSIONS

MAIN FINDINGS

- ◆ A *complex* relationship between collaboration and topic switches

CONCLUSIONS

MAIN FINDINGS

- ◆ A *complex* relationship between collaboration and topic switches
- ◆ *Prominent* authors are more likely to induce topic switches

CONCLUSIONS

MAIN FINDINGS

- ◆ A *complex* relationship between collaboration and topic switches
- ◆ *Prominent* authors are more likely to induce topic switches
- ◆ Having too many collaborators *dilutes* the influence

CONCLUSIONS

MAIN FINDINGS

- ◆ A *complex* relationship between collaboration and topic switches
- ◆ *Prominent* authors are more likely to induce topic switches
- ◆ Having too many collaborators *dilutes* the influence

FUTURE DIRECTIONS

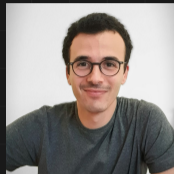
- ◆ Incorporate *institutional affiliations* of authors
- ◆ Relate topic switching probability with *semantic* similarity
- ◆ Model *higher-order* diffusion effects



Sara Venturini
University of Padova



Francesco Rinaldi
University of Padova



Francesco Tudisco
Gran Sasso Science Inst.



Santo Fortunato
Indiana University

Thanks!



[arXiv:2304.06826](https://arxiv.org/abs/2304.06826)



ssikdar@iu.edu



[@satyaki30](https://twitter.com/satyaki30)

DATA

<i>Topic</i>	<i># Windows</i>	<i>Interaction Window</i>		<i>Activation Window</i>	
		<i># Papers</i>	<i># Authors</i>	<i># Papers</i>	<i># Authors</i>
Dark matter	13	6,433	8,348	9,203	12,346
Fluid dynamics	16	5,290	11,950	7,231	16,960
Mobile computing	13	6,356	13,844	6,828	15,827
Cryptography	15	9,706	15,181	14,865	25,218
Alzheimer's disease	23	9,313	22,628	11,723	31,624
Neurology	23	9,260	26,046	12,795	39,515