Springer Proceedings in Complexity

Stefano Battiston
Francesco De Pellegrini
Guido Caldarelli
Emanuela Merelli *Editors*

Proceedings of ECCS 2014

European Conference on Complex Systems

Springer

Springer Proceedings in Complexity

Springer Complexity

Springer Complexity is an interdisciplinary program publishing the best research and academic-level teaching on both fundamental and applied aspects of complex systems—cutting across all traditional disciplines of the natural and life sciences, engineering, economics, medicine, neuroscience, social, and computer science.

Complex Systems are systems that comprise many interacting parts with the ability to generate a new quality of macroscopic collective behavior the manifestations of which are the spontaneous formation of distinctive temporal, spatial, or functional structures. Models of such systems can be successfully mapped onto quite diverse "real-life" situations like the climate, the coherent emission of light from lasers, chemical reaction–diffusion systems, biological cellular networks, the dynamics of stock markets and of the Internet, earthquake statistics and prediction, freeway traffic, the human brain, or the formation of opinions in social systems, to name just some of the popular applications.

Although their scope and methodologies overlap somewhat, one can distinguish the following main concepts and tools: self-organization, nonlinear dynamics, synergetics, turbulence, dynamical systems, catastrophes, instabilities, stochastic processes, chaos, graphs and networks, cellular automata, adaptive systems, genetic algorithms, and computational intelligence.

The three major book publication platforms of the Springer Complexity program are the monograph series "Understanding Complex Systems" focusing on the various applications of complexity, the "Springer Series in Synergetics", which is devoted to the quantitative theoretical and methodological foundations, and the "SpringerBriefs in Complexity" which are concise and topical working reports, case-studies, surveys, essays, and lecture notes of relevance to the field. In addition to the books in these two core series, the program also incorporates individual titles ranging from textbooks to major reference works.

More information about this series at http://www.springer.com/series/11637

Stefano Battiston · Francesco De Pellegrini
Guido Caldarelli · Emanuela Merelli
Editors

Proceedings of ECCS 2014

European Conference on Complex Systems

Springer

Editors
Stefano Battiston
RAF
University of Zürich
Zürich
Switzerland

Guido Caldarelli
IMT Lucca
Lucca
Italy

Francesco De Pellegrini
CREATE-NET
Trento
Italy

Emanuela Merelli
School of Science and Technology
University of Camerino
Camerino
Italy

ISSN 2213-8684 ISSN 2213-8692 (electronic)
Springer Proceedings in Complexity
ISBN 978-3-319-29226-7 ISBN 978-3-319-29228-1 (eBook)
DOI 10.1007/978-3-319-29228-1

Library of Congress Control Number: 2016934958

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG Switzerland

Foreword

In the past decade the field of Complexity Science has moved into a new stage of its life. The big data and information technology revolutions are finally providing the necessary data, numerical experiments and validation tests to the many conceptual and theoretical advances that complex systems science has already provided to a large number of scientific disciplines. These fast paced developments are augmenting complex systems science with an "applied" dimension. Our increasing capability to solve many open problems, in a large diversity of scientific fields, has made it possible that Complex Systems Science becomes one of the conceptual and methodological keys to understand and deal with important real-world challenges that range from epidemics and traffic congestions, to systemic risks and cultural evolution, to cite a few.

In this framework, it is no wonder that the Complex Systems Society, gathering all researchers engaged in complex systems research has grown and developed along the same lines. The general Society conference is annually gathering about 1,000 scientists from all disciplines and it is a meeting point where every scientist interested in complex systems research can network the collective with a vibrant research community.

The annual conference on Complex Systems of 2014, organized at the IMT School for advanced studies in Lucca, was a smashing success, breaking many records for attendance, number of presentations—more than 200—and parallel workshops. The Lucca conference is certainly a milestone in the life of the field and the Complex Systems Society. We are extremely glad to see that the chairmen of the conference Guido Caldarelli and Stefano Battiston—Chairmen of the Lucca's conference—have teamed up with Francesco De Pellegrini and Emanuela Merelli to edit a book that collects a selection of 27 papers presented at the conference. The final result is a proceedings volume that is truly representative of the wide range of problems addressed by the community and the depth of the technical approaches used to tackle them. It speaks loudly for itself and we are sure that it will also become a reference for those that want to grasp what the community is doing nowadays.

On behalf of the Complex Systems Society and its members we thank the organizers of the conference and the editors of this Proceedings of ECCS 2014 for all their work, the exemplary engagement and their service to Complex Systems Science.

Alessandro Vespignani
President of the Complex Systems Society 2012–2015
Boston, MA, USA

Yamir Moreno
President of the Complex Systems Society 2016–2019
Zaragoza, Spain

Preface

This volume collects a series of multidisciplinary contributions in the field of complex systems science. Several works presented in this collection pivot on the theory and applications of formal and computational approaches. These methods are suitable to construct and simulate models of complex systems so as to analyse their properties. This is indeed an emerging research area encompassing a broad range of fields including—but not limited to—physics, computer science and mathematics, economics, business, political science, biology, sociology, neuroscience and medicine. The collection is thus addressed to the new generation of transdisciplinary researchers.

The work contains contributions which have been initially discussed in the *European Conference on Complex Systems* (ECCS'14) held at IMT, Lucca from 22 to 26 September 2014, under the sponsorship of the Complex Systems Society. ECCS'14 is a major international conference in the area of Complex Systems and interdisciplinary science in general. The main aim is to offer unique opportunities to study novel foundational approaches in a multitude of application areas. Thus, it spans from Complexity in ICT and Social Systems, to Complexity in Infrastructures, Complexity in Environment and Cities, Complexity in Natural Sciences, Complexity in Humanities, Linguistics and Society Complexity in Economics and Finance.

The project had an internal call for papers presented at the ECCS14 Conference. It contains a selection of 27 papers which originated from the conference oral presentations and poster sessions. All the manuscripts are extended versions of the contributions presented there and went through an independent review process.

The editors express their thanks to all authors of the articles submitted to this special issue. They also acknowledge the efforts of our many reviewers for their help in selecting the papers published in this special issue.

