The plant cytoskeleton as a transportation network

Dissertation

zur Erlangung des akademischen Grades
Doctor rerum naturalium
in der Wissenschaftsdisziplin “Systembiologie”

eingereicht in kumulativer Form
an der Mathematisch-Naturwissenschaftlichen Fakultät
der Universität Potsdam von

David Breuer

Potsdam 2016

Published online at the
Institutional Repository of the University of Potsdam:
URN urn:nbn:de:kobv:517-opus4-93583
http://nbn-resolving.de/urn:nbn:de:kobv:517-opus4-93583

1 Contents

2 Abstract

The cytoskeleton is an essential component of living cells. It is composed of different types of protein filaments that form complex, dynamically rearranging, and interconnected networks. The cytoskeleton serves a multitude of cellular functions which further depend on the cell context. In animal cells, the cytoskeleton prominently shapes the cell's mechanical properties and movement. In plant cells, in contrast, the presence of a rigid cell wall as well as their larger sizes highlight the role of the cytoskeleton in long-distance intracellular transport. As it provides the basis for cell growth and biomass production, cytoskeletal transport in plant cells is of direct environmental and economical relevance. However, while knowledge about the molecular details of the cytoskeletal transport is growing rapidly, the organizational principles that shape these processes on a whole-cell level remain elusive.

This thesis is devoted to the following question: How does the complex architecture of the plant cytoskeleton relate to its transport functionality? The answer requires a systems level perspective of plant cytoskeletal structure and transport. To this end, I combined state-of-the-art confocal microscopy, quantitative digital image analysis, and mathematically powerful, intuitively accessible graph-theoretical approaches.

This thesis summarizes five of my publications that shed light on the plant cytoskeleton as a transportation network: (1) I developed network-based frameworks for accurate, automated quantification of cytoskeletal structures, applicable in, e.g., genetic or chemical screens; (2) I showed that the actin cytoskeleton displays properties of efficient transport networks, hinting at its biological design principles; (3) Using multi-objective optimization, I demonstrated that different plant cell types sustain cytoskeletal networks with cell-type specific and near-optimal organization; (4) By investigating actual transport of organelles through the cell, I showed that properties of the actin cytoskeleton are predictive of organelle flow and provided quantitative evidence for a coordination of transport at a cellular level; (5) I devised a robust, optimization-based method to identify individual cytoskeletal filaments from a given network representation, allowing the investigation of single filament properties in the network context. The developed methods were made publicly available as open-source software tools.

Altogether, my findings and proposed frameworks provide quantitative, system-level insights into intracellular transport in living cells. Despite my focus on the plant cytoskeleton, the established combination of experimental and theoretical approaches is readily applicable to different organisms. Despite the necessity of detailed molecular studies, only a complementary, systemic perspective, as presented here, enables both understanding of cytoskeletal function in its evolutionary context as well as its future technological control and utilization.

Zusammenfassung

Das Zytoskelett ist ein notwendiger Bestandteil lebender Zellen. Es besteht aus verschiedenen Arten von Proteinfilamenten, die ihrerseits komplexe, sich dynamisch reorganisierende und miteinander verknüpfte Netzwerke bilden. Das Zytoskelett erfüllt eine Vielzahl von Funktionen in der Zelle. In Tierzellen bestimmt das Aktin-Zytoskelett maßgeblich die mechanischen Zelleigenschaften und die Zellbewegung. In Pflanzenzellen hingegen kommt dem Aktin-Zytoskelett eine besondere Bedeutung in intrazellulären Transportprozessen zu, bedingt insbesondere durch die starre pflanzliche Zellwand sowie die Zellgröße. Als wesentlicher Faktor für Zellwachstum und somit auch die Produktion von Biomasse, ist Zytoskelett-basierter Transport daher von unmittelbarer ökologischer und ökonomischer Bedeutung. Während das Wissen über die molekularen Grundlagen Zytoskelett-basierter Transportprozesse beständig wächst, sind die zugrunde liegenden Prinzipien zellweiter Organisation bisher weitgehend unbekannt.

Diese Dissertation widmet sich daher folgender Frage: Wie hängt die komplexe Architektur des pflanzlichen Zytoskeletts mit seiner intrazellulären Transportfunktion zusammen? Eine Antwort auf diese Frage erfordert eine systemische Perspektive auf Zytoskelettstruktur und -transport. Zu diesem Zweck habe ich Mikroskopiedaten mit hoher raumzeitlicher Auflösung sowie Computer-gestützte Bildanalysen und mathematische Ansätzen der Graphen- und Netzwerktheorie kombiniert.

Die vorliegende Dissertation umfasst fünf meiner Publikationen, die sich einem systemischen Verständnis des pflanzlichen Zytoskeletts als Transportnetzwerk widmen: (1) Dafür habe ich Bilddaten-basierte Netzwerkmodelle entwickelt, die eine exakte und automatisierte Quantifizierung der Architektur des Zytoskeletts ermöglichen. Diese Quantifizierung kann beispielsweise in genetischen oder chemischen Versuchen genutzt werden und für eine weitere Erforschung der genetischen Grundlagen und möglicher molekularer Interaktionspartner des Zytoskeletts hilfreich sein; (2) Ich habe nachgewiesen, dass das pflanzliche Aktin-Zytoskelett Eigenschaften effizienter Transportnetzwerk aufweist und Hinweise auf seine evolutionären Organisationsprinzipien liefert; (3) Durch die mathematische Optimierung von Transportnetzwerken konnte ich zeigen, dass unterschiedliche Pflanzenzelltypen spezifische und optimierte Organisationsstrukturen des Aktin-Zytoskeletts aufweisen; (4) Durch quantitative Analyse des Transports von Organellen in Pflanzenzellen habe ich nachgewiesen, dass sich Transportmuster ausgehend von der Struktur des Aktin-Zytoskeletts vorhersagen lassen. Dabei spielen sowohl die Organisation des Zytoskeletts auf Zellebene als auch seine Geometrie eine zentrale Rolle. (5) Schließlich habe ich eine robuste, optimierungs-basierte Methode entwickelt, die es erlaubt, individuelle Filamente eines Aktin-Netzwerks zu identifizieren. Dadurch ist es möglich, die Eigenschaften einzelner Zytoskelettfilamente im zellulären Kontext zu untersuchen. Die im Zuge dieser Dissertation entwickelten Methoden wurden frei und quelloffen als Werkzeuge zur Beantwortung verwandter Fragestellung zugänglich gemacht.

Insgesamt liefern die hier präsentierten Ergebnisse und entwickelten Methoden quantitative, systemische Einsichten in die Transportfunktion des Zytoskeletts. Die hier etablierte Kombination von experimentellen und theoretischen Ansätzen kann, trotz des Fokusses auf das pflanzliche Zytoskelett, direkt auf andere Organismen angewendet werden. Als Ergänzung molekularer Studien bildet ein systemischer Blickwinkel, wie er hier entwickelt wurde, die Grundlage für ein Verständnis sowohl des evolutionären Kontextes als auch zukünftiger Kontroll- und Nutzungsmöglichkeiten des pflanzlichen Zytoskeletts.

3 Introduction

The cytoskeleton is an essential intracellular component of organisms from all kingdoms of life. It is composed of different components, typically comprising actin filaments, microtubules, and intermediate filaments, which are protein polymers that may rearrange dynamically through (de-)polymerization processes and form intricate network structures (Section 3.1). While many aspects of cytoskeletal composition and function are conserved across species, there are notable differences, in particular between animal and plant cells (Section 3.2). In plants, the actin cytoskeleton provides the basis for cellular transport and, hence, cell maintenance and growth (Section 3.3). Despite the importance of actin-based cellular transport, the vast majority of studies has focused on elucidating its molecular details, neglecting a systems perspective. However, only a systems perspective can enable understanding of the organizational principles and of the intricate coordination of cellular transport at a cellular level (Section 3.5). Although system-wide actin-based cellular trafficking has been tackled in a few theoretical studies, there has been no attempt to integrate imaging data of the actin cytoskeleton and cellular transport. In particular, a network-based framework that naturally captures the filamentous structure of the cytoskeleton and incorporates biological data is lacking (Section 3.7). Therefore, the aim of this thesis is the development of an image- and network-based framework for a systems perspective on actin-based transport (Section 3.9). Such a framework allows quantitative analyses and interpretations of real-life cytoskeletal network structures. Thus, it provides a cornerstone towards comprehension of cytoskeletal functions in its evolutionary context as well as its future technological control and utilization.

3.1 Basics of the cytoskeleton with a focus on its actin component

The concept “cytoskeleton” was hypothesized at the beginning of the 19th century [Dujardin, 1835], as a necessary cellular structure for the positioning of organelles, although its name was only introduced around a century later [Wintrebert, 1931]. Nowadays, it is known that the cytoskeleton is an essential component of living cells that form a scaffold-like structure spanning the cytoplasm [Cooper, 2000; Liu, 2010]. In particular, the cytoskeleton serves a multitude of different functions including, indeed, organelle positioning (cf. Section 3.3). Despite its name, the cytoskeleton is not a rigid and stable structure, but is composed of different arrays of protein polymers that rearrange dynamically through (de-)polymerization processes. The cytoskeleton in cells of most organisms comprises actin filaments, microtubules, and intermediate filaments (Figs. 3.1A and B; [Cooper, 2000; Liu, 2010]). In this thesis, we focus on the actin cytoskeleton and its specific role in cellular transport in plant cells.

The basic unit of the actin cytoskeleton are (globular) G-actin proteins (Fig. 3.2A, upper). G-actin is notable by itself in at least three ways: It is one of the most abundant proteins in eukaryotes [Lodish et al., 2000], its amino acid sequence is highly conserved across species [Gunning et al., 2015], and its structure enables interaction with an exceptionally large range of other proteins [Dominguez, 2004].

The monomeric G-actin may further form (filamentous) F-actin, or actin filaments, which constitute the

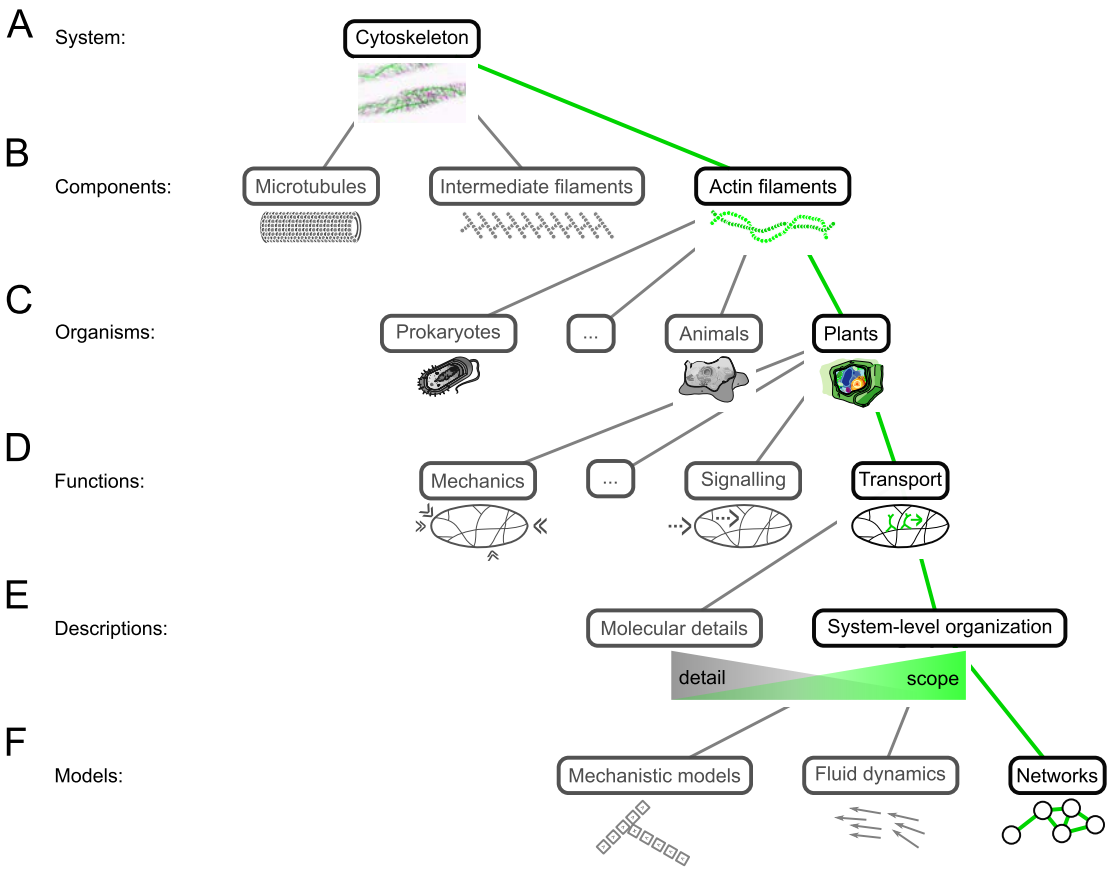


Figure 3.1: **Overview of the thesis' focus on system-level transport by the actin cytoskeletal networks in plant cells and related areas of research.** The research presented in this thesis focuses on the path highlighted in dark green. **(A)** System of interest: Cytoskeleton. Confocal recording of actin cytoskeleton (light green) and microtubules (magenta) of two hypocotyl *Arabidopsis thaliana* plant cells. **(B)** Components: Microtubules, intermediate filaments, and actin filaments. Schematics of different polymeric structures of the cytoskeletal components. Actin filaments are flexible, linear polymer chains that may align and form stiff bundles. **(C)** Organisms: Prokaryotes and eukaryotes such as animals and plants. Schematics of representative cells, whereby plant cells are distinguished among other things by a rigid cell wall and the presence of multiple Golgi bodies. **(D)** Functions: Mechanics, signaling, transport, and others. Cellular transport in plant cells relies on the actin cytoskeleton which serves as tracks for motor proteins that carry different types of cargo such as Golgi bodies. **(E)** Descriptions: Due to the complexity of many biological systems, such as the acto-myosin transport system in plant cells, a trade-off must be considered between the amount of incorporated details of the systems and the scope of their description. **(F)** Models: Mechanistic models, fluid-dynamical models, and networks. The network framework naturally captures the discrete filamentous structure of the actin cytoskeleton, while enabling both intuitive and computationally efficient investigation of its organizational principles as well as comparison to other types of biological and human-made transportation networks.

network-like structure of the actin cytoskeleton (Fig. 3.2A, lower; [Staiger et al., 2000; Carlsson, 2010; Liu, 2010]). Filament formation involves adenosine triphosphate (ATP)-driven nucleation and elongation processes that depend on G-actin concentrations and regulatory actin-binding proteins [Staiger and Blanchoin, 2006]. In the formed filaments, G-actin monomers establish filament polarity by aligning such that the exposed ATP binding sites (called pointed or "-" end) and their opposite (called barbed or "+" end) point toward either end of the filament. Actin filaments exhibit a helical, single-stranded structure of around 7 nm in diameter. Due to different (de-)polymerization rates at the "+" and "-" ends, actin filaments may grow, shrink, or, in equilibrium, show treadmilling behavior.