Zürich, Switzerland — Stefano Battiston
Trento, Italy — Francesco De Pellegrini
Lucca, Italy — Guido Caldarelli
Camerino, Italy — Emanuela Merelli

Contents

Contributors

Madalena Alarcão Faculty of Psychology and Education Sciences, University of Coimbra, Coimbra, Portugal

Talayeh Aledavood Aalto University School of Science, Espoo, Finland

Hideaki Aoyama Graduate School of Sciences, Kyoto University, Kyoto, Japan

Edmund Barter Department of Engineering Mathematics, University of Bristol, University Walk, Bristol, UK

Timothy Bates Physics Department, Assumption College, Worcester, MA, USA

Stefano Battiston Department of Banking and Finance, University of Zurich, Zürich, Switzerland

Igor Borovikov Nekkar.net: International Labs, Foster City, CA, USA

Dmitry Bratsun Theoretical Physics Department, Perm State Pedagogical University, Perm, Russia

Markus Brede Institute for Complex Systems Simulation, University of Southampton, Southampton, UK

F. Caravelli Invenia Technical Computing, Winnipeg, Canada; Department of Computer Science, UCL, London, UK; London Institute of Mathematical Sciences, London, UK

T. Carletti Department of Mathematics and naXys, University of Namur, Namur, Belgium

Alexander Casey Physics Department, Assumption College, Worcester, MA, USA; University of Notre Dame, Notre Dame, IN, USA

Filippo Castiglione Institute for Applied Mathematics (IAC) CNR, Rome, Italy

A. Chakraborti School of Computational and Integrative Sciences (SCIS), Jawaharlal Nehru University, New Delhi, India

Anna Chernyshova Siberian Federal University, Krasnoyarsk, Russia

Michael Daly Physics Department, Assumption College, Worcester, MA, USA; Meditech, Framingham, MA, USA

Ana Teixeira de Melo Faculty of Psychology and Education Sciences of the University of Coimbra, Coimbra, Portugal

Antonio De Nicola ENEA-CR Casaccia, Rome, Italy; University of Rome Tor Vergata, Rome, Italy

Francesco De Pellegrini CREATE-NET, Trento, Italy

Alonso Espinosa Mireles de Villafranca Department of Engineering Mathematics, University of Bristol, University Walk, Bristol, UK

Lei Dong School of Architecture, Tsinghua University, Beijing, China

Robin I.M. Dunbar Department of Experimental Psychology, University of Oxford, Oxford, UK

Gregorio D'Agostino ENEA-CR Casaccia, Rome, Italy; Center for Polymer Studies, Boston University, Boston, MA, USA

Yoshi Fujiwara Graduate School of Simulation Studies, University of Hyogo, Kobe, Japan

Georgi Yordanov Georgiev Physics Department, Assumption College, Worcester, MA, USA; Physics Department, Tufts University, Medford, MA, USA; Department of Physics, Worcester Polytechnic Institute, Worcester, MA, USA

Erin Gombos Physics Department, Assumption College, Worcester, MA, USA; National Cancer Institute, Bethesda, MD, USA

E. Heinsalu National Institute of Chemical Physics and Biophysics (NICPB), Tallinn, Estonia

Kaitlin Henry Physics Department, Assumption College, Worcester, MA, USA

Jorge Hidalgo Departamento de Electromagnetismo y Física de la Materia e Instituto Carlos I de Física Teórica y Computacional, Universidad de Granada, Granada, Spain

Masafumi Hino NEC Corporation, Minato-ku, Tokyo, Japan

Masato Hisakado Financial Services Agency, Chiyoda-ku, Tokyo, Japan

Johanne Hizanidis National Center for Scientific Research "Demokritos", Athens, Greece; Crete Center for Quantum Complexity and Nanotechnology, Department of Physics, University of Crete, Heraklion, Greece

Philipp Hövel Institut für Theoretische Physik, Technische Universität Berlin, Berlin, Germany; Bernstein Center for Computational Neuroscience Berlin, Humboldt-Universität zu Berlin, Berlin, Germany

Hiroshi Iyetomi Department of Mathematics, Niigata University, Niigata, Japan

Mark P. Johnson Ryan Institute, National University of Ireland Galway, Galway, Ireland

Andreas Karpf Université Paris 1 Panthéon-Sorbonne, Centre d'Économie de la Sorbonne/Paris School of Economics, Paris, France

K. Kaski Department of Computer Science, Aalto University School of Science, Aalto, Finland

Espen Knoop Department of Engineering Mathematics, University of Bristol, University Walk, Bristol, UK; Bristol Robotics Laboratory, Bristol, UK

Ruiqi Li School of Systems Science, Beijing Normal University, Beijing, China

Eduardo López CABDyN Complexity Center, Saïd Business School, University of Oxford, Oxford, UK

Antoine Mandel Université Paris 1 Panthéon-Sorbonne, Centre d'Économie de la Sorbonne, Paris, France

Dimitris Maniadakis Department of Informatics and Telecommunications, University of Athens, Athens, Greece

Ramona Marguta Departamento de Física, Biosystems Integrative Sciences Institute (BioISI), Faculdade de Ciências da Universidade de Lisboa, Campo Grande, Lisbon, Portugal

L. Marzola National Institute of Chemical Physics and Biophysics (NICPB), Tallinn, Estonia; Laboratory of Theoretical Physics, Institute of Physics, University of Tartu, Tartu, Estonia

Antoni Matyjaszkiewicz Department of Engineering Mathematics, University of Bristol, University Walk, Bristol, UK

Christopher McWilliams Department of Engineering Mathematics, University of Bristol, University Walk, Bristol, UK

F. Medda QASER Lab, UCL, London, UK

Emanuela Merelli School of Science and Technology, University of Camerino, Camerino, Italy

Daniele Miorandi CREATE-NET, Trento, Italy

Paolo Moretti Institute of Materials Simulation (WW8), Friedrich-Alexander-University, Erlangen-Nünberg, Fürth, Germany

Shintaro Mori Department of Physics, Kitasato University, Sagamihara, Kanagawa, Japan

Esteban Moro Departamento de Matemáticas & GISC, Universidad Carlos III de Madrid, Leganés, Spain

Igor Mozetič Jožef Stefan Institute, Ljubljana, Slovenia

Miguel A. Muñoz Departamento de Electromagnetismo y Física de la Materia e Instituto Carlos I de Física Teórica y Computacional, Universidad de Granada, Granada, Spain

James T. Murphy Sorbonne Universités, UPMC Univ Paris 6, UMR 7144, Station Biologique de Roscoff, Roscoff, France; CNRS, UMR 7144, Equipe Div&Co, Station Biologique de Roscoff, Roscoff, France; Ryan Institute, National University of Ireland Galway, Galway, Ireland

D. Nicolay Department of Mathematics and naXys, University of Namur, Namur, Belgium

Iryna Omelchenko Institut für Theoretische Physik, Technische Universität Berlin, Berlin, Germany