Furthermore, multiple actin filaments may in turn form bundles via actin-binding, cross-linking proteins [Bartles, 2000; Claessens et al., 2006; Thomas et al., 2009]. These actin bundles may contain filaments of different polarity and can be up to 150 nm in diameter [Haviv et al., 2008]. Actin bundles play a major role in cellular transport (cf. Section 3.3; [Thomas et al., 2009; Akkerman et al., 2011]).

3.2 Comparison of actin cytoskeleton in plants and other organisms

Many features of the cytoskeleton are conserved across Eukaryotes [van den Ent et al., 2001; Wickstead and Gull, 2011]. For example, both actin filaments and microtubules are found in cells of most eukaryotes [Cooper, 2000; Liu, 2010], and evidence suggests that also bacteria [Ausmees et al., 2003; Shih and Rothfield, 2006; Bagchi et al., 2008] and plant cells contain intermediate filament-like proteins [Hargreaves et al., 1989] that were previously attributed only to animal cells. Moreover, as mentioned previously, the amino acid sequence of actin is among the most highly conserved sequences [Gunning et al., 2015].

Nevertheless, beyond these basic similarities, there are substantial differences in actin cytoskeleton organization and dynamics, especially between plant and non-plant species (Fig. 3.1C; [Staiger and Lloyd, 1991; Staiger et al., 2000; Liu, 2010]). These differences range from lower ratios of F-actin and G-actin in plant cells (between $1-10\%$ [Gibbon et al., 1999; Snowman et al., 2002] as compared to $30-70\%$ in animal cells [McGrath et al., 1998, 2000]) to substantially higher growth rates of actin filaments (up to $10-100$ times faster than in many non-plant cells [Staiger et al., 2009]) to different functions (cf. Section 3.3). Moreover, in plants, intracellular transport depends mainly on the actin cytoskeleton while in animals, it is largely microtubule-based [Nebenführ et al., 1999; Vale, 2003; Liu, 2010].

The differences between plant and non-plant actin cytoskeletons can be attributed to three aspects (Fig. 3.2B; [Wasteneys, 2000; Ehrhardt and Shaw, 2006]): The presence of a large, central vacuole in plant cells restricts large parts of the actin cytoskeleton to the cell cortex and implies a largely two-dimensional, cylindrical geometry that strongly differs from the pervasive, three-dimensional actin architecture in animal cells [Kost and Chua, 2002]. The rigid cell wall of immobile plant cells renders the mechanical effect of the actin cytoskeleton on cell stability less relevant as compared to flexible and mobile animal cells [Kost and Chua, 2002]. Moreover, plants cells are often larger than animal cells, which affects the geometry of the actin cytoskeleton and increases the necessity for long-range cellular transport (cf. Section 3.3; [Shimmen and Yokota, 2004; Liu, 2010]).

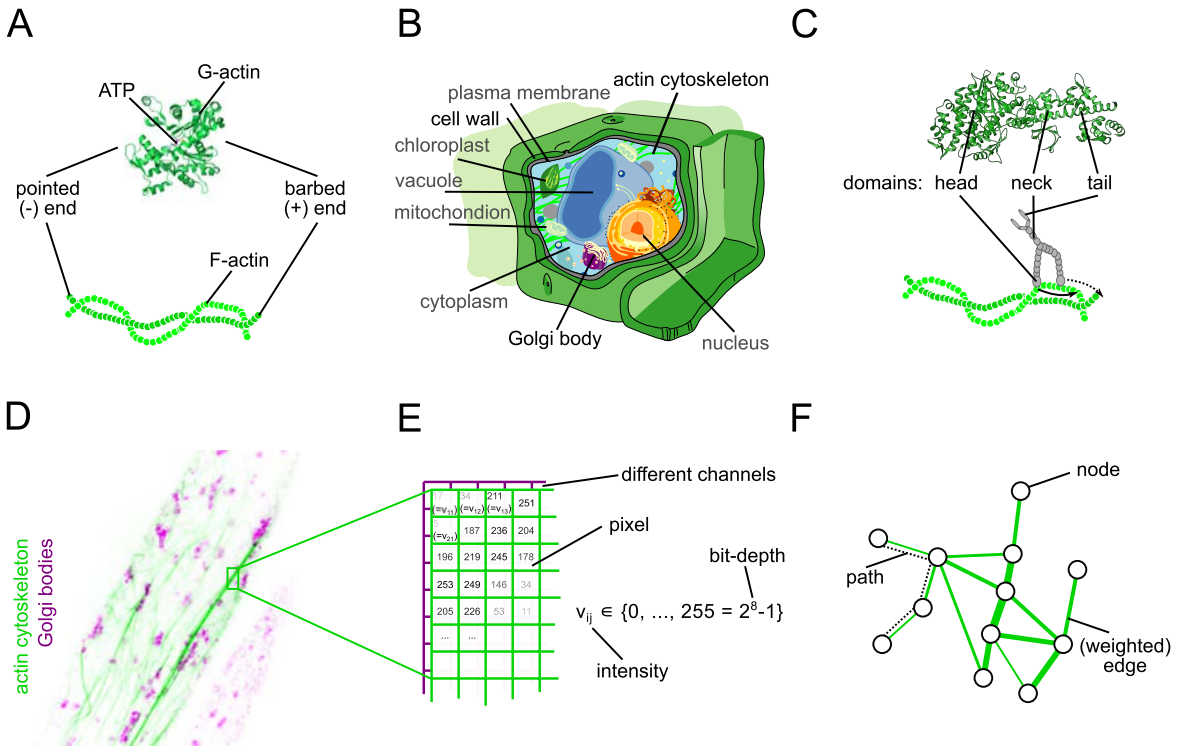


Figure 3.2: **Details of different components of the acto-myosin transport system in plant cells, and schematics of the employed imaging and network approaches.** **(A)** Molecular structure of (globular) G-actin with bound ATP that prevents the G-actin's loss of structural integrity (upper; CC BY-SA 3.0 Thomas Splettstoesser), whereby the ATP-binding site and its opposite are referred to as pointed (or "-") and barbed (or "+") end, respectively. G-actin monomers may polymerize and form a helical, (filamentous) F-actin strand (lower). **(B)** Schematic of a typical plant cell with its rigid cell wall, the actin cytoskeleton, and several Golgi bodies. **(C)** Molecular structure of myosin, i.e., a motor protein with a head domain that binds filamentous actin, a neck domain, and a tail domain that binds cellular cargo (upper; CC BY-SA 3.0 Thomas Splettstoesser). Through ATP hydrolysis at the catalytic region in the neck domain, the two myosin heads propel the motor protein towards the barbed end of the actin filament. **(D)** Confocal recording of plant actin cytoskeleton (green) and Golgi bodies (magenta). **(E)** Digital representation of image data. Intensities of different channels (green and magenta) are stored separately as matrices where each entry corresponds to a pixel, here in an 8-bit encoding. **(F)** Mathematical representation of cytoskeletal network structures as a graph. The edges reflect the filament segments, their weights capture the average filament segment thickness, and the nodes are the endpoints of filament segments. Example of a path, i.e., an ordered sequence of adjacent edges that may be used to describe cytoskeletal filaments.

The majority of research on the actin cytoskeleton has focused on animal, fungal, or prokaryotic cells [Cooper, 2000]. However, the above-mentioned differences in the actin cytoskeleton between plant and non-plant cells clearly demonstrate that focused studies are needed to unravel the specific organization and function of the actin cytoskeleton in plant cells.

3.3 The plant actin cytoskeleton facilitates cellular transport and various other tasks

The plant actin cytoskeleton is highly multi-functional, not least due to the diverse interaction opportunities of actin proteins [Dominguez, 2004], and may perform many functions simultaneously (Fig. 3.1D, [Volkmann and Baluska, 1999; Nick, 2007; Liu, 2010]). It enables anisotropic cell expansion [Cosgrove, 1996; Smith and Oppenheimer, 2005], it is central for intracellular regulatory processes [Staiger, 2000] as well as for signal transduction of external stimuli [Volkmann and Baluska, 1999; Shafrir and Forgacs, 2002]. Moreover, the actin cytoskeleton plays an important role in plant cell division and the formation alignment of the cell plate [Staehelin and Hepler, 1996; Hussey et al., 2006]. As opposed to animal cells, the rigid cell wall of plant cells renders actin cytoskeletal support against mechanical stress and strain less relevant [Kost and Chua, 2002]. However, especially due to their larger sizes [Cooper, 2000; Liu, 2010], plant cells have a strong need for long-range cellular transport which is also met by the actin cytoskeleton [Staiger and Lloyd, 1991; Shimmen and Yokota, 2004; Prokhnevsky et al., 2008; Akkerman et al., 2011].

The basic requirement of cells for long-range transport systems is a consequence of the heterogeneous and crowded cytoplasm that comprises a large range of molecules and organelles [Luby-Phelps, 2000; Ellis, 2001]. Since diffusion through this complex environment is not sufficient to match varying demands for cell maintenance and growth, intricate cellular transport schemes have evolved [Geisler et al., 2008; Goldstein et al., 2008; Wightman and Turner, 2010; Goldstein and van de Meent, 2015; Wang and Hussey, 2015]. In plants, this transport relies on the acto-myosin system, i.e., the movement of myosin along the actin cytoskeleton [Staiger and Lloyd, 1991; Shimmen and Yokota, 2004; Prokhnevsky et al., 2008; Akkerman et al., 2011]. Myosins are ATP-driven, actin-binding motor proteins that carry cargo (typically) towards the barbed end of actin filaments (Fig. 3.2C; [Cooper, 2000; Shimmen et al., 2000; Cai et al., 2014]). The myosin family comprises a large number of different classes [Shimmen et al., 2000; Peremyslov et al., 2011], whereby myosin XI has been shown to be the major driver of cellular transport in plants [Holweg and Nick, 2004; Avisar et al., 2008b; Peremyslov et al., 2010, 2008; Cai et al., 2014].

The cellular cargo of the acto-myosin transport system is diverse. It comprises secretory vesicles for the delivery of cellular material to growing parts of the cell [Mathur, 2004], vesicles for trafficking between endomembrane compartments [Wightman and Turner, 2008; Gutierrez et al., 2009; Wang and Hussey, 2015], small cargo, like mRNAs [Bassell and Singer, 1997], and also large organelles such as mitochondria, chloroplasts, and Golgi (Fig. 3.2D; [Nebenführ et al., 1999; Staiger, 2000; Nebenführ and Staehelin, 2001; Avisar et al., 2008b; Akkerman et al., 2011]). In contrast to animal cells with a single Golgi apparatus, plant cells often comprise many separate, mobile Golgi bodies [Griffing, 1991; Nebenführ and Staehelin, 2001]. Golgi bodies synthesize for example hemicelluloses and pectins, two of the main constituents of the plant cell wall

[Gibeaut and Carpita, 1994; Dupree and Sherrier, 1998; Somerville et al., 2004], that are subsequently incorporated into the cell wall via exocytosis [Geisler et al., 2008]. An impaired plant actin cytoskeleton leads to Golgi aggregation and both reduced exo- and endocytosis [Nebenführ et al., 1999; Avisar et al., 2008a; Akkerman et al., 2011; Sampathkumar et al., 2013]. Therefore, actin-based Golgi transport is crucial for cell growth, especially in young hypocotyl cells which are a well-established model system for cell growth [Hussey et al., 2006; Akkerman et al., 2011; Sampathkumar et al., 2011].

Another special feature of plant cells as compared to animal cells is the directed flow of cytosol and organelles called cytoplasmic streaming [Shimmen and Yokota, 2004; Shimmen, 2007; Goldstein and van de Meent, 2015]. Beyond the direct movement of cargo by the acto-myosin system, the high velocities of plant myosins [Peremyslov et al., 2008; Liu, 2010] and the fast (de-)polymerization dynamics of plant actin filaments [Staiger et al., 2009], in combination with the viscous cytoplasmic environment [Luby-Phelps, 2000; Ellis, 2001] may establish bulk flow streaming patterns at a cellular level. Interestingly, different cell type exhibit different streaming patterns [Shimmen and Yokota, 2004; Goldstein and van de Meent, 2015]. The best studied type of streaming is rotational streaming in large internodal cells of characean algae [Kachar and Reese, 1988; Foissner, 2004; Woodhouse and Goldstein, 2013], which are closely related to higher plants. Plant cells that exhibit longitudinal tip-growth, like root hairs or pollen tubes, typically show fountain-like streaming where the flow along the central cell axis is opposite to the flow towards the cell membrane [Allen and Allen, 1978; Goldstein and van de Meent, 2015]. However, the most common form of cytoplasmic streaming involves a less organized, more erratic and saltatory form of cytoplasmic movement in which cytoplasmic particles or organelles jump larger distances [Nebenführ et al., 1999; Goldstein and van de Meent, 2015]. Thus, direct actin-based transport of cargo in plant cells may be substantially supported by indirect bulk flow movement. Indeed, for certain organelles is it still under debate whether they move directly through binding to motor proteins or whether they are carried on indirectly via cytoplasmic streaming [Buchnik et al., 2014; Cai and Cresti, 2012].

3.4 Methods 1: Systems microscopy and quantitative image analysis

Many fields of biological research, and cell biology in particular, increasingly rely on microscopy data [Swedlow and Eliceiri, 2009; Eliceiri et al., 2012]. Apart from the preference of humans for visual input, this increase is due to three reasons: First, technical advances continue to increase the spatial and temporal resolution of microscopes while keeping the damage to the organism low. Thus, these advances enable imaging of processes at a molecular scale in living systems [Stephens and Allan, 2003; Sheahan et al., 2004; Li et al., 2015]. Second, imaging approaches capture the organization and dynamics of cellular components on a systems level, enabling the investigation of global organizational and functional principles beyond local genetic and biochemical interactions (cf. Section 3.5 below). Third, by providing digital data, microscopy studies allow fast, automated, quantitative, and unbiased analyses [Swedlow and Eliceiri, 2009; Eliceiri et al., 2012]. Therefore, in combination with advances in machine learning methods and decreasing costs for computing power, imaging-based approaches provide a powerful means towards answering questions from cell biology.

Microscopy data are in the form of images that may capture different regions (e.g., z-layers that capture slices

at varying depths in the sample) or time points of a given object as well as different channels (e.g., RGB in color images or channels for different emission wavelengths of fluorescent markers in confocal microscopy [Chalfie et al., 1994; Lippincott-Schwartz and Patterson, 2003]). In digital image analysis, each image consists of discrete pixels, arranged on a rectangular grid, conveniently represented as table, or matrix, where each entry denotes the intensity of the respective pixel (Fig. 3.2E; [Gonzalez and Woods, 2009]). Typically, images are stored in an 8-bit (or 16-bit) encoding that keeps image memory requirements low, while providing sufficient intensity resolution. This digital representation is inherently quantitative and allows various options for automated image processing and analysis.

In most classical bioimaging studies, scalar values are measured, such as absolute or relative intensities of certain regions (e.g., in co-localization studies [Boevink et al., 1998; Gonzalez and Woods, 2009; Sampathkumar et al., 2011; North, 2006]) or sizes of specific structures (e.g., for determining organ or plant sizes [Gonzalez and Woods, 2009; Gonzalez et al., 2012; Ivakov and Persson, 2013; Apelt et al., 2015]). The cytoskeleton and other complex structures, however, cannot be described using simple scalar values and more sophisticated frameworks are required (cf. Section 3.6 below).

3.5 Understanding actin-based transport in plant cells: Molecular details or systems perspective

Cells are the building blocks of terrestrial life [Ingber, 1998; Cooper, 2000]. At the same time, already cells by themselves are complex, multi-scale systems: They comprise different and interlinked layers of molecular, i.e., genetic, metabolic, and regulatory processes (on a nm – μm scale) and contain many interacting compartments (on a μm scale). Furthermore, they rely on diverse processes on a cellular level, such as cellular trafficking and transport (on a μm – mm scale) which are in turn based on molecule-level processes. Due to this multi-scale complexity of the cell, cell biology is faced with an inherent trade-off between a detailed understanding of the molecular mechanisms and a comprehensive view of the system-level organizational principles (Fig. 3.1E).

Most biological studies of actin-based transport in plant cells have previously focused on elucidating its molecular features. Besides the accumulating knowledge about details of actin and actin filament formation, dynamic behavior of actin filaments defined as buckling and straightening have been reported in plant interphase cells and severing is suggested to be a major control mechanism for actin cytoskeleton organization [Staiger et al., 2009]. Besides the molecular details of structure and function of individual motor proteins, the interaction of multiple motor proteins [Badoual et al., 2002; Müller et al., 2008; Guérin et al., 2010], their movement at junctions of cytoskeletal filaments [Bálint et al., 2013; Osunbayo et al., 2015], as well as the local increases in transport velocities due to bundling of cytoskeletal filament are increasingly well understood [Akkerman et al., 2011].

Contrary to this wealth of molecular details, a system-level perspective of cytoskeletal transport in general, or actin-based transport in plants in particular, lags behind. So far, only a few imaging-based studies have addressed cytoskeletal organization and transport at a systems level. These studies are discussed in Section 3.7 below.

3.6 Methods 2: Networks, networks everywhere

Many biological and technical systems are complex systems, i.e., they are composed of units that exhibit non-trivial interactions. Examples cover metabolic networks [Fell and Wagner, 2000; Jeong et al., 2000], neural networks [Felleman and Van Essen, 1991; Koch and Laurent, 1999], and transportation systems like road [Barthélemy and Flammini, 2008; Barthélemy, 2011], air transportation [Guimerà et al., 2005; Bianconi et al., 2009], or leaf venation networks [West et al., 1999; Katifori et al., 2010] to name but a few.

An intuitive and flexible framework to capture such complex systems or networks is via graphs. A graph is a mathematical structure in which nodes, representing the system's units, are connected by edges, representing the units' interactions (Fig. 3.2F; [West, 2001; Newman, 2009]). While often the terms network and graph are used synonymously, we refer to "network" as a modeling framework for the underlying system and to "graph" as its mathematical representation. Since nodes and edges of a graph may be assigned different attributes, the graph flexibly captures various aspects of the underlying system. For example, in spatially embedded systems, the nodes may be assigned coordinates (e.g., crossings or hubs in transportation networks). Similarly, the edges may be assigned weights and directions (e.g., reflecting speed limits and directions of travel in transportation networks). Thereby, the network framework and its graph-theoretical basis enable mathematically powerful and quantitative analyses. Due to these benefits of the network framework, it has found application in diverse fields of research, ranging from data mining [Han et al., 2011] and dynamical systems theory [Boccaletti et al., 2006; Newman, 2009] to systems biology [Fell and Wagner, 2000; Barabási and Oltvai, 2004; Tero et al., 2010] and even cultural studies [Schich et al., 2014].

3.7 State of the art in theoretical modeling of actin-based transport

We distinguish three classes of approaches to model and understand cytoskeletal organization and transport on a systems level, i.e., mechanistic, fluid-dynamical, and network-based models (Fig. 3.1F; [Vaziri and Gopinath, 2008; Banerjee and Park, 2015]). As described in the previous section, any description of a given complex system is subject to a trade-off between the level of details included in the description and its scope. This trade-off is also reflected by the different classes of models, whereby mechanistic models typically incorporate the highest level of detail, while network models achieve the highest level of abstraction. While, ultimately, multi-scale models spanning all these levels are desired, they face the challenges already present at the respective levels and are computationally not yet feasible.

The models in the first class capture the different cytoskeletal molecular components as well as their biophysicochemical interactions. While such mechanistic models have successfully reproduced the self-organized formation of aligned microtubule arrays in plant cells [Hawkins et al., 2010; Tindemans et al., 2010], the rapid and diverse dynamics of the actin cytoskeleton still impede such modeling approaches in this setting. In particular, since microtubules in plant cells play a minor role in cellular transport, no mechanistic models for cytoskeletal transport are currently available. Generally, the high level of biological detail renders mechanistic models of whole-cell cytoskeletal systems computationally expensive. Moreover,

mechanistic simulations require many biological parameters whose experimental measurement is often challenging.

The models in the second class capture both cytoplasmic streaming and cytoskeletal components as continuous, interacting flow fields [Woodhouse and Goldstein, 2013; Goldstein and van de Meent, 2015]. These fluid dynamical models have been used to analyze the interplay between cytoplasmic streaming and actin organization in algae cells, revealing the emergence of self-organized, rotational streaming patterns [Woodhouse and Goldstein, 2013; Goldstein and van de Meent, 2015]. Since this approach does not incorporate molecular interactions, it is typically less parameter-intensive and the fluid-dynamical differential equations can be solved using established and efficient solvers [Versteeg and Malalasekera, 2007]. However, fluid-dynamical approaches neglect the discrete, filamentous structure of the cytoskeleton. Consequently, they do not incorporate microscopy data of actual cytoskeletal structures. Moreover, while current fluid-dynamic approaches assume a single flow field, the cytosol and different types of cargo may show very different patterns of movement (cf. Section 3.3).

In the third class, a number of network-based models have been developed that take into account the discrete, filamentous structure of the cytoskeleton (cf. Section 3.6). In a random network of filament segments, different regimes of transport were found, depending on the contribution from diffusion or motor-protein driven transport [Neri et al., 2013]. In a similar model, high transport rates of were shown to depend rather on filament polarities than orientations [Ando et al., 2015]. Moreover, the impact of motor protein movements on the surrounding cytoplasm was demonstrated in a model of two parallel lattices [Houtman et al., 2007]. A major advantage of these network-based model is the low number of employed biological parameters. Furthermore, these abstract models typically consider the movement of motor-proteins only and it is not clear whether their findings are transferable to real-life transport of cargo. This uncertainty arises from incomplete knowledge about interactions of motor proteins and cargo [Cai and Cresti, 2012; Buchnik et al., 2014], the non-trivial cooperation and competition of multiple motor proteins [Badoual et al., 2002; Müller et al., 2008; Guérin et al., 2010], as well as the lack of accompanying experimental data of cellular transport. Even more severely, the current network-based studies did not incorporate biological data of cytoskeletal structures.

3.8 Methods 3: Extraction of system-level network representations from image data

On the one hand, image analysis enables key insights in cell biology; however, capturing and understanding how complex cellular features such as cytoskeletal organization emerge remains challenging (cf. Sections 11 and 12). On the other hand, networks provide suitable representations of the cytoskeleton; but so far, biological data about real-life cytoskeletal structures have not been incorporated in this modeling framework (cf. Sections 13 and 13). Therefore, a combination of imaging data and network representations is needed for an accurate and quantitative description of the cytoskeleton and its functions.

Currently there is a number of different approaches for the reconstruction of biological networks from image data that may be divided into two classes: The first class of approaches relies on global image segmentation

to find the center lines of the imaged network structure, and consequent extraction of the network edges and nodes as center line segments and segment end points, respectively. The second class of approaches employs local optimization schemes to connect given seed points on the imaged network structure via paths that follow the underlying network filaments, and consequent connection of touching filaments. Generally, the existing methods for image-based network extraction are designed for specific types of networks. For example, methods from the first class have been developed for the extraction of tree-like networks from plant root systems [Pound et al., 2013], fungal networks [Baumgarten and Hauser, 2012; Obara et al., 2012b], or leaf venation networks [Dhondt et al., 2012]. Methods from the second class have been prominently developed for the reconstruction of neuronal topologies, or the connectome [Mayerich and Keyser, 2008; Meijering, 2010; Longair et al., 2011; Peng et al., 2015b], and a few studies have addressed extraction of cytoskeletal networks [Smith et al., 2010; Xu et al., 2011, 2014, 2015].

However, many of these approaches require user input, rendering them unfeasible for fast and unbiased large-scale analyses [Smith et al., 2010; Dhondt et al., 2012; Pound et al., 2013; Longair et al., 2011]. At the same time, fully automated approaches are usually tailored to specific image sources and challenged by low signal-to-noise ratios, as observed for the rapidly rearranging cytoskeleton in living plant cells (cf. Fig. 3.1A; [Sheahan et al., 2004; Riedl et al., 2008; Staiger et al., 2009]), which may strongly affect the resulting networks. Furthermore, some of the methods for extraction of cytoskeletal networks from image data are only feasible for small networks [Smith et al., 2010; Xu et al., 2011] and all of them rely on local instead of global optimization schemes [Smith et al., 2010; Xu et al., 2011, 2014, 2015]. As a result, the outcome of these approaches may be biased with respect to the selection of seed points as well as the definitions of local filament properties that are used in the optimization procedure. In particular, the existing methods for the extraction of networks from image data do not incorporate the underlying image intensity as an estimator of actin bundle thickness that is otherwise concealed by the diffraction limit of light microscopy (cf. Fig. 3.1A; [Sampathkumar et al., 2011; Moseley, 2013; Li et al., 2015]).

3.9 Research aims

We conclude that, complementary to a better understanding of the molecular mechanisms underlying cytoskeletal architecture and functionality, there is need for a system-level approach to the cytoskeleton that addresses the following five issues:

1. Description of the discrete filamentous structure of cytoskeletal networks *in vivo* on a system-wide, cellular level.
2. Understanding of biological principles underlying the organization of the actin cytoskeleton.
3. Identification of individual actin filaments as the basic unit of the actin cytoskeleton.
4. Quantification of the actin-based cellular transport dynamics of organelles.
5. Investigation of the coordination of organelle transport by the actin cytoskeleton.

Resolving these issues will be an important step towards a complete understanding of cytoskeletal transport, from its cellular coordination to its evolutionary context and its future technological control and utilization.

4 Results

This chapter comprises five of our publications that address and resolve the issues described in the Introduction (cf. Chapter 3) in order to shed light on the plant cytoskeleton as a transportation network. To that end, we developed and employed a robust image-based grid approximation of cytoskeleton networks, revealing organizational principles of the cytoskeleton (Section 4.1, [Breuer et al., 2014]). To simplify the extraction of networks from images, we presented an open-source software tool that implements the developed grid-based method (Section 4.2, [Breuer and Nikoloski, 2014]). We used multi-objective optimization combined with image data of different plant cell types and demonstrated that different cell types sustain cytoskeletal networks with cell-type specific organization (Section 4.3, [Breuer and Nikoloski, 2015a]). We devised a method to accurately segment and represent the cytoskeleton as a network and used automated tracking of organelles to show how the system-wide organization of the actin cytoskeleton drives cellular transport (Section 4.4, [Breuer et al., 2015]). Moreover, for a given network representation of any filamentous system, we developed an optimization-based approach to detect individual filaments, enabling non-invasive investigation of individual filament properties in their native network context (Section 4.5, [Breuer and Nikoloski, 2015b]). Overview of included publications:

4.1 Quantitative analyses of the plant cytoskeleton reveal underlying organizational principles
David Breuer, Alexander Ivakov, Arun Sampathkumar, Florian Hollandt, Staffan Persson, Zoran Nikoloski
J. R. Soc. Interface, 2014, 11(97):20140362

4.2 img2net: Automated network-based analysis of imaged phenotypes
David Breuer, Zoran Nikoloski
Bioinformatics, 2014, 30(22):3291-3292

4.3 Cell type-specific organization and optimality of the plant actin cytoskeleton
David Breuer, Yi Zhang, Staffan Persson, Zoran Nikoloski
in preparation

4.4 System-wide organization of the actin cytoskeleton drives organelle transport in plant cells
David Breuer, Alexander Ivakov, Jacqueline Nowak, Staffan Persson, Zoran Nikoloski
submitted

4.5 DeFiNe: an optimization-based method for robust disentangling of filamentous networks
David Breuer, Zoran Nikoloski
Sci. Rep., 2015, 5:18267

4.1 Quantitative analyses of the plant cytoskeleton reveal underlying organizational principles

Publication: *J. R. Soc. Interface, 2014, 11(97):20140362*

Authors: David Breuer[1,2], Alexander Ivakov[3], Arun Sampathkumar[4], Florian Hollandt[1], Staffan Persson[3,5], Zoran Nikoloski[1,*]

Affiliations: [1]Systems Biology and Mathematical Modeling, Max Planck Institute of Molecular Plant Physiology, Am Muehlenberg 1, 14476 Potsdam, Germany

[2]Institute of Biochemistry and Biology, University of Potsdam, Karl-Liebknecht-Straße 24-25, 14476 Potsdam, Germany

[3]Plant Cell Walls, Max Planck Institute of Molecular Plant Physiology, Am Muehlenberg 1, 14476 Potsdam, Germany

[4]Sainsbury Laboratory, University of Cambridge, Bateman Street, Cambridge CB2 1LR, United Kingdom

[5]ARC Centre of Excellence in Plant Cell Walls, School of Botany, University of Melbourne, Grattan Street, Parkville, Victoria 3010, Australia

Contact: *nikoloski@mpimp-golm.mpg.de

4.1.1 Abstract

The actin and microtubule cytoskeletons are vital structures for cell growth and development across all species. While individual molecular mechanisms underpinning actin and microtubule dynamics have been intensively studied, principles that govern the cytoskeleton organization remain largely unexplored. Here, we captured biologically relevant characteristics of the plant cytoskeleton through a network-driven imaging-based approach allowing to quantitatively assess dynamic features of the cytoskeleton. By introducing suitable null models, we demonstrate that the plant cytoskeletal networks exhibit properties required for efficient transport, namely, short average path lengths and high robustness. We further show that these advantageous features are maintained during temporal cytoskeletal re-arrangements. Interestingly, man-made transportation networks exhibit similar properties, suggesting general laws of network organization supporting diverse transport processes. The proposed network-driven analysis can be readily used to identify organizational principles of cytoskeletons in other organisms.