Andrea Parisi Departamento de Física, Biosystems Integrative Sciences Institute (BioISI), Faculdade de Ciências da Universidade de Lisboa, Campo Grande, Lisbon, Portugal

M. Patriarca National Institute of Chemical Physics and Biophysics (NICPB), Tallinn, Estonia

Marco Pettini Centre de Physique Théorique, Aix-Marseille University, Marseille, France

Thong Pham Osaka University, Osaka, Japan

Felix Reed-Tsochas CABDyN Complexity Center, Saïd Business School, University of Oxford, Oxford, UK; Department of Sociology, University of Oxford, Oxford, UK

R. Reinanda Faculty of Science, University of Amsterdam, Amsterdam, The Netherlands

Sam G.B. Roberts Department of Psychology, University of Chester, Chester, UK

Lewis Roberts Department of Engineering Mathematics, University of Bristol, University Walk, Bristol, UK

A. Roli Department of Computer Science and Engineering (DISI), University of Bologna, Campus of Cesena, Bologna, Italy

Matteo Rucco School of Science and Technology, University of Camerino, Camerino, Italy

Michael Sadovsky Institute of Computational Modelling SB RAS, Krasnoyarsk, Russia

Jari Saramäki Aalto University School of Science, Espoo, Finland

Luís F. Seoane ICREA-Complex Systems Lab, Universitat Pompeu Fabra-PRBB, Barcelona, Spain; Institut de Biologia Evolutiva, CSIC-UPF, Barcelona, Spain

Paul Sheridan The University of Tokyo, Tokyo, Japan

Hidetoshi Shimodaira Osaka University, Osaka, Japan

Borut Sluban Jožef Stefan Institute, Ljubljana, Slovenia

Jasmina Smailović Jožef Stefan Institute, Ljubljana, Slovenia

Ricard Solé ICREA-Complex Systems Lab, Universitat Pompeu Fabra-PRBB, Barcelona, Spain; Institut de Biologia Evolutiva, CSIC-UPF, Barcelona, Spain; Santa Fe Institute, Santa Fe, NM, USA

Massimo Stella Institute for Complex Systems Simulation, University of Southampton, Southampton, UK

Taiki Takahashi Department of Behavioral Science, Faculty of Letters, Hokkaido University, Sapporo, Japan; Center for Experimental Research in Social Sciences, Hokkaido University, Sapporo, Hokkaido, Japan

V.A. Traag CWTS, Leiden University, Leiden, The Netherlands

G. van Klinken KITLV, Leiden, The Netherlands

Dimitris Varoutas Department of Informatics and Telecommunications, University of Athens, Athens, Greece

Frédérique Viard Sorbonne Universités, UPMC Univ Paris 6, UMR 7144, Station Biologique de Roscoff, Roscoff, France; CNRS, UMR 7144, Equipe Div&Co, Station Biologique de Roscoff, Roscoff, France

Pablo Villegas Departamento de Electromagnetismo y Física de la Materia e Instituto Carlos I de Física Teórica y Computacional, Universidad de Granada, Granada, Spain

Andrea Vüllings Institut für Theoretische Physik, Technische Universität Berlin, Berlin, Germany

Xinran Wang College of Resources Science and Technology, Beijing Normal University, Beijing, China

Hiroshi Yoshikawa Graduate School of Economics, The University of Tokyo, Tokyo, Japan

Andrey Zakharov Theoretical Physics Department, Perm State Pedagogical University, Perm, Russia

Mattia Zeni Department of Information Engineering and Computer Science, University of Trento, Trento, Italy

Jiang Zhang School of Systems Science, Beijing Normal University, Beijing, China

Chapter 1
Detection of Non-self-correcting Nature of Information Cascade

Shintaro Mori, Masafumi Hino, Masato Hisakado and Taiki Takahashi

Abstract We propose a method of detecting non-self-correcting information cascades in experiments in which subjects choose an option sequentially by observing the choices of previous subjects. The method uses the correlation function $C(t)$ between the first and the $t+1$th subject's choices. $C(t)$ measures the strength of the domino effect, and the limit value $c \equiv \lim_{t\to\infty} C(t)$ determines whether the domino effect lasts forever ($c > 0$) or not ($c = 0$). The condition $c > 0$ is an adequate condition for a non-self-correcting system, and the probability that the majority's choice remains wrong in the limit $t \to \infty$ is positive. We apply the method to data from two experiments in which T subjects answered two-choice questions: (i) general knowledge questions ($T_{avg} = 60$) and (ii) urn-choice questions ($T = 63$). We find $c > 0$ for difficult questions in (i) and all cases in (ii), and the systems are not self-correcting.

1.1 Introduction

Herding phenomena are ubiquitous in human and animal behavior [1, 2]. An example is an information cascade, in which a person observes others' choices and chooses the majority's choice even though the person's private signal contradicts it [3, 4]. It is a rational behavior for people who are uncertain about choosing. If an information

S. Mori (✉)
Department of Physics, Kitasato University, 1-15-1 Kitasato, Sagamihara, Kanagawa 252-0373, Japan
e-mail: mori@sci.kitasato-u.ac.jp

M. Hino
NEC Corporation, Siba 5-7-1, Minato-ku, Tokyo 108-8001, Japan

M. Hisakado
Financial Services Agency, Kasumigaseki 3-2-1, Chiyoda-ku, Tokyo 100-8967, Japan

T. Takahashi
Department of Behavioral Science, Faculty of Letters, Hokkaido University, Sapporo, Japan

T. Takahashi
Center for Experimental Research in Social Sciences, Hokkaido University, Kita 10, Nishi 7, Kita-ku, Sapporo, Hokkaido 060-0810, Japan

S. Battiston et al. (eds.), *Proceedings of ECCS 2014*, Springer Proceedings in Complexity, DOI 10.1007/978-3-319-29228-1_1

cascade occurs, the same mechanism applies to later decision-makers, and the majority's choice tends to prevail. In some cases, the successive choices are wrong, and the cascade leads to irrational herding behavior [5].