Keywords: cytoskeletal networks, cytoskeletal transport, plant cell walls, complex networks, organizational principles

4.1.2 Introduction

Complex systems can be represented by networks that capture the underlying components, as nodes, and their interactions, as links. Network representations have provided insights into the organizational principles of a variety of systems, ranging from man-made to systems shaped by evolution, such as: metabolic networks [Fell and Wagner, 2000; Jeong et al., 2000], neural networks [Felleman and Van Essen, 1991; Koch and Laurent, 1999], food webs [Williams and Martinez, 2000; Jordano et al., 2003], and transportation systems, including: vascular [Gazit et al., 1995; West et al., 1997] and leaf venation networks [West et al., 1999; Katifori et al., 2010].

The cytoskeleton represents yet another type of biological network. It is composed of actin filaments (AFs), microtubules (MTs), and intermediate filaments that form intricate interconnected arrays. Plant cells lack intermediate filaments, and their actin and microtubule cytoskeleton exhibits structural and functional differences to that of animal and yeast cells. These differences may be due to the presence of a rigid cell wall, a large central vacuole, the absence of discrete cytoskeleton organizing centers, or the general need of plants, as sessile organisms, to cope with changing environmental conditions [Wasteneys, 2000; Ehrhardt and Shaw, 2006].

In plant interphase cells, AFs exhibit extraordinarily dynamic behaviors [Staiger et al., 2009]. A major function of the actin cytoskeleton is to support cytoplasmic streaming, the directed flow of cytosol and organelles, which is mainly powered by ATP-driven myosin movement of compartments along the actin cytoskeleton [Shimmen and Yokota, 2004; Szymanski, 2009]. Furthermore, recent studies have shown that transportation of organelles depends on the micro-environment of the actin structures, where organelles are rapidly transported along actin bundles and display reduced motility when surrounded by thin AFs [Akkerman et al., 2011; Sampathkumar et al., 2013].

The behavior of single MTs, as well as MT arrays, has been described throughout the cell cycle and for different cell types [Wasteneys and Ambrose, 2009; Buschmann et al., 2011; Jacques et al., 2013b]. Their dynamics has been well-characterized [Hush et al., 1994; Ehrhardt and Shaw, 2006; Ehrhardt, 2008], and can lead to the formation of self-organized patterns that largely explain the MTs' orientation in growing cells [Zumdieck et al., 2005; Allard et al., 2010; Tindemans and Mulder, 2010; Lindeboom et al., 2013]. While MTs sustain vesicle motility in certain plant cells [Collings, 2008], they are typically located at the cell cortex and support the synthesis of cellulose microfibrils in interphase cells [Paredez et al., 2006]. Nevertheless, there is emerging evidence for transport along MTs also in these cell types, e.g., Golgi and small cellulose containing compartments have been reported to track along cortical MTs [Crowell et al., 2009].

Several studies have investigated the mechanical properties of the cytoskeleton in yeast and animal cells both experimentally [Wagner et al., 2006; Lieleg et al., 2007] and theoretically [MacKintosh et al., 1995; Wagner et al., 2006; Benetatos and Zippelius, 2007]. Models of AFs as a system of stiff, spring-connected rods have demonstrated a percolation-related transition in the viscoelastic properties [Ziemann et al., 1994; Forgacs, 1995], similar to signal propagation in a cytoskeleton model of connected rods [Shafrir et al., 2000; Shafrir and Forgacs, 2002].

The above studies employed a bottom-up approach in which the behavior of a system is explained based on the dynamics of its components. However, the interconnected structure and the rapid dynamics of the

cytoskeleton lend themselves to a top-down approach, which is independent of detailed molecular knowledge and better suited for uncovering the principles underlying cytoskeletal organization. Two studies have used such an approach in animal systems: AF arrays have been described as a superposition of different tessellation models [Fleischer et al., 2007] and a theoretical investigation of cytoskeletal transport in a system with passive diffusion and active transport along a random network of segments has demonstrated different regimes of transport [Neri et al., 2013]. Therefore, there is need for a network-based representation of the cytoskeleton that: (1) captures its complex network structure, (2) is based on biologically solid ground, (3) can be used to describe dynamic network processes, and (4) may uncover organizational principles of the cytoskeleton in plant and animal cells.

In this study, we propose a novel framework that captures the structure and dynamics of the actin and microtubule cytoskeleton as complex networks. We used this framework to quantify and compare the behavior of AFs and MTs in plant interphase cells under different conditions. We tested the hypothesis that the cytoskeleton is well-suited to support transport processes. By developing suitable null models as references, we show that the cytoskeleton indeed exhibits biologically desirable transport-related properties, such as short average path lengths and high robustness against disruptions of the network. Finally, we demonstrate that man-made transportation networks display similar properties. The developed framework is readily applicable to study the cytoskeleton of other organisms or under different conditions.

4.1.3 Results

Reconstruction of complex networks from cytoskeletal images

To investigate the networks of AFs and MTs, we grew Arabidopsis (*A. thaliana*) FABD:GFP and TUA5:mCherry dual-labeled seedlings [Sampathkumar et al., 2011] in the dark and imaged elongating hypocotyl cells. To capture rapid changes and to minimize bleaching, we used a spinning-disk confocal microscope (Fig. 4.1.1A and B; cf. Appendix 6.1.1). To generate complex networks from the cytoskeleton image series we followed a two-step procedure: We placed a grid over the cytoskeleton which covers the cell's cytoskeleton (e.g. Fig. 4.1.1C). From the grid, we constructed an edge-weighted network in which nodes represent the grid's junctions, and edges represent the grid's links. We assigned a weight to each edge by creating convolution kernels with Gaussian profiles for each edge (Fig. 4.1.1D), thus projecting the cytoskeleton onto the overlaid grid. This results in a weighted, undirected network (Fig. 4.1.1E) where the weights reflect the intensity of the underlying filaments/bundles. Using confocal z-stack image series, these steps were also used to construct three-dimensional cytoskeletal networks (Fig. 4.1.1G). The procedure was repeated for all images of the recorded actin and microtubule time series, separately. As a result, each network captures information of the time-dependent cytoskeletal component whose properties may be readily investigated.

To determine if the studied network properties carry a biological signal, we developed several null models that randomize parts of the cytoskeletal structures while preserving the total amount of cytoskeleton in the cell (cf. Appendix 6.1.2). If a given network property is significantly higher or lower than expected by chance we conclude that the underlying cytoskeletal organization is non-random and, therefore, biological relevant. This may suggest that the cytoskeleton is tuned to guarantee such values of the structural or functional network property.

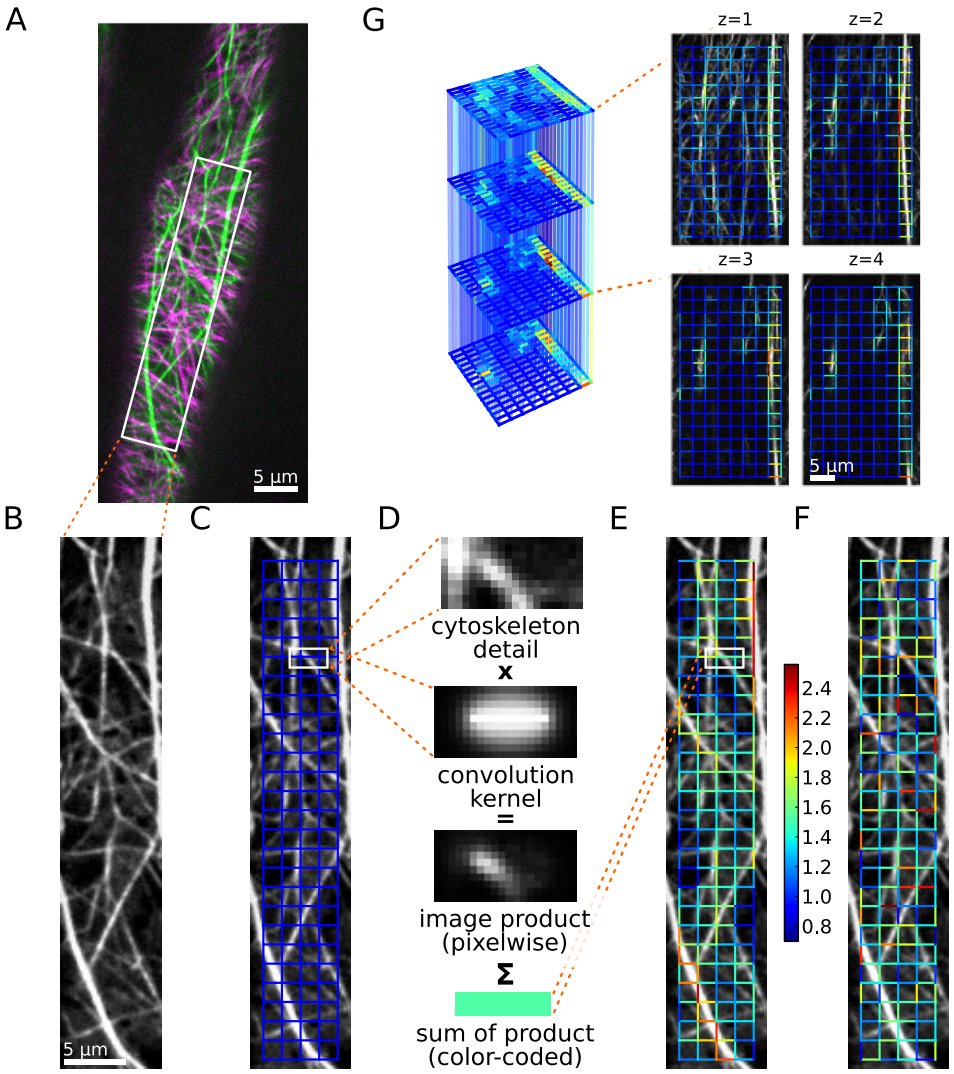


Figure 4.1.1: **From fluorescence image to reconstructed network, null model, and 3D extension.** **(A)** Colored overlay of unprocessed snapshots of the AFs (green) and MTs (magenta) of a dual-labeled, three-day-old *A. thaliana* hypocotyl cell. **(B)** Single preprocessed image of one cytoskeletal component, here AFs. **(C)** Grid used for network reconstruction (a uniform rectangular grid with 10 pixels spacings is shown in blue). **(D)** Using convolution kernels, the links of the grid are assigned scalar values by pixelwise multiplication of the kernel with the cytoskeleton image and subsequent summation. **(E)** Weighted, undirected network with edges given by the links of the chosen grid type and weights obtained via the kernel method (weights are color-coded from blue to red). **(F)** To assess the biological relevance of various properties of the cytoskeletal network, a null model is introduced through an ensemble of networks with shuffled edge weights (one exemplary realization is depicted). **(G)** From confocal z-stack recordings, a three-dimensional cytoskeletal network is reconstructed (grid spacings are 20 pixels; edges connecting different z-layers are set transparent for better visibility of the full network).