An experimental setup demonstrates a situation in which an information cascade occurs [6]. There are two urns, A and B, and urn A (B) contains two a (b) balls and one b (a) ball. In each run of the experiment, an urn is randomly chosen initially and called X. Then, the subjects guess whether urn X is A or B and choose sequentially. They get a reward for the correct choice. In the course of the experiment, each subject draws a ball from X, which is his private signal. If the ball is a (b), urn X is more likely to be A (B). He also observes the choices of the previous subjects. If the difference between the numbers of subjects who choose each urn exceeds two, the private signal cannot overcome the majority's choice. An information cascade starts if someone chooses the majority's choice although his private signal suggests the minority's one. As the probability that the first two persons both choose the wrong option is non-zero, the probability for the onset of a cascade where the majority's choice is wrong is positive.

We now consider whether the wrong cascade continues [5]. If it continues forever, the majority's choice converges to the wrong option. Information cascades were initially considered to be fragile phenomena. As the trigger of the cascade is a small imbalance, people can be dissuaded from following the majority's choice [3]. In addition, an agent model with a Bayesian update of the private belief showed that the information cascade is self-correcting [8]. As the number of agents tends toward infinity, the wrong cascade disappears, and the majority's choice converges to the optimal option.

Using an information cascade experiment with a general knowledge two-choice quiz, we have shown that a phase transition occurs between a one-peak phase and a two-peak phase [9]. If the questions are easy, the ratio $z(t)$ of the correct choices of t subjects converges to a value $z_+ > 1/2$ in the limit $t \to \infty$. As there is only one peak in the probability distribution function of $z(t)$, we call the corresponding phase the one-peak phase [10, 11]. If the questions are difficult and most people do not know the answers, $z(t)$ converges to $z_+ > 1/2$ or $z_- < 1/2$. One cannot predict the value in $\{z_+, z_-\}$ to which $z(t)$ converges. We call the corresponding phase the two-peak phase. In the two-peak phase, the wrong cascade does not necessarily disappear, and the system is not self-correcting.

It was recently shown that the limit value of the normalized correlation function is the order parameter of the phase transition [14]. The normalized correlation function shows how the first subject's choice propagates to later subjects. It provides a measure of the domino effect. In addition, the positiveness of the limit value is a sufficient condition for a non-self-correcting system. By extrapolating the results for a finite system to infinity, we can determine whether the system is self-correcting. We report on the application of the method to data from two types of information cascade experiments. In Sect. 1.2, we define the normalized correlation function. We also explain the behavior of the function in each phase and the extrapolation method used to estimate its limit. We present the results of the data analysis in Sect. 1.3. Section 1.4 summarizes the results.

1.2 Correlation Function and Asymptotic Behaviors

We consider a typical information cascade experiment. T subjects answer a two-choice question sequentially in each run. We denote the order of the subjects as t, where $t = 1, 2, \ldots, T$. We denote the choice of subject t by $X(t) \in \{0, 1\}, t = 1, 2, \ldots, T$. If the choice is true (false), $X(t)$ takes 1 (0).

The correlation function $C(t)$ is defined as the covariance between $X(1)$ and $X(t+1)$ divided by the variance of $X(1)$:

$$C(t) \equiv \mathrm{Cov}(X(1), X(t+1))/\mathrm{Var}(X(1)).$$

$C(t)$ can be expressed as the difference of two conditional probabilities.

$$C(t) = \Pr(X(t+1) = 1|X(1) = 1) - \Pr(X(t+1) = 1|X(1) = 0). \tag{1.1}$$

$C(t)$ shows the degree to which the first subject's choice is transmitted to later subjects. It is a measure of the domino effect in an information cascade.

$C(t)$ is generally positive, and its asymptotic behavior depends on the phase of the system and the shape of the response function $q(z)$. Here $q(z)$ represents the dependence of the probability of the correct choice by subject $t+1$ on the ratio $z(t)$ of the correct choices of the previous t subjects.

$$q(z) \equiv \Pr(X(t+1) = 1|z(t) = z), \quad z(t) = \frac{1}{t}\sum_{s=1}^{t} X(s).$$

With the definition of $q(z)$, the stochastic process $\{X(t)\}, t = 1, 2 \ldots$ becomes a generalized Pólya urn process [12]. If there is one solution for $z = q(z)$ at z_+ (left panel in Fig. 1.1), $z(t)$ converges to z_+. $C(t)$ shows power-law decay for large t with two constants, c' and l, as

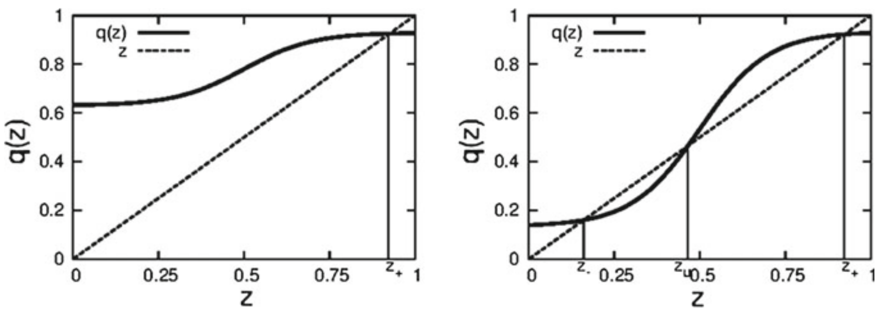

Fig. 1.1 Response function $q(z)$ versus z. *Left* panel shows the one-peak phase, in which there is one solution, z_+, for $z = q(z)$. *Right* panel shows the two-peak phase, in which there are three solutions, $z_- < z_u < z_+$, for $z = q(z)$

$$C(t) \simeq c' \cdot t^{l-1} \quad l < 1.$$

Here, l is the exponent for the power-law decay and is less than 1. The value of l is given by $g'(z_+)$ [11, 13]. If there are three solutions for $z = q(z)$ at $z_- < z_u < z_+$ (right panel in Fig. 1.1), the system is in the two-peak phase; $\lim_{t\to\infty} z(t) = z_+$ or z_- [12]. The limit value $c \equiv \lim_{t\to\infty} C(t)$ is positive, and the first subject's choice propagates to an infinite number of later subjects [14]. $C(t)$ behaves asymptotically as

$$C(t) \sim c + c' \cdot t^{l-1}. \tag{1.2}$$

Here $c' \cdot t^{l-1}$ is the subleading term of $C(t)$, and l is given by the larger value among $\{g'(z_+), g'(z-)\}$. Further, c acts as an order parameter of the phase transition, and (1.2) is the general asymptotic behavior of $C(t)$ [15].

As it is difficult to estimate c using $c \equiv \lim_{t\to\infty} C(t)$ with empirical data, where the system size and number of samples are strictly limited, we introduce two quantities for the estimation. First, we define the nth moment $m_n(t)$ for $C(t)$ as $m_n(t) \equiv \sum_{s=0}^{t-1} C(s)(s/t)^n$. We define the integrated correlation time $\tau(t)$ as $\tau(t) = m_0(t)$. We also define the second moment correlation time $\xi(t)$ as $\xi(t) \equiv t \cdot \sqrt{m_2(t)/m_0(t)}$. Using the asymptotic behavior of $C(t)$, we estimate the subsequent asymptotic behavior of $\tau(t)/t$ and $\xi(t)/t$.

$$\tau(t)/t \simeq c + \frac{c'}{l} \cdot t^{l-1} \tag{1.3}$$

$$\xi(t)/t \to \begin{cases} \sqrt{l/l+2} & c = 0 \\ \sqrt{1/3} & c > 0 \end{cases} \tag{1.4}$$

As $\tau(t)/t$ is defined as the summation of $C(s)$ over $0 \le s < t$ divided by t, the standard error becomes smaller than that of $C(t)$. The asymptotic behavior of $\tau(t)/t$ in (1.3) provides a more reliable estimate of c and l than the fitting of $C(t)$ to (1.2). $\xi(t)/t$ also provides a reliable estimate for l [15]. If $c > 0$, the leading term of $C(t)$ is the constant c, and l should be interpreted as $l = 1$.

We define whether the system is self-correcting according to whether $z(t)$ always converges to z_+. In the one-peak (two-peak) phase, the system is (non-)self-correcting. If $c > 0$, the system is in the two-peak phase and is non-self-correcting. However, $c = 0$ does not necessarily mean that the system is self-correcting. For the system to be self-correcting, $q(z) = z$ has to have only one solution, z_+.

1.3 Domino Effect and Detection of Non-self-correcting Nature

We study the domino effect and non-self-correction in information cascades. We discuss two types of information cascade experiments.

In experiment 1 (EXP-I), subjects answered a general knowledge two-choice quiz. First, the subjects answered using only their own knowledge. Then, they observed the choices of previous subjects and answered the question again. The average length of the sequence of subjects is $T = 60$, and the number of choice sequences is 240. The choice sequences are classified into four bins according to the ratio of correct choices $z_0(T)$ of the first answers without observation as $z_0(T) = 50 \pm 5, 60 \pm 5, 70 \pm 5$, and $80 \pm 5\,\%$, and the number of samples in each bin is $38(50 \pm 5\,\%), 52(60 \pm 5\,\%), 38(70 \pm 5\,\%)$, and $38(80 \pm 5\,\%)$, respectively [16].

Experiment 2 (EXP-II) is similar to the situation explained in the Introduction. There are two urns, A and B, which contain a and b balls in different configurations. We use two configuration patterns: (i) two a balls and one b ball in urn A versus one a ball and two b balls in urn B and (ii) five a balls and four b balls in urn A versus four a balls and five b balls in urn B. Urn $X \in \{\text{A,B}\}$ is chosen at random at the beginning of each run, and subjects are asked to choose between A or B. Each subject draws one ball from X and checks whether it is a or b. The ball corresponds to the type of urn X with probability $q = 2/3(5/9)$ for (i) [(ii)]. In addition, the subject also observes the choices of previous subjects. Our results, unlike those of previous experiments [6–8], show the summary statistics of the number of subjects who have chosen each urn. The length T and number of questions I are 63 and 200, respectively, for $q \in \{2/3, 5/9\}$ [17].

We denote the choice sequences in each bin as $\{X(i,t)\}, i = 1, \ldots, I, t = 1, \ldots, T(i)$. Here, the length of the sequence depends on question i in EXP-I; we denote it as $T(i)$. The number of samples I also depends on the bins. In EXP-II, $T(i) = 63$, and $I = 200$. First, we estimate $C(t)$ and its standard error $\Delta C(t)$ using (1.1). We denote the estimate and standard error of the probabilities as $q_x(t+1) = \Pr(X(t+1) = 1|X(1) = x)$ and $\Delta q_x(t+1)$, respectively. They are estimated from experimental data $\{X(i,t)\}$ as

$$q_x(t+1) = \frac{1 + \sum_{i=1}^{I} X(i,t+1)\delta_{X(i,1),x}}{N_x + 2},$$

$$N_x = \sum_{i=1}^{I} \delta_{X(i,1),x},$$

$$\Delta q_x(t+1) = \sqrt{\frac{q(x,t+1)(1 - q_x(t+1))}{N_x + 3}}.$$

Here, we use the expectation value and standard deviation obtained from the posterior probability distribution for the probabilities. $C(t)$ is then estimated as

$$C(t) = q_1(t+1) - q_0(t+1).$$

The error bars of $C(t)$ are given as

$$\Delta C(t) = \sqrt{\Delta q_1(t+1)^2 + \Delta q_0(t+1)^2}. \tag{1.5}$$

Using $C(t)$ and $\Delta C(t)$, we estimate the error bars of $m_n(t)$ as

$$\Delta m_n(t) = \sqrt{\sum_{s=1}^{t-1} \Delta C(s)^2 (s/t)^{2n}}.$$

Here we assume that $\Delta C(s)$ and $\Delta C(s')$ are independent of each other if $s \neq s'$. We estimate the error bars of $\tau_t(t)$ and $\xi_t(t)$ as

$$\begin{aligned} \Delta\tau_t &= \frac{1}{t}\Delta m_0(t), \\ \Delta\xi_t &= \sqrt{\xi_t}(\Delta m_2(t)/2m_2(t) + \Delta m_0(t)/2m_0(t)). \end{aligned} \tag{1.6}$$

In the estimation of $\Delta\xi_t$, we assume that $\Delta m_2(t)$ and Δm_0 are completely correlated.

1.3.1 EXP-I: General Knowledge Quiz Case

Figure 1.2 plots $C(t)$ versus t. The value of $C(t)$ generally decreases from its initial value of 1 with increasing t. Because the sample number is restricted, $\Delta C(t)$ is large. We see that for difficult questions with $z_0(T) = 50 \pm 5$ and 60 ± 5 %, $C(t)$ is positive for large values of t. On the other hand, for easy questions with $z_0(T) = 70 \pm 5$ and 80 ± 5 %, $C(t)$ decreases to zero with increasing t. These results suggest that the system is in the two-peak phase for difficult questions. For $z_0(T) = 70 \pm 5$ and 80 ± 5 %, an analysis of $q(z)$ showed that the system was in the one-peak phase [16].