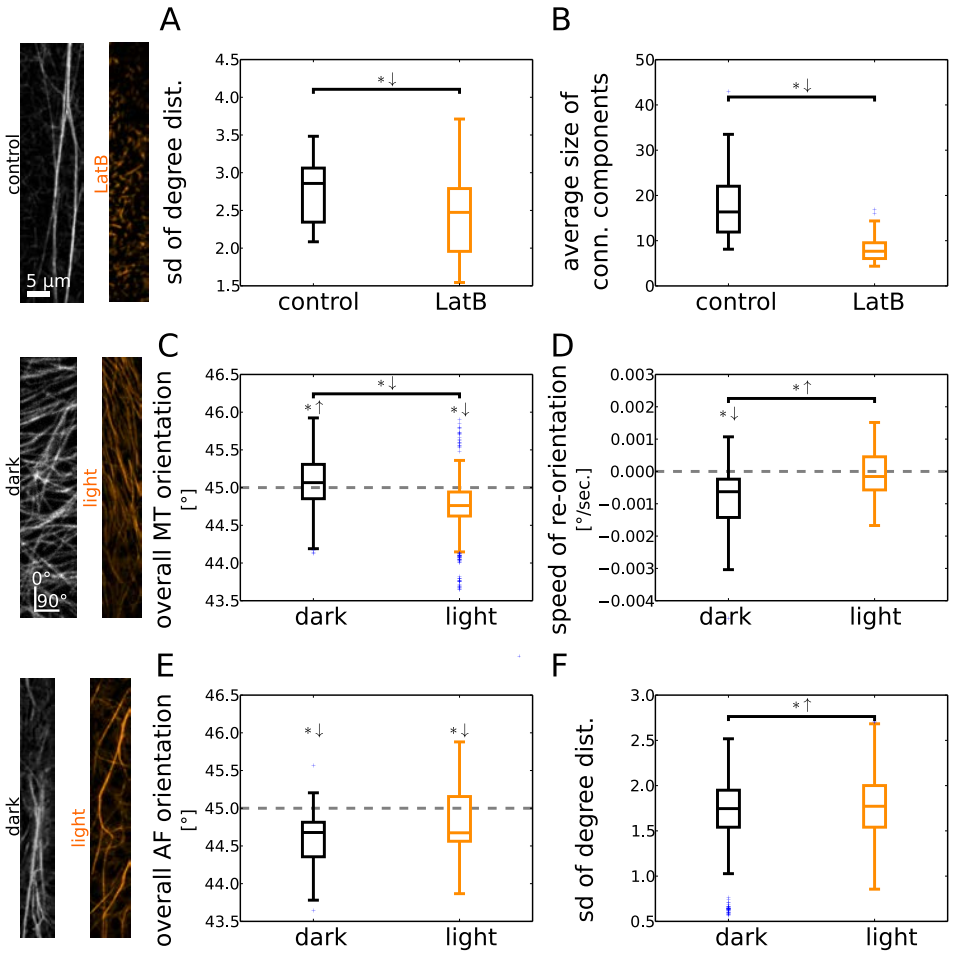


Figure 4.1.2: **Network properties capture biologically relevant aspects of cytoskeletal organization for different scenarios.** The first 20 frames of the image series are used for the analysis. $*\downarrow$ and $*\uparrow$ above a bar denote a decrease or increase of the properties of the treated relative to those of the control plants (independent two-sample t-test: $p-$value < 0.05). $*\downarrow$ and $*\uparrow$ above a box plot denote network properties that fall below or exceed a reference value marked by a gray, dotted line (one-sample two-sided t-test: $p-$value < 0.05). **(A)** The AF network of Latrunculin B-treated plants displays a smaller standard deviation of the network's degree distribution. **(B)** The connected patches of AFs are smaller after Latrunculin B treatment. **(C)** The orientation of MTs is predominantly horizontal in dark-grown plants and vertical in light-exposed plants. **(D)** Computing the change in MT orientation per unit time shows a difference between control and light-treated plants. In dark-grown plants, a significant change towards a vertical orientation is observed, which is absent for light-treated plants. **(E)** The (horizontal) orientation of AFs is not altered in plants exposed to light. **(F)** Light induces a dispersion of AFs which yields a broader degree distribution.

The reconstructed networks capture biologically relevant features of the actin and microtubule cytoskeletal components

To test whether the proposed network-based approach captures biologically meaningful features, we used chemical treatments and environmental stimuli to alter the behavior of the cytoskeleton. First, we quantified the effect of the actin-disrupting drug Latrunculin B on the actin cytoskeleton. This drug binds to monomeric actin and, thereby, inhibits AF formation [Yarmola et al., 2000]. We reconstructed the AF networks for both the control and the treated plants for each frame of the image series. The structure of the AFs and their drug-induced fragmentation was quantified by two network properties which can be related to the biological phenomenon (see Appendix 6.1.3 for a mathematical description and detailed interpretation of these quantities): The standard deviation of the degree distribution [West, 2001] captures the spatial heterogeneity of the distribution of actin structures, i.e., images with regions of low and high cytoskeletal intensities yield both small and large edge weights and consequently a broader degree distribution (cf. Fig. 4.1.1E). Since the edge weights integrate intensities of possibly multiple filaments, our approach does not resolve differences in thicknesses or numbers of individual filament but only a combination thereof. By comparing the standard deviations of the degree distributions of control and treated plants, we found a statistically significant reduction by Latrunculin B (Fig. 4.1.2A; independent two-sample t-test: $p-\text{value} = 7.0 \cdot 10^{-9}$; for treated and non-treated plants, respectively, we pooled the standard deviations of the degree distributions across the first 20 time points of the image series). We then determined the average number of nodes per connected, non-trivial network component after thresholding the edge weights, providing an estimate for the extent to which the cytoskeletal filaments form connected networks. By using the 50th percentile as a threshold, we found that Latrunculin B reduces the average size of the resulting connected components (Fig. 4.1.2B; independent two-sample t-test: $p-\text{value} = 2.9 \cdot 10^{-42}$). These findings are in agreement with visual reports on the fragmented actin structure of Latrunculin B-treated cytoskeletons [Yarmola et al., 2000].

By using the reconstructed MT network, we quantified the overall orientation of MTs in plants that had been exposed to light several hours before imaging. Light is one of the environmental factors that determine plant growth, and it is well established that the MT array rapidly changes from largely transverse to a generally longitudinal when seedlings are exposed to light [Wymer and Lloyd, 1996; Paradez et al., 2006; Sambade et al., 2012]. As our method does not detect individual filaments, we inferred the MT orientation indirectly (see Appendix 6.1.4 for a detailed derivation): By placing an imaginary rod of a specific length and orientation over the grid we calculated its contributions to the weights of edges with different orientations by using our kernel method (cf. Fig. 4.1.1D). Here, we solved the inverse problem to obtain the overall MT orientation α from the weight distribution of edges with different orientations . Angles $\alpha \in [0°, 45°)$ and $\alpha \in (45°, 90°]$ indicate overall vertical and horizontal orientations of the MTs, respectively. We estimated the MT orientation for seedlings grown under dark and light conditions and found a significant difference (Fig. 4.1.2C; independent two-sample t-test: $p-\text{value} = 5.8 \cdot 10^{-52}$) with a horizontal and longitudinal orientation, respectively (one-sample two-sided t-tests: dark $p-\text{value} = 6.3 \cdot 10^{-8}$, light $p-\text{value} = 4.7 \cdot 10^{-45}$). These findings are in agreement with known results [Granger and Cyr, 2001; Paradez et al., 2006; Sampathkumar et al., 2011; Sambade et al., 2012]. They further revealed that despite the strong correlation between light exposure and longitudinal MT orientation, there are also deviants (cf. Fig. 4.1.2C): Under dark condition, a fraction of about 40% of the MT networks shows an, unexpected, overall vertical orientation while under light conditions,

about 20% of the analyzed MT networks display an overall horizontal orientation, contrary to expectations. . These deviations from the expected results may highlight the inherent variability of cytoskeletal responses to external stimuli and support the view that hypocotyl cells may be in different stages of growth [Gendreau et al., 1997; Le et al., 2005; Sambade et al., 2012].

We also studied the speed of MT reorientation under the microscope by computing the slope of the average orientation time series via a linear regression. There was a significant difference between the two treatments (Fig. 4.1.2D; independent two-sample t-test: $p-$value $= 2.5 \cdot 10^{-3}$), i.e., while the change in orientation is negative in dark, it does not significantly differ from zero in light (one-sample two-sided t-test: dark $p-$value $= 2.5 \cdot 10^{-5}$, light $p-$value $= 0.6$). Therefore, we conclude that over the range of five minutes of the experiment the confocal laser light does induce a reorientation of the MTs towards the longitudinal cell axis in dark-grown plants; however, this is not the case in plants exposed to light before imaging as reorientation of MTs has already progressed further in these cells.

Next, we employed our network-based framework to investigate the behavior of AFs in response to light in growing hypocotyls. Like for MTs, we inferred the overall orientation of the AFs for dark-grown and light-treated plants, respectively. We found no significant difference in actin orientation between the treatments (Fig. 4.1.2E; independent two-sample t-test: $p-$value $= 6.6 \cdot 10^{-2}$), with a consistent major longitudinal orientation (one-sample two-sided t-test: dark $p-$value $= 3.8 \cdot 10^{-7}$, light $p-$value $= 4.4 \cdot 10^{-2}$). To quantify the heterogeneity of the actin distribution, we computed the standard deviation of the degree distributions for both scenarios. In light-treated plants, the actin cytoskeleton displayed a more heterogeneous distribution across the cell than in dark-grown plants (Fig. 4.1.2F; independent two-sample t-test: $p-$value $= 1.5 \cdot 10^{-4}$), implying the prevalence of bundles. These findings agree with reports on the impact of light on the organization of AFs in maize coleoptiles [Waller and Nick, 1997]. However, they do not agree with the qualitative findings in a different species, i.e., rice, where light was shown to promote a change in AF orientation from transverse to longitudinal and to disperse actin bundles [Holweg and Nick, 2004]. Interestingly, the rearrangement of AFs under light has been linked to that of MTs [Sampathkumar et al., 2011]. Therefore, our findings suggest that a change in environmental conditions would impose a need for rapid redistribution of cellular material in the cell, which are known to be facilitated by actin bundles [Akkerman et al., 2011].

Accessibility and robustness of cytoskeletal networks

After demonstrating the inherent ability of our network approach to capture biologically relevant information on cytoskeletal organization, we focused on identifying network properties that reflect the functions of the cytoskeleton. To investigate the transport efficiency of the AF and MT networks (Fig. 4.1.3A), we computed average path length (as a measure for the cellular accessibility of the cytoskeleton) and algebraic connectivity (as a measure for the cytoskeletons' robustness against disruptions) (see Appendix 6.1.3 for the mathematical formulation and a detailed interpretation of the properties).

The average (shortest) path length (APL) [West, 2001] is the average of the minimum distances between all pairs of nodes in a (edge-weighted) network. Here, the length of an edge is given by the inverse of its weight, i.e., thick actin bundles or tubulin filaments yield small edge lengths. This is reasonable since cytoskeletal bundles typically facilitate faster transport compared to thinner filaments [Akkerman et al., 2011] which may, in general, depend on the size of the cargo. The APL provides an estimate of how close any two nodes

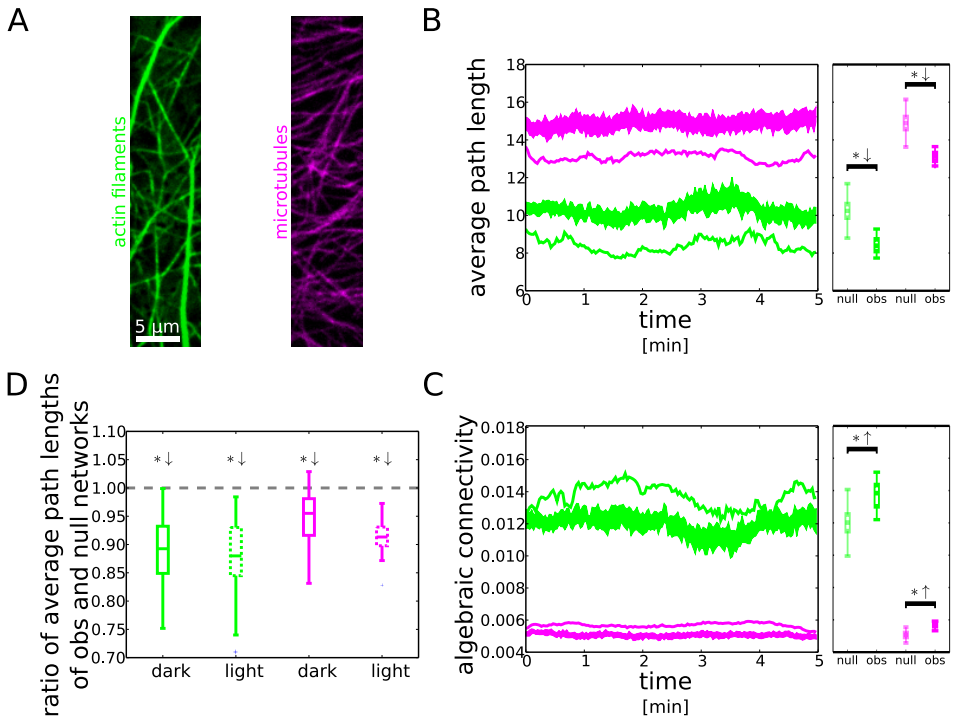


Figure 4.1.3: **Time-resolved average path length and algebraic connectivity of cytoskeletal networks and null model networks.** In (B) and (C), the results for the observed networks (solid lines) of AFs (green) and MTs (magenta) are compared to those of the null model (dashed lines: ensemble-mean; shaded regions: ensemble-mean ± standard deviation). The box plots show the distribution of values of a network property used in the statistical test. **(A)** Green- and magenta-colored images of AFs and MTs, respectively. **(B)** The average path lengths of AF and MT networks (solid lines) fluctuate over time and stay well below the average path lengths of the null model (dashed lines and shaded regions). **(C)** The algebraic connectivity is consistently larger for both AFs and MTs in the observed cytoskeletal networks (solid lines) than in the null model networks (dashed lines and shaded regions). **(D)** Comparing the ratios of average path lengths of the observed networks and their null model networks for AFs and MTs yields no statistically significant difference between dark and light conditions.

are expected to be and, hence, the accessibility in the cytoskeleton. By computing the APLs for the sequence of the AF and MT networks, we obtained two time series (Fig. 4.1.3B, green and magenta, solid lines). The corresponding values largely reflect the overall intensity distribution of the images and, by themselves, carry little information about the underlying network structure. As a reference, we calculated the APLs for ensembles of AF or MT null model networks, i.e., networks obtained by shuffling edge weights (Fig. 4.1.3B, green and magenta, dashed lines and shaded regions). We found that the observed networks exhibit significantly smaller APLs than their respective null models (independent two-sample t-test: AF $p-$value = $7.8 \cdot 10^{-273}$, MT $p-$value < $2.2 \cdot 10^{-308}$).

The algebraic connectivity (AC) [West, 2001] is the second smallest eigenvalue of the network's graph Laplacian, which is closely related to the weight matrix of the network, and reflects how well-knit the network is. While a vanishing AC indicates the decomposition of the network into two or more disconnected components, larger values correspond to a higher robustness of the network against disruptions. By comparing the AC of the observed actin and microtubule networks (Fig. 4.1.3C, green and magenta, solid lines) to their null model counterparts (Fig. 4.1.3C, green and red, dashed lines and shaded regions), we found that the observed networks yield significantly larger algebraic connectivities than their respective null model networks (independent two-sample t-test: AF $p-$value = $7.5 \cdot 10^{-139}$, MT $p-$value = $2.9 \cdot 10^{-278}$).