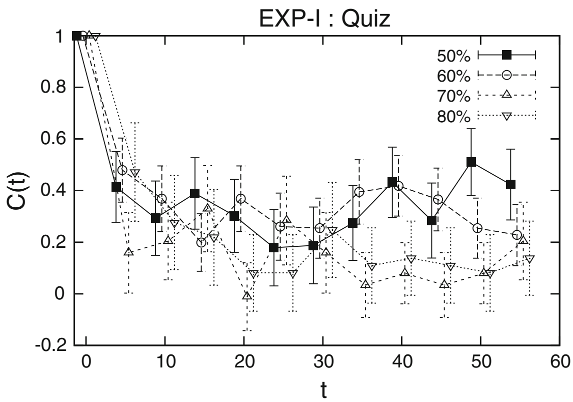

Fig. 1.2 $C(t)$ versus t for EXP-I. The sample choice sequences are classified according to the value of $z_0(T)$ as $z_0(T) = 50 \pm 5$ % (*filled quare*), 60 ± 5 % (*opened circle*), 70 ± 5 % (*opened triangle*), and 80 ± 5 %(*opened down triangle*). We plot only data with the interval $\Delta t = 5$. To see the behavior clearly, we slightly shift the data horizontally

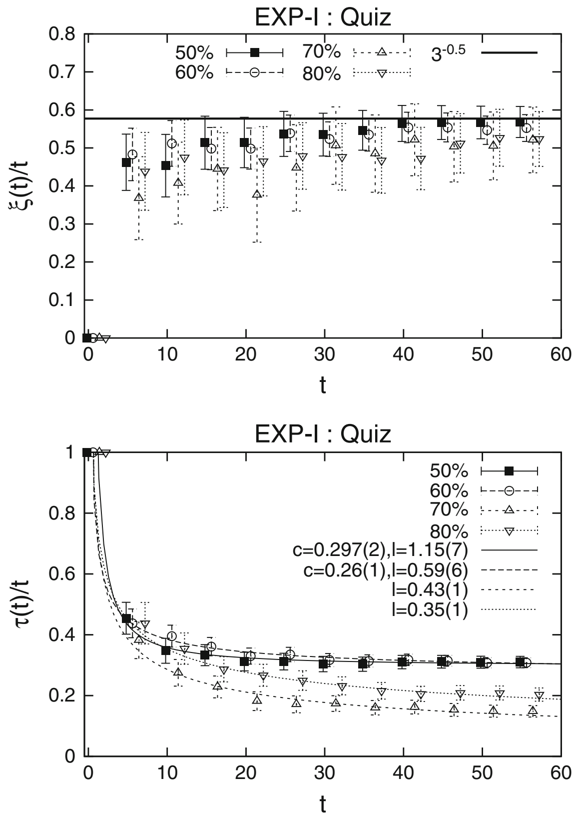

Fig. 1.3 $\xi(t)/t$ and $\tau(t)/t$ versus t for EXP-I with the interval $\Delta t = 5$. We also plot the fitted results for $\tau(t)/t$

Figure 1.3 shows plots of $\xi(t)/t$ and $\tau(t)/t$ versus t. The standard errors for $\xi(t)/t$ are larger than those for $\tau(t)/t$ because $\xi(t)$ is calculated with the second moment $m_2(t)$. For large values of t, $\xi(t)/t$ takes $\sqrt{1/3}$ for difficult questions with $z_0(T) = 50 \pm 5$ and $60 \pm 5\,\%$. The results suggest that the system is in the two-peak phase. For easy questions with $z_0(T) = 70 \pm 5$ and $80 \pm 5\,\%$, $\xi(t)/t \simeq 0.5$ for large values of t. As $\xi(t)/t \simeq \sqrt{l/l+2}$, $l \simeq 0.7$ for easy questions. As l is smaller than 1, the system is in the one-peak phase.

As the system is considered to be in the two-peak phase for $z_0(T) = 50 \pm 5$ and $60 \pm 5\,\%$, we assume $\tau(t)/t = c + d \cdot t^{l-1}$ and estimate c, l, d using the least square fit. We find that $c = 0.297(2)$ for $z_0(T) = 50 \pm 5\,\%$ and $c = 0.26(1)$ for $z_0(T) = 60 \pm 5\,\%$. For $z_0(T) = 70 \pm 5$ and $80 \pm 5\,\%$, we assume $\tau(t)/t = d \cdot t^{l-1}$ and estimate l and d. We find that $l = 0.43(1)$ for $z_0(T) = 70 \pm 5\,\%$ and $l = 0.35(1)$ for $z_0(T) = 80 \pm 5\,\%$, which differ slightly from the value of $l \simeq 0.7$ estimated from $\xi(t)/t$.

1.3.2 EXP-II: Urn Choice Case

Figure 1.4 shows plots of $C(t)$, $\xi(t)/t$, and $\tau(t)/t$ versus t for $q \in \{2/3, 5/9\}$. As the number of samples is larger than that in EXP-I, the standard errors are smaller than the symbols' size for $\tau(t)/t$ and large t. We see that $C(t)$ is positive for large values of t for both cases of q, where $q \in \{2/3, 5/9\}$. In addition, $\xi(t)/t$ for large values of t converges to $\sqrt{1/3}$, and the exponent l for $C(t) \sim t^{l-1}$ is almost one. These results suggest that the system is in the two-peak phase for both values of q. We assume $\tau(t)/t = c + d \cdot t^{l-1}$ and estimate c, l, d using the least square fit. We find that $c = 0.261(1)$ for $q = 2/3$ and $c = 0.207(1)$ for $q = 5/9$.

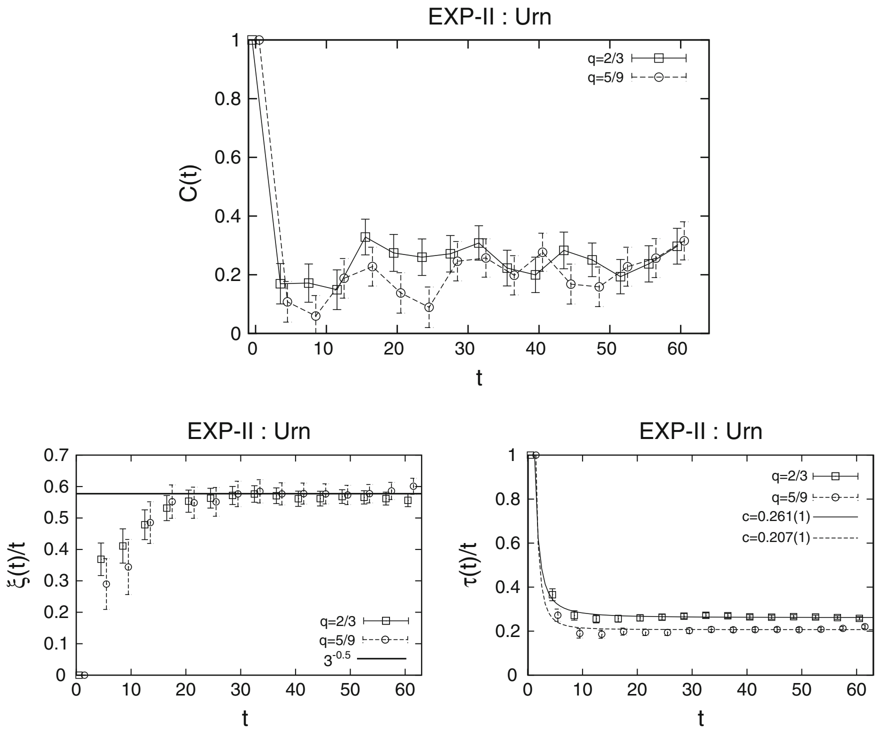

Fig. 1.4 $C(t)$, $\xi(t)/t$, and $\tau(t)/t$ versus t for EXP-II. We use the symbol *opened square* (*opened circle*) for $q = 2/3(5/9)$. We plot only data with the interval $\Delta t = 4$. To see the behavior clearly, we slightly shift the data horizontally

1.4 Conclusion

We studied the self-correcting nature of information cascades. We proposed the use of the normalized correlation function $C(t)$, which shows how the first subject's choice is propagated to later subjects and measures the strength of the domino effect in information cascades. $c \equiv \lim_{t\to\infty} C(t) > 0$ is a sufficient condition for a non-self-correcting information cascade. In this case, the domino effect continues infinitely. The system is in the two-peak phase, and the probability that $z(t)$ converges to $z_- < 1/2$ is positive. We used data from two types of information cascade experiment: EXP-I, which used a general knowledge quiz, and EXP-II, which used urns. The accuracy q of the private signal is $q \in \{2/3, 5/9\}$ in EXP-II. We estimate $C(t)$ and its integrated quantities $\tau(t)$ and $\xi(t)$. In EXP-I, when the questions were difficult, $c > 0$. In EXP-II, $c > 0$ for both cases of q where $q \in \{2/3, 5/9\}$. In these cases, the system is non-self-correcting.

We focus on the study of the non-self-correcting nature of information cascades. Although $c > 0$ is a sufficient condition for a non-self-correcting cascade, $c = 0$ is not a sufficient condition for a self-correcting cascade. To verify this, one should study the response function $q(z)$ and count the number of solutions for $z = q(z)$. Alternatively, it is necessary to study the limit value of the variance of $z(t)$. If there is only one solution, $z_+ > 1/2$, or the limit value is zero, the system is self-correcting. In EXP-I, we studied these points and concluded that the system is self-correcting for $z_0(T) = 70 \pm 5$ and $80 \pm 5\,\%$ [16]. Our experiment for EXP-II and its analysis are under way [17].

Acknowledgments This work was supported by Grant-in-Aid for Challenging Exploratory Research 25610109.

References

1. Catellano, C., Fortunato, S., Loreto, V.: Statistical physics of social dynamics. Rev. Mod. Phys. **81**, 591–646 (2009)
2. Fernández-Gracia, J., Sucheki, K., Ramasco, J.J., Miguel, M.S., Eguíluz, V.M.: Is the Voter Model a model for voters? Phys. Rev. Lett. **112**, 158701–158705 (2014)
3. Bikhchandani, S., Hirshleifer, D., Welch, I.: A theory of fads, fashion, custom, and cultural changes as informational cascades. J. Polit. Econ. **100**, 992–1026 (1992)
4. Devenow, A., Welch, I.: Rational herding in financial economics. Euro. Econ. Rev. **40**, 603–615 (1996)
5. Lee, I.H.: On the convergence of informational cascades. J. Econ. Theory **61**, 395–411 (1993)
6. Anderson, L.R., Holt, C.A.: Information cascades in the laboratory. Am. Econ. Rev. **87**, 847–862 (1997)
7. Kübler, D., Weizsäcker, G.: Limited depth of reasoning and failure of cascade formation in the laboratory. Rev. Econ. Stud. **71**, 425–441 (2004)
8. Goeree, J.K., Palfrey, T.R., Rogers, B.W., McKelvey, R.D.: Self-correcting information cascades. Rev. Econ. Stud. **74**, 733–762 (2007)
9. Mori, S., Hisakado, M., Takahashi, T.: Phase transition to two-peaks phase in an information cascade voting experiment. Phys. Rev. E **86**, 026109–026118 (2012)

10. Hisakado, M., Mori, S.: Digital herders and phase transition in a voting model. J. Phys. A **44**, 275204–275220 (2011)
11. Hisakado, M., Mori, S.: Two kinds of phase transitions in a voting model. J. Phys. A, Math. Theor. **45**, 345002–345016 (2012)
12. Hill, B., Lane, D., Sudderth, W.: A strong law for some generalized urn processes. Ann. Prob. **8**, 214–226 (1980)
13. Hod, S., Keshet, U.: Phase-transition in binary sequences with long-range correlations. Phys. Rev. E **70**, 015104–015109 (2004)
14. Mori, S., Hisakado, M.: Finite-size scaling analysis of binary stochastic processes and universality classes of information cascade phase transition. J. Phys. Soc. Jpn. **84**, 054001–054013 (2015)
15. Mori, S., Hisakado, M.: Correlation function for generalized Pólya urns: Finite-size scaling analysis. Phys. Rev. E. **92**, 052112–052121 (2015)
16. Mori, S., Hisakado, M., Takahashi, T.: Collective adoption of max-min strategy in an information cascade voting experiment. J. Phys. Soc. Jpn. **82**, 084004–084013 (2013)
17. Hino, M., Hisakado, M., Takahashi, T., Mori, S.: Detection of phase transition in generalized Plya urn in Information cascade experiment. J. Phys. Soc. Jpn. **85**, 034002–034013 (2016)

Chapter 2
Fitting Planar Proximity Graphs on Real Street Networks

Dimitris Maniadakis and Dimitris Varoutas

Abstract Due to the rising progress of sustainable urban infrastructures, modeling realistic street networks is a fundamental challenge. This study contributes to this modeling direction, by suggesting the utilization of planar proximity graphs, and specifically the β-skeleton graphs. Their goodness of fit on producing real-like urban street networks is verified by comparison to real data. In particular, the basic topological and geometrical properties derived from synthetic β-skeleton planar graphs are compared to the properties of five urban street network datasets, all represented using the Primal approach. A good agreement with empirical patterns is found and a possible explanation is discussed.