These findings may be graphically explained as follows: Networks of MTs, and even more so AFs, possess filaments and bundles that stretch across large parts of the cell. These structures establish connected paths in the networks with large weights and therefore small lengths. In the computation of shortest path lengths, they act as "highways" that efficiently connect spatially distant regions. Furthermore, the AF and MT networks exhibit larger regions that are particularly strongly linked and result in a higher robustness of the networks against disruption. To support this interpretation, we computed the degree assortativity [Newman, 2009] given by the correlation between degrees of nodes and those of their neighbors. It quantifies the extent to which nodes of (dis-)similar degree are connected to each other. Both the AF and MT networks exhibit significantly higher assortativity than their corresponding null model networks and are hence more spatially clustered (independent two-sample t-tests: AF $p-$value < $2.2 \cdot 10^{-308}$, MT $p-$value < $2.2 \cdot 10^{-308}$).We note that both APL and AC are summary statistics that do not capture differences in local connectivity patterns but reflect network properties that relate to the network's overall transport capacity.

Interestingly, despite the differences in the network architecture of AFs and MTs under dark and light conditions (cf. Fig. 4.1.2), there were no significant differences in the ratios of APLs of observed and null model networks (Fig. 4.1.3D). Moreover, these ratios stay consistently below one throughout the time series (one-sample two-sided t-tests: $p-$values < 0.05 for AF/MT dark/light), reflecting small effort for reaching any part of the cytoskeletal networks. Thus, the cytoskeleton preserves its advantageous transport properties over time and across conditions.

While the MT network can largely be captured at the cell cortex in interphase cells, the actin cytoskeleton constitutes a three-dimensional structure that spans the expanding cell. To ensure that we captured the volumetric behavior of the AFs, we also recorded confocal z-stack image series of such cells and used our framework to reconstruct the AF network as a three-dimensional network (cf. Fig. 4.1.1G). To assess if the additional information about AFs below the cortical plane changes the transport efficiency of the AF network, we compared the APL and AC of the three-dimensional network to that of the two-dimensional network ob-

tained by averaging the intensities of edges at the same x-y-position across all z-layers. In both cases, the ratio of network properties of the observed network and the null model networks stays well below one (one-sample two-sided t-test: $p-$values < 0.05 for 2D/3D APL/AC; see Appendix 6.1.2 for details). Analogously to the bundle structures in the 2D networks, the actin structures that reach deeper into the cell provide strong connections in the 3D reconstruction (see Fig. 4.1.1G) that naturally equip the network with shorter paths and higher robustness.

Taken together, AF as well as MT networks display highly non-random features. In particular, short APL (good accessibility) and large AC (high robustness) are preserved over time and across different environmental conditions. These findings provide quantitative support for the idea that plants reliably establish and maintain cytoskeletal structures that are optimized for transport processes throughout the cell [Shimmen and Yokota, 2004; Paredez et al., 2006; Sampathkumar et al., 2013].

The cytoskeleton and the German autobahn exhibit similar network properties

We next asked whether the observed efficiency in network properties is unique to the cytoskeleton or if it can also be observed in other transportation networks. As a prominent example, we generated images of the German autobahn with color-coded speed limits (Fig. 4.1.4A; The maximum speed limit is set to 200 km/h, the speed limit outside of the autobahn is set to 50 km/h. However, our findings are largely independent of this choice, see Appendix 6.1.5). By using our established framework, we obtained a weighted network for the autobahn (Fig. 4.1.4B) and corresponding null model networks by shuffling edge weights. For the APL and the AC, we computed the ratios for the respective network properties of the autobahn and its null model networks and we found that the autobahn network exhibits shorter APLs and higher algebraic connectivities than expected by chance (Fig. 4.1.4I, black). These results are similar to what we found for the cytoskeleton (Fig. 4.1.4I, green and magenta). Furthermore, we note that the degree distributions of the cytoskeletal networks and the autobahn are unimodal and peak around their means (Fig. 4.1.4E and F).

To differentiate these networks from networks with different structural and transport-related properties, we further studied a contrived network with a stronger local structure and weaker long-ranged connections (Fig. 4.1.4C). The contrived network displays a heavy-tailed degree distribution (Fig. 4.1.4G) as well as properties associated with poor transport efficiency, namely, longer APL and smaller AC than expected from the null model (Fig. 4.1.4I, orange).

Another interesting comparison is that of the cytoskeleton and the autobahn to networks in which one or several transport-related properties are optimized. The weight distribution of a network with a fixed sum of edge weights and maximal AC may be computed efficiently by solving a semi-definite optimization problem whose solution is unique [Boyd, 2006; Sun et al., 2006] (Fig. 4.1.4D). Such an optimally robust network outperforms the cytoskeletal and the autobahn networks by a factor of eight with respect to the ratio of AC of observed and null model networks (Fig. 4.1.4I, blue). However, it is less efficient in terms of its APL, which is higher than expected by chance, demonstrating a trade-off between different measures of network optimality.

While there are certainly differences in the structure as well as the function of the cytoskeleton and the autobahn, the network properties studied here are summary statistics and it is not possible to infer local

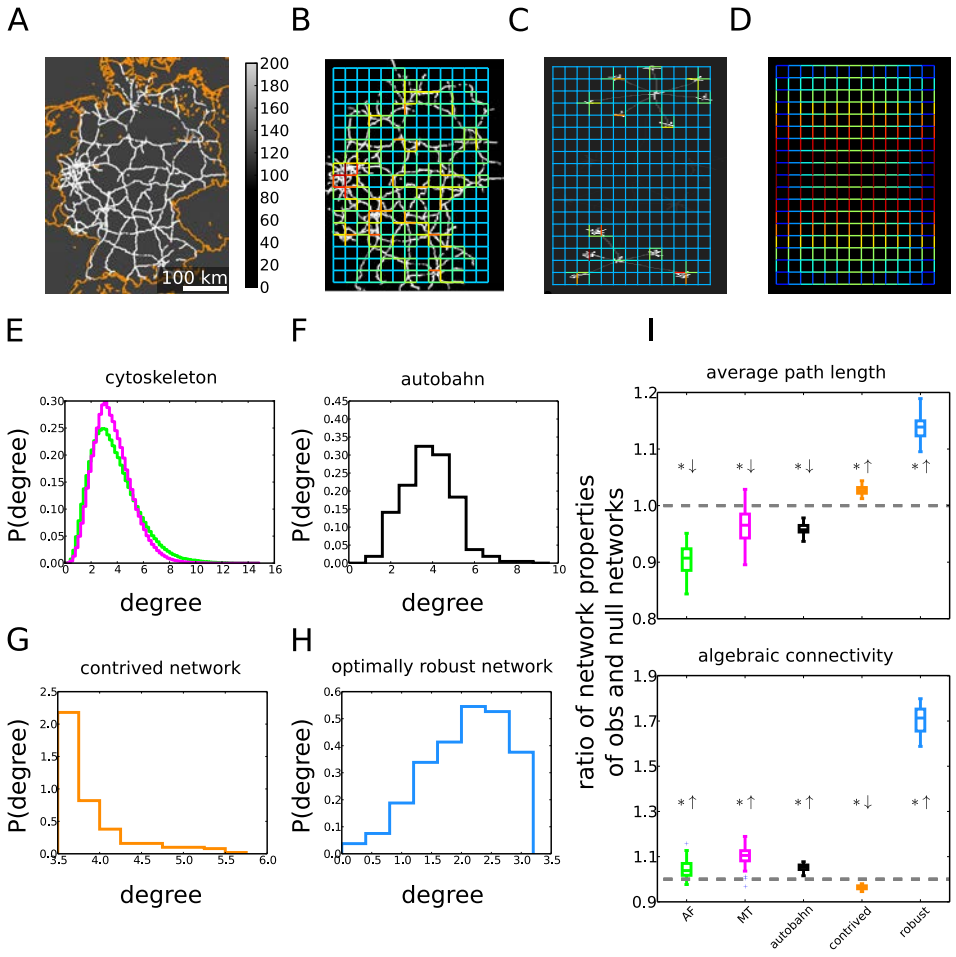


Figure 4.1.4: **Comparison of cytoskeleton, German autobahn, and other types of networks.** **(A)** The autobahn network of Germany with color-coded speed limits and national borders (orange) for guidance. **(B)** Network reconstructed from the autobahn image. **(C)** Reconstructed network for a contrived underlying structure with different structural and transport-related properties. **(D)** Network with maximal algebraic connectivity for the given grid-topology and normalized sum of weights (here, the network is constructed via an optimization procedure and not inferred from an underlying image intensity distribution; the black background is for better visibility only). **(E)** The degree distributions of the AF (green) and MT (magenta) networks are unimodal and peak around their means (excess kurtosis > 0; the resolution of the histograms is higher as they include all networks of an image series, cf. Fig. 4.1.3). **(F)** The degree distribution of the autobahn network is unimodal and peaks around its mean (excess kurtosis > 0). **(G)** For the contrived network, the degree distribution is also peaked (excess kurtosis > 0) but shows a more heavy right tail compared to (E) and (F). **(H)** The degree distribution of the optimally robust network is broader around the peak and has thinner tails (kurtosis < 0). **(I)** The ratios of average path lengths of observed and null model networks are below one for the cytoskeletal and the autobahn networks and above one for the contrived and the optimally robust network. The algebraic connectivities are bigger than expected by chance for all studied networks, except for the contrived networks that show a smaller algebraic connectivity. All deviations from the unit ratio are statistically significant (one-sample two-sided t-tests: all $p-$values < 0.05).

structural differences from them. In particular, we do not study their absolute values but the relative efficiency of these networks with respect to their respective null model which may point to its organizational principles.

To conclude, both the actin and microtubule cytoskeleton display characteristics typical of transportation networks, such as the autobahn, and exhibit structures which may not be aimed at optimizing a single property indicative of efficient transport. Our data therefore provide quantitative measures to support a view of the plant interphase cytoskeleton as an efficient transportation network.

4.1.4 Discussion

Though many studies have analyzed the cytoskeleton, most of them have relied on qualitative observations or manual tracking of up to some dozens of AFs and MTs. The rapid dynamics, as well as strong variability of cytoskeletal organization across different cell types, stages of cell life, and environmental conditions, necessitate a framework that allows for a fast and objective quantification of the cytoskeletal components in living cells. This would then allow for biologically meaningful interpretations that go beyond strictly theoretical studies to investigate the structure of the system. Here, we described one such framework that captures biologically relevant variations.

Many studies have used a bottom-up approach to the cytoskeleton, in which the molecular principles are presupposed and used to infer the behavior of the system. Instead, we pursued a top-down strategy to represent the cytoskeletal organization without the need for detailed molecular knowledge. Hence, our approach hints at the underlying organizational principles of the cytoskeleton. More specifically, by choosing a representation through complex networks, we could exploit the well-equipped toolkit from graph theory to investigate the structure of the cytoskeleton and its relation to, e.g., efficient transport processes in the cell.

For a careful interpretation of these findings, we need to bear in mind several points: (1) Our network reconstruction method creates nodes at positions given by the chosen grid (cf. Fig. 4.1.1). Hence, not all nodes correspond to crossings of the filaments. Moreover, despite our focus on the largely planar cortical cytoskeletal even apparently crossing filament maybe separated by hundreds of nanometers in z-direction. Using our current imaging techniques, such distances may not be resolved in the 3D reconstruction either (see outlook below). In particular, such distances prohibit the switching of motor proteins (cf. e.g. [Bálint et al., 2013] for MT). Yet, the edge weights in our method agglomerate local intensities that may originate from multiple filaments in different depths. More importantly, typical cargo such as mitochondria, Golgi bodies, or chloroplasts range from several 100 nm to several μm in size and may thus easily bridge even larger distances between filaments, thereby justifying our assumption of transport along edges and via nodes. (2) In our approach, all edges are undirected, i.e., they allow bi-directional transport. While bi-directional transport may occur along single actin or microtubular filaments, e.g., due to different motor proteins or fluctuations [Gross, 2004; Lee et al., 2004; Caviston and Holzbaur, 2006], bundles of filaments typically allow uni-directional only. This uni-directionality is further amplified by the cytoplasmic bulk flow generated by the coordinated movement of motor proteins [Shimmen and Yokota, 2004]. Yet again, our reconstruction methods assigns edge weights by integrating local intensities of possibly multiple filaments with different

orientations. Thus, fully bi-directional of transport is unlikely (but not excluded) and, since it can not be inferred from the cytoskeletal images alone, we use it as an approximation of the potential transport capacity. (3) This potential transport capacity is modeled to be higher in regions with many/thick cytoskeletal filaments, as described by the edge weights. However, the edge weights do not quantify the speed/amount of cargo that is really transported (and we do not measure it, see outlook below). (4) Finally, we note that AF and MT networks generally transport different cargo (cf. [Shimmen and Yokota, 2004; Paredez et al., 2006]), although there is evidence for transport of, e.g., small cellulose containing vesicle along both structures [Goode et al., 2000; Gutierrez et al., 2009]. In addition, different types of cells may require different modes of transportation [Hussey et al., 2006]. Here, we focused on the potential transport capacity of the cytoskeleton in interphase hypocotyl cells, but our framework may readily be used to study other scenarios.

There is a rich literature on the comparison of structures of different networks. Many biological and man-made networks show scale-free degree distributions, i.e., there are a few nodes with many neighbors [Bornholdt and Röhl, 2003], e.g., airway networks [Guimerà et al., 2005]. However, nodes in other transportation networks are restricted regarding the number of potential neighbors due to the physical limitations. Road and railway networks display degree distributions that peak around their average values [Barthélemy, 2011], which we also demonstrated for the cytoskeletal networks. Despite the apparently diverse principles underlying man-made transportation networks, studies have revealed strong agreement in a number of their properties, e.g., degree distribution. This agreement may be explained by costs associated with the establishment of new nodes and links [Barthélemy and Flammini, 2008; Courtat et al., 2011; Louf et al., 2013]. Our findings suggest that comparable cost-related restrictions may play a role in the formation of the cytoskeleton, leading to similar structures and transport properties as in man-made networks.

In summary, our framework captures the complex network structure of filamentous cytoskeletal components. We used this framework to derive organizational principles of the cytoskeleton. We further showed that AF and MT networks display biologically desirable characteristics, such as short APLs and high robustness, similar to characteristics found in non-biological transportation networks. In particular, these features of efficient transportation networks are maintained over time and across conditions.

Possible directions of future efforts are manifold: (1) Our framework can be employed to quantify the complex structures of AF and MT networks, and thus enables an automated and objective comparison of the complex structures of cytoskeletal networks in other biological systems, e.g., focusing on the cytoskeleton connecting the nucleus to other parts of the cell. (2) The resolution of the fine cytoskeletal structures may be improved by using more advanced imaging techniques like total internal reflection fluorescence microscopy, at least for the cortical cytoskeleton. (3) Another promising direction is the comparison of reconstructed cytoskeletal networks to networks that optimize one or several seminal network properties. As different network structures favor specific properties, the cytoskeleton may represent an evolutionarily shaped compromise between them. While such a balance has been suggested, e.g., between the speed and the sensitivity in the polarization of the cytoskeleton [Hawkins et al., 2010], quantitative evidence for a trade-off in the cytoskeleton's transport properties is lacking. We note that besides its vital role in cellular transport processes the plant cytoskeleton strikingly determines the mechanical properties of the cell. (4) Finally, our work paves the way for direct studies of the cytoskeleton as a transportation network. Employing actin and organelle dual-labeled plants, it is appealing to correlate actual biological transport processes with flow-related network measures. While several studies have investigated the transport of organelles and vesicles

along the cytoskeleton [Goode et al., 2000; Rogers and Gelfand, 2000; Shimmen and Yokota, 2004; Bálint et al., 2013], none have quantitatively linked it to the complex structure of the cytoskeletal network. Answering these questions may contribute to a better understanding of the organizing and dynamic principles of the cytoskeleton.

4.1.5 Materials and Methods

The experiment setup includes dual-labeled *Arabidopsis thaliana* Columbia-0 seedlings to which different treatments were applied prior to imaging in a spinning disk confocal microscope setup. For further details, refer to Appendix 6.1.1. The computational network-based investigation of the image series, as illustrated in Fig. 1, includes: (1) the preprocessing of the images in Fiji [Schindelin et al., 2012], (2) the creation and quantification of the weighted cytoskeletal and null model networks, and (3) their statistical analyses in Python [Van Rossum and Drake, 2011] (using SciPy [Olivier et al., 2002], NumPy [Oliphant, 2006], NetworkX [Hagberg et al., 2008] and the Matplotlib [Hunter, 2007] libraries). The construction of an optimally robust network was performed by solving a semi-definite optimization problem using the Cvxopt Python package [Dahl and Vandenberghe, 2006]. Detailed descriptions of these steps are given in Appendix 6.1.2 and the studied network properties are described in detail in Appendix 6.1.3. The overall orientation of cytoskeletal components is inferred from the network's weight distribution as described in Appendix 6.1.4. The data of the German autobahn, as depicted in Fig. 4.1.4, are collected from OpenStreetMap and filtered as explained in Appendix 6.1.5.

4.1.6 Acknowledgments

Support: We thank Dr. Tijs Ketelaar and Dr. Georg Basler for valuable comments on the manuscript.

Funding: D.B., A.I., S.P., and Z.N. were supported by the Max Planck Society.

Contributions: A.I., F.H., S.P., and Z.N. designed the study; D.B., A.I. and F.H. performed the experiments; D.B. implemented the method and analyzed the data; D.B., A.I., S.P., and Z.N. wrote the manuscript.

4.2 img2net: Automated network-based analyses of imaged phenotypes

Publication: *Bioinformatics, 2014, 30(22):3291-3292*

Authors: David Breuer[1,2,*], Zoran Nikoloski[1]

Affiliations: [1]Systems Biology and Mathematical Modeling, Max Planck Institute of Molecular Plant Physiology, Am Muehlenberg 1, 14476 Potsdam, Germany

[2]Institute of Biochemistry and Biology, University of Potsdam, Karl-Liebknecht-Straße 24-25, 14476 Potsdam, Germany

Contact: *breuer@mpimp-golm.mpg.de

4.2.1 Abstract

Summary: Automated analysis of imaged phenotypes enables fast and reproducible quantification of biologically relevant features. Despite recent developments, recordings of complex, networked structures, such as: leaf venation patterns, cytoskeletal structures, or traffic networks, remain challenging to analyze. Here we illustrate the applicability of img2net to automatedly analyze such structures by reconstructing the underlying network, computing relevant network properties, and statistically comparing networks of different types or under different conditions. The software can be readily used for analyzing image data of arbitrary 2D and 3D network-like structures.

Availability and Implementation: img2net is open-source software under the GPL and can be downloaded from http://mathbiol.mpimp-golm.mpg.de/img2net/, where supplementary information and data sets for testing are provided.

Keywords: image processing, networks, phenotyping, cytoskeleton

4.2.2 Introduction

Biological and man-made systems, ranging from biochemical reactions to neural and social interactions, can often be represented as networks, with nodes and edges representing the components and their interactions, respectively. Network representations facilitate not only intuitive visualization, but also quantitative studies of the systems' structure and dynamics.

Spatial networks constitute an import subclass which includes networks, such as: leaf venation [Dodds et al., 2010], cellular cytoskeleton [Volkmann and Baluska, 1999], and city streets [Barthélemy, 2011]. While some of these networks, like city infrastructure, have been well-characterized, others, like biological spatial networks, remain poorly understood.

Snapshots of spatial networks can be obtained by imaging technologies. The computational challenge is that of extracting the the underlying networks from the gathered images in a fast and reliable fashion, and of examining the reconstructions to reveal the underlying organizational principles.

The existing image-based methods for reconstruction of biological networks are typically designed for specific types of networks: Tree-like networks for plant root architectures [Pound et al., 2013]; fungal or leaf venation networks [Obara et al., 2012a]; or neuronal topology of the human connectome [Meijering, 2010; Longair et al., 2011]. While most of these approaches require user input, rendering them unfeasible for high-throughput studies, fully automated algorithms are usually tailored to specific image sources and challenged by low signal-to-noise ratios, which may strongly affect the resulting networks.

Here, we present a robust approach to reconstruct 2D and 3D (non-)biological spatial networks from gray-scale image data. By constructing weighted networks and computing their seminal network properties, img2net allows an extensive quantification and statistical comparison of network topologies.

4.2.3 Methods and Functionality

We extract the networks from image data in two steps: Starting from a gray-scale image, where high intensities reflect strong links in the network, we place a (arbitrary) grid which covers the region of interest. For each grid-edge, we determine a weight by convolving a Gaussian kernel with the original image to capture the edge's capacities to carry certain traffic, the speed of transportation, or combinations thereof (cf. Results).

To facilitate comparative analyses, different seminal properties of the resulting weighted network are computed, including: degree distribution, path lengths, or random-walk-related properties. To evaluate the biological importance of the calculated properties, we developed several null models which randomize the network while preserving the distribution of edge weights. They provide the basis for revealing the principles underlying the network organization [Breuer et al., 2014].

img2net is written in Python and provides a graphical user interface (GUI; Fig. 4.2.1A) for the selection of the image input data and the parameters of the network reconstruction procedure. It operates on folders of .png, .jpg, or .tiff files that display the networked structure of interest, with a gray-scale representation of edge strengths. The directory tree divides the images into different treatments and experiments, e.g.:

root/treatment_003/experiment_005/image_z001_t001.tiff

The usage of rectangular, triangular, or hexagonal grids is supported; grid spacings, periodic boundary conditions, the widths of the convolution kernel and the number of layers for 3D image data can be specified by the user. Different null models are available and the number of null model realizations can be set. Finally, img2net can be run on multiple cores.

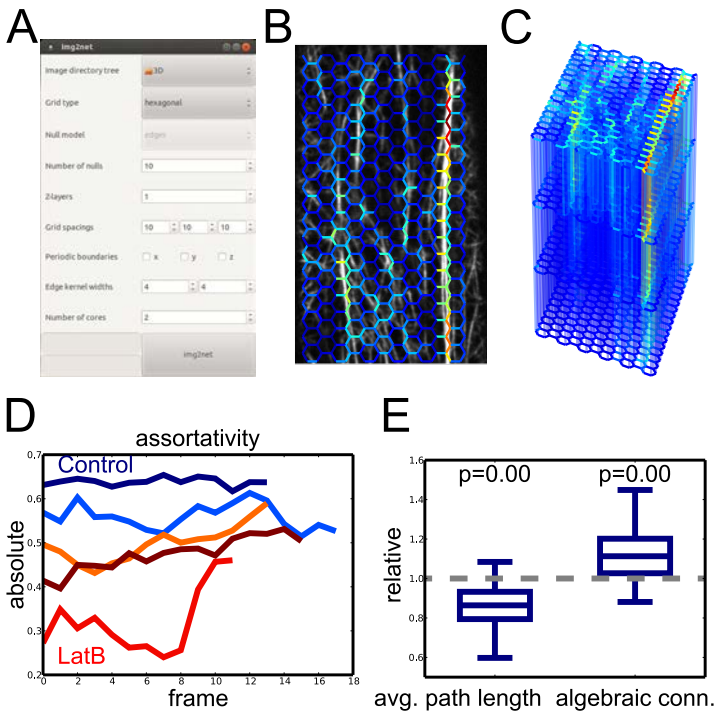


Figure 4.2.1: **Img2net: graphical user interface and output visualizations.** **(A)** GUI to set parameters for the network reconstruction. **(B)** Actin network reconstructed with a hexagonal grid. **(C)** 3D reconstruction for image data with four z-slices. **(D)** Time series of the assortativity for multiple untreated (blue) and treated (red) plants. **(E)** Ratios of average path lengths and algebraic connectivities of observed and null model networks.

An output folder is generated in the root directory of the chosen image directory tree: A human-readable document is created which contains the values of the network properties for the analyzed networks and their null models to allow further analyses and visualizations of the results. img2net generates standard plots as .svg files: For each experiment, a reconstructed network overlaying the original image data is saved. For all analyzed network properties, plots of the resulting time series are generated to facilitate network comparisons.

4.2.4 Results

For testing, img2net was used to compare confocal recordings of the actin cytoskeleton (Fig. 4.2.1B and C) of untreated plant cells and cells treated with an actin disrupting drug [Breuer et al., 2014]. We find that the actin networks of treated cells display a consistently lower assortativity (Fig. 4.2.1D), indicating a drug-related filament fragmentation. Further, for untreated cells, the observed networks display significantly smaller average path lengths ("accessibility") and significantly higher algebraic connectivities ("robustness") than expected by chance (p-values estimated from the null model via one-sample t-tests; Fig. 4.2.1E), suggesting a biological basis for the maintenance of short and robust transportation routes.

We also used img2net to analyze the network structure of the German autobahn which was obtained from

OpenStreetMap as an image by using a gray-scale coding of the speed limits [Breuer et al., 2014]. We find that, similar to the actin cytoskeleton, the autobahn shows significantly smaller path lengths and a higher algebraic connectivity than expected by chance.

4.2.5 Conclusion

Since networks provide an intuitively accessible as well as mathematically sound framework for the representation of complex systems, the reconstruction of spatial networks from image data is useful in biological and technical research. img2net implements an automated method for fast and robust reconstruction of arbitrary 2D and 3D (non-)biological spatial networks.

As our approach relies on fixed grid topologies, img2net offers different types of grids to verify that the findings are grid-independent. For small grid spacings, the grid typically approaches the "true" structure of the underlying network. The main benefit of this approach is its robustness against noise and flaws in the input images. For example, small ruptures in the underlying network do not disrupt the corresponding edges but only weaken them. Hence, our approach does not require sophisticated error-correcting image processing steps. Furthermore, img2net is directly applicable to a wide range of image data from different sources.

The analyses of the reconstructed networks implemented in img2net allow a quantification of structural properties, comparisons of networks, e.g., under different conditions, and an assessment whether or not the network properties reflect underlying organizational principles.

4.2.6 Acknowledgements

Support: We thank Alexander Ivakov and Staffan Persson for help with the biological experiments and discussions on cytoskeletal networks.

Funding: D.B. and Z.N. were supported by the Max Planck Society.

Contributions: D.B. and Z.N. designed the study; D.B. implemented the method; D.B. and Z.N. wrote the manuscript.

4.3 Cell-specific organization and optimality of the plant actin cytoskeleton

Publication: *in preparation*

Authors: David Breuer[1,2,*], Yi Zhang[3], Staffan Persson[4,5,6], Zoran Nikoloski[1]

Affiliations: [1]Systems Biology and Mathematical Modeling, Max Planck Institute of Molecular Plant Physiology, Am Muehlenberg 1, 14476 Potsdam, Germany

[2]Institute of Biochemistry and Biology, University of Potsdam, Karl-Liebknecht-Straße 24-25, 14476 Potsdam, Germany

[3]Central Metabolism, Max Planck Institute of Molecular Plant Physiology, Am Muehlenberg 1, 14476 Potsdam, Germany

[4]Max Planck Institute of Molecular Plant Physiology, Am Muehlenberg 1, 14476 Potsdam, Germany

[5]School of Biosciences, University of Melbourne, Grattan Street, Parkville, Victoria 3010, Australia

[6]ARC Centre of Excellence in Plant Cell Walls, School of Botany, University of Melbourne, Grattan Street, Parkville, Victoria 3010, Australia

Contact: *breuer@mpimp-golm.mpg.de

4.3.1 Abstract

Cytoskeletal networks are essential cellular components that serve a multitude of functions, from shaping cell mechanical properties to enabling cellular transport. From an evolutionary standpoint, the cytoskeleton may be optimized towards its, likely, cell-type specific functions, although quantitative evidence to this end is largely missing. Here, we employ an image-based network representation of the cytoskeleton that provides the basis for systematic, quantitative, and automated investigation of cytoskeletal organization. Our comparative analysis of the actin cytoskeleton from four plant cell types supports the hypothesis of cell-specific actin organization. We demonstrate that elongating hypocotyl and roots cells accommodate dispersed actin cytoskeletal networks that facilitate fast cellular transport, while mature cells from these organs exhibit stronger actin bundling that may restrain further cell elongation. We show that the performance of the cytoskeletal networks follows a trade-off between three network properties that reflect the reachability in the actin networks as well as their robustness against disruptions and cell coverage. Moreover, by employing multi-objective network optimization, we show that the cytoskeleton in elongating cells, in comparison

to randomized networks, operates close to the Pareto frontier which represents the set of optimal compromises between these different network properties. Thus, our findings provide quantitative insights into the multi-faceted design principles of the plant actin cytoskeleton and its near-optimal organization.