2.1 Introduction

There are broad agreements that the street patterns shape overlay infrastructure deployment since they define a basic template which strongly constrains the further development of other webs (e.g., power grid or communication networks). Due to the rising progress of sustainable urban infrastructures, understanding and modeling the structure of street networks is an elementary challenge. Despite a large number of studies on street networks, the existing modeling methodologies are mostly long, random-based and simulation-based, which require several assumptions for generating a realistic street layout, e.g., [1].

On the other hand, the construction of planar proximity graphs can be straightforward by using analytical or simulation methods. Planar proximity graphs are planar graphs (edges intersect only in the points/nodes) where two points in Euclidean plane are connected by an edge if they are close in some sense. Each pair of points is assigned a certain neighborhood, and the points of the pair are connected by an

D. Maniadakis (✉) · D. Varoutas
Department of Informatics and Telecommunications, University of Athens, Athens, Greece
e-mail: D.Maniadakis@di.uoa.gr

D. Varoutas
e-mail: D.Varoutas@di.uoa.gr

S. Battiston et al. (eds.), *Proceedings of ECCS 2014*, Springer Proceedings in Complexity, DOI 10.1007/978-3-319-29228-1_2

edge if their neighborhood is empty. The Delaunay Triangulation (DT), the Relative Neighborhood Graph (RNG), the Gabriel Graph (GG) and the Minimum Spanning Tree (MST) are well known examples of proximity graphs. These are constructed by parameter-less algorithms, given the nodes positions. Specifically, the DT for a set of points in a plane is a triangulation such that no point is inside the circumcircle of any triangle; the RNG is defined by connecting two points whenever there does not exist a third point closer to both points; the GG is a graph where two points have an edge between them if no other point exists in the circumball containing the two points; last, the MST is a tree consisting of all points while having the minimum total weight (length). Though, the β-skeleton graphs [2] constitute a parameterized family of planar proximity graphs where different β values give rise to different graphs.

This study contributes to the urban street modeling, examining the fitness of planar proximity graphs, particularly the β-skeleton graphs, on real street networks with complex characteristics. Additionally, a possible explanation is discussed concerning the findings of the analysis.

The rest of this paper is structured as follows. Section 2.2, contains some preliminaries on the β-skeleton concept. The datasets and the methodology used are described in Sect. 2.3, while the results of applying the methodology are presented in Sect. 2.4. Section 2.5 discusses a possible explanation of the findings and finally Sect. 2.6 concludes the study.

2.2 The β-Skeleton Graphs

In the lune-based neighborhoods approach [2], given a spatial distribution of points S in two-dimensional space, two points u and v are connected by an edge whenever the intersection of the two disks of radius r, centered at the points $c1$ and $c2$, contains no points of S (see Fig. 2.1).

The case $\beta = 0$ corresponds to the DT, $\beta = 1$ corresponds to the GG and $\beta = 2$ corresponds to the RNG. For $1 \leq \beta < \infty$, the radius and the disk centers are defined as follows:

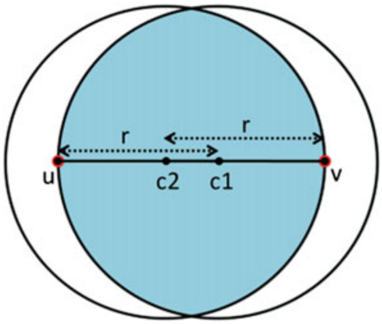

Fig. 2.1 Definition of β-skeleton in the lune-based variant for $1 \leq \beta < \infty$

$$r = \frac{\beta \cdot D(u, v)}{2} \tag{2.1}$$

$$c1 = \left(1 - \frac{\beta}{2}\right) \cdot u + \left(\frac{\beta}{2}\right) \cdot v \tag{2.2}$$

$$c2 = \left(\frac{\beta}{2}\right) \cdot u + \left(1 - \frac{\beta}{2}\right) \cdot v \tag{2.3}$$

while for $0 < \beta < 1$ the two disks pass through both u and v, with radius given by:

$$r = \frac{D(u, v)}{2 \cdot \beta} \tag{2.4}$$

The parameter β determines the size and shape of the lune-based neighbourhood. With the increase of β, the number of edges in the β-skeleton decreases (see Fig. 2.2).

A β-skeleton of a random planar set usually becomes a disconnected graph for $\beta > 2$ and continues losing its edges with further increase of β [3]. On the other hand, as β approaches zero, more and more edges are added to the β-skeleton until it eventually forms the complete geometric graph. For $1 \leq \beta \leq 2$, the following relationships among the different proximity graphs hold for any finite set of points S in the plane:

$$DT(S) \supseteq GG(S) \supseteq \beta - skeleton(S, \beta) \supseteq RNG(S) \supseteq MST(S) \tag{2.5}$$

Since urban street networks are usually connected networks neither DT-like, nor MST-like [4], it is thus of interest to answer to the following questions; (a) is there sufficient accuracy when using β-skeletons with $1 \leq \beta \leq 2$ to reproduce urban street networks? (b) is there a particular β value or subrange of values for which the accuracy is better? (c) what is the possible mechanism that leads real street networks to be associated with particular β values?

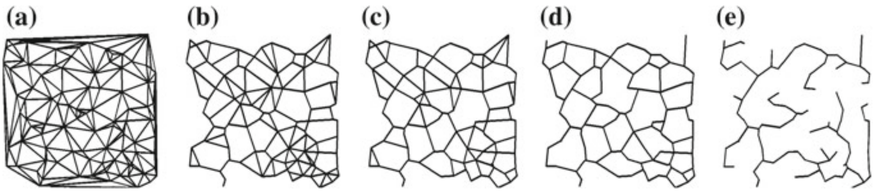

Fig. 2.2 Graph visualizations for the same set of 100 points: **a** delaunay triangulation ($\beta = 0$), **b** Gabriel graph ($\beta = 1$), **c** β-skeleton (here $\beta = 1.4$), **d** relative neighborhood graph ($\beta = 2$), **e** minimum spanning tree