Keywords: cytoskeleton, actin, multi-objective optimization, Pareto frontier, networks

4.3.2 Introduction

Most complex technical and biological systems perform different and often opposing tasks that lead to inherent trade-offs. The presence of trade-offs results in deviations from optimality for single task. For instance, transport systems, like road networks [Barthélemy and Flammini, 2008; Barthélemy, 2011; Louf et al., 2013], exhibit trade-offs between construction costs, travel times between different locations, and the robustness of the network against disruptions. Beyond these transport-related compromises, biological transport systems, like leaf venation [Bohn et al., 2002; Dodds et al., 2010; Katifori et al., 2010; Onoda et al., 2011; Sack and Scoffoni, 2013] or cytoskeletal networks [Staiger et al., 2000; Nick, 2007; Fletcher and Mullins, 2010; Breuer et al., 2014, 2015], often entail additional trade-offs with respect to mechanical objectives such as leaf stability or cell mechanics.

For the cytoskeleton, consequently, it has been suggested that cell types with different functional requirements accommodate different cytoskeletal structures. For example, in plants, young elongating hypocotyl cells have been suggested to harbor dynamic and diverse actin cytoskeletons that support efficient transport processes [Staiger et al., 2009; Breuer et al., 2014]. In contrast, it has been suggested that the actin cytoskeleton in mature hypocotyl cells is more static, comprising few thick actin bundles that constrain further cell elongation [Hussey et al., 2006; Sampathkumar et al., 2011]. However, while most of these studies relied on visual inspection, a systematic, quantitative, and fully automated investigation of cell-specific organization of the actin cytoskeletal remains elusive.

Moreover, considering the strong conservation of cytoskeletal proteins and functions throughout evolution [Cooper, 2000; van den Ent et al., 2001; Gunning et al., 2015], it has been suggested that cytoskeletal networks are optimized for their specific tasks [Breuer et al., 2014; Recho et al., 2014; Ando et al., 2015]. However, most studies were only theoretical and did not incorporate real-life cytoskeletal architectures [Recho et al., 2014; Ando et al., 2015]. Another study compared image-based network representations of the plant actin cytoskeleton to randomized null model networks [Breuer et al., 2014], suggesting that different transport-related properties of the actin cytoskeleton do not arise by chance but reflect underlying, yet unknown, organizational principles. However, the challenge remains to identify the key tasks that these cytoskeletal networks aim to optimize, and how close to optimality these networks may operate with respect to fulfilling these tasks.

Here, we study trade-offs in cytoskeletal organization in different cell types and investigate their optimality with respect to the multiple tasks the fulfill. To this end, we employ the image-based network representation of the cytoskeleton developed in [Breuer et al., 2014; Breuer and Nikoloski, 2014] to quantify and analyze cytoskeletal phenotypes in various cell types of *Arabidopsis thaliana* (Arabidopsis) plants. Moreover, by employing multi-objective optimization, we construct networks that optimize different seminal network properties (e.g., average path length or “reachability”, algebraic connectivity or “robustness”, Herfindahl index

of edge weights or "homogeneity" of cell coverage). Our findings demonstrate that the studied cytoskeletal architectures are cell specific. In addition, we show that elongating cells exhibit near-optimality with respect to some, but not all, studied network properties. Our approach is readily applicable to investigate these questions in other types of spatial networks, given that their structure is known or they are accessible to imaging efforts.

4.3.3 Results

Extracted actin cytoskeletal networks display cell-type specific organization

To study actin organization in various plant cell types, we imaged the cortical cytoskeleton of FABD2-GFP-labeled Arabidopsis seedlings using confocal microscopy (Fig. 4.3.1A; cf. Materials and Methods). Next, to robustly quantify the actin organization, we extracted grid-based weighted network representations of the actin cytoskeleton from these image data as established and described in [Breuer et al., 2014; Breuer and Nikoloski, 2014] (Fig. 4.3.1A; cf. Materials and Methods). Such network representations naturally capture the filamentous structure of the cytoskeleton and enable quantitative analyses and linking of different network properties to cellular functionality. We recorded at least six cells of four different cell types, i.e., elongating and mature hypocotyl and root cells (Figs. 4.3.1B and C).

While certain differences in actin organization between the studied cell types seem visually apparent, we computed several seminal network properties to quantify these differences in an automated and unbiased way (Fig. 4.3.1D; cf. Materials and Methods). Namely, we focused on the average path length (measuring the "reachability" of a network), the algebraic connectivity (measuring the "robustness" of a network against disruptions), the Herfindahl index of the edge weights (measuring the "homogeneity" of the actin distribution), the degree assortativity (measuring the "bundling" of filaments in an actin network), and the average size of the connected components after removal of edges with low weights (measuring the "connectedness" of a network). In the following, $v_{i,j,k}$ refers to network property k of the actin network in the jth studied cell of cell type i. The studied network properties allow unequivocal identification of each of the four cell types. For example, computing the algebraic connectivity and the Herfindahl index, we found that actin networks in elongating cells showed larger values than in mature cells. These findings indicate stronger dispersal of the actin cytoskeleton in elongating cells and stronger bundling in mature cells. Alternatively, computing the average size of the connected components after removal of edges with low weights, we found consistently smaller values for this property in actin networks of root cells than in hypocotyl cells. This distinction indicates stronger fragmentation of the actin cytoskeleton in root cells as opposed to hypocotyl cells, which may be explained by the more horizontal alignment of actin filaments and bundles in the former (cf. Figs. 4.3.1A and C).

To study these differences more systematically, and for better comparability, we computed the relative average values $u_{i,k}$ of our five network properties k in the studied cell types i (Fig. 4.3.1E; $u_{i,k} = \left(v_{i,j,k} - \mathrm{E}_{i,j}\left[v_{i,j,k}\right]\right)/\mathrm{SD}_{i,j}\left[v_{i,j,k}\right]$, where $\mathrm{E}_{i,j}$ and $\mathrm{SD}_{i,j}$ are averages and standard deviations across cell types i and cells j, respectively). We found that the average path length, algebraic connectivity, and the Herfindahl index reliably distinguished between different cell maturation states, while the average size of the connected components allowed distinction of hypocotyl and root cells. The assortativity did not allow

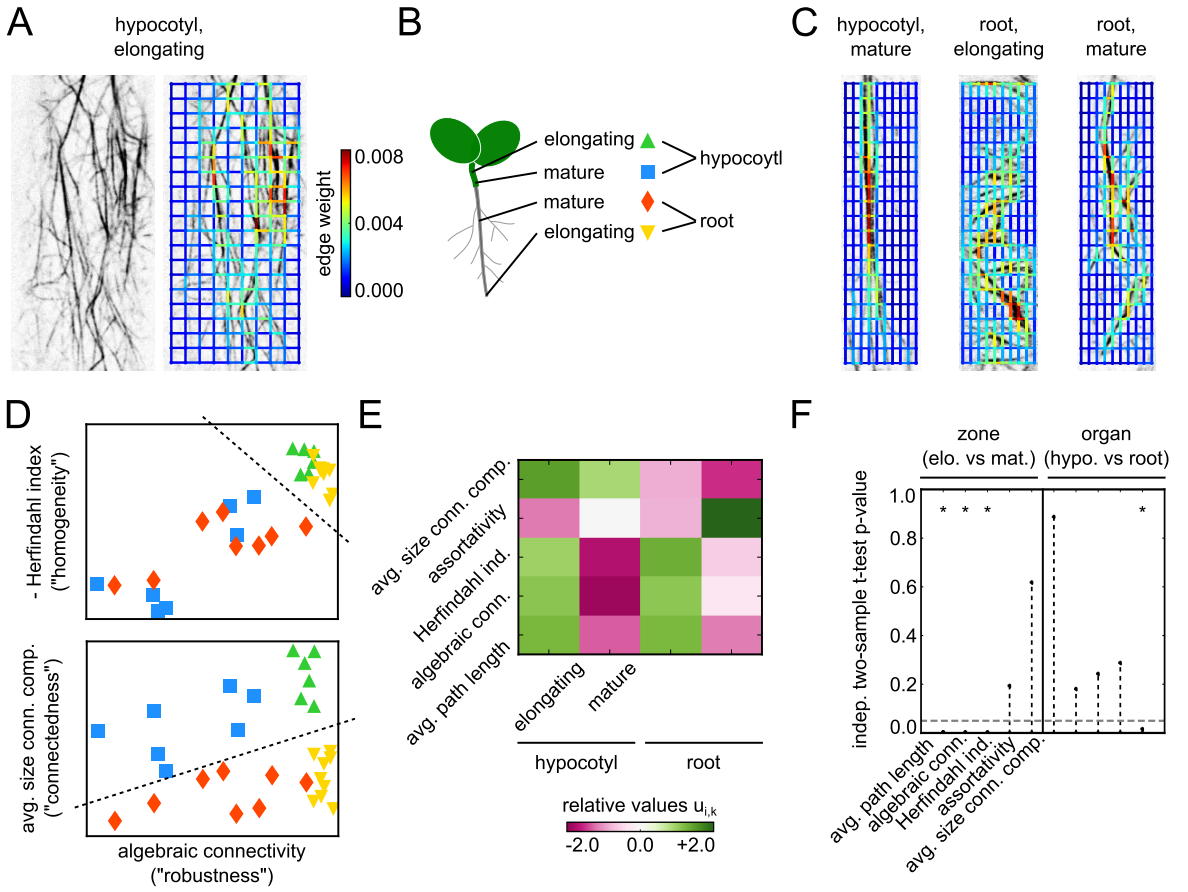


Figure 4.3.1: **Extraction of actin cytoskeletal network from image data and cell-type specific actin organization in elongating and mature plant hypocotyls and roots.** **(A)** Average over 10 slices of confocal z-stack recording of cortical actin cytoskeleton in an elongating Arabidopsis hypocotyl cell (left panel). Overlay of this actin image and extracted network representation with color-coded edge weights reflecting the underlying intensity (right panel). **(B)** Schematic of Arabidopsis seedling and studied cell types, i.e., elongating and mature cells from plant hypocotyl and root. **(C)** Overlay of confocal recordings of actin cytoskeletons and weighted network representation from mature hypocotyl cell (left panel), as well as elongating (middle panel) and mature root cells (right panel). **(D)** Sections through space $v_{i,j,k}$ of network properties k for actin networks j from different cell types i (cf. (B)). Actin networks in elongating cells showed larger algebraic connectivities ("robustness") and Herfindahl indices ("homogeneity") than in mature cells (upper). Actin networks from hypocotyl cells showed larger average sizes of the connected components ("connectedness") than in root cells. **(E)** Relative average values $u_{i,k}$ of different network properties, including assortativity ("bundling") and average path length ("reachability"), across the four studied cell types. Averages and standard deviations of a given network property across cell types were normalized to zero and one for better comparability, respectively. **(F)** Bar plot of independent two-sample t-test p-values between elongating and mature (left) and hypocotyl and root cells (right), respectively, whereby p-values< 0.05 were considered significant (cf. stars). While average path length, algebraic connectivity, and Herfindahl may be used to distinguish elongating from mature cells, the average size of the connected components may be used to distinguish hypocotyl from root cells.

a conclusive distinction of different cell types. These findings were statistically confirmed by independent two-sample t-tests between the different classes of cells (Fig. 4.3.1F). Thus, our findings imply that the studied network properties can be used for discriminant purposes, e.g., classification of different cell types by the organization of their actin cytoskeleton.

Actin cytoskeletal networks are near-optimal in elongating, but not in mature cells

Beyond organizational differences of actin cytoskeletal networks in different plant cell types, we investigated their optimality. Optimality was assessed with respect to the introduced network properties that reflect different potential cytoskeletal functions. We employed multi-objective optimization to construct networks that optimize one or multiple network properties (Figs. 4.3.2A; cf. Materials and Methods). Generally, such multi-objective optimization yields a set of networks that form a hyper-surface in the space of network properties called Pareto frontier (Fig. 4.3.2A, black surface; [Ehrgott, 2006; Handl et al., 2007; Shoval et al., 2012]). For a fixed network property and value, the Pareto frontier is monotonically decreasing in all other properties, which reflects the trade-off between the different properties.

The proximity of the real-life actin networks to the Pareto frontier indicates their optimality with respect to the trade-offs. For example, studying the average path length, Herfindahl index, and algebraic connectivity of actin networks from different cell types, we found that elongating cells were closer to the Pareto frontier than mature cells (Fig. 4.3.2A). Moreover, to assess the distance from the Pareto frontier, we computed the network properties for an ensemble of randomized networks, demonstrating that the actin networks were significantly closer to the Pareto frontier than expected by chance (Fig. 4.3.2A, gray dots; cf. Materials and Methods; [Breuer et al., 2014; Breuer and Nikoloski, 2014]). Taken together, our findings suggest that the actin cytoskeleton in elongating cells is optimized for cellular transport, as captured by the network properties studied here, while mature cells have a reduced need for efficient, uniform cellular transport. Moreover, each of the studied real-life actin networks in elongating cells did not optimize a single network property, but a combination thereof. This trade-off is supported by artificial, optimal networks which optimize only a single network property and which differ visually from real-life actin networks. For instance, the network optimizing algebraic connectivity shows a large region of strong edge weight in the center of the network (Fig. 4.3.2B), the network optimizing average path length shows highway-like filamentous structures in the center (Fig. 4.3.2C), and the network optimizing the Herfindahl index shows a uniform distribution of edges weights (Fig. 4.3.2D). Moreover, all optimal networks exhibit vertical and horizontal symmetry in the distribution of edge weights. The trade-off between various network properties, i.e., average path length, algebraic connectivity, and Herfindahl index, in elongating cells is further supported by their negative correlations (Fig. 4.3.2E, cf. black dashed rectangle; Pearson correlation coefficients $p_{k,k'}$ between network properties k and k'), suggesting that the networks operate close to the Pareto frontier where an increase in one property implies a decrease in the other properties. In mature cells, in contrast, these correlations are positive which is only possible if some of the networks are further away from the Pareto frontier. The strongly negative correlations between the assortativity and the remaining network properties for the mature cells are due to the relatively strong actin bundling (cf. Fig. 4.3.1) which favors high assortativities but decreases the other network properties.

To further quantify the optimality of the actin networks in different plant cell types i, we computed the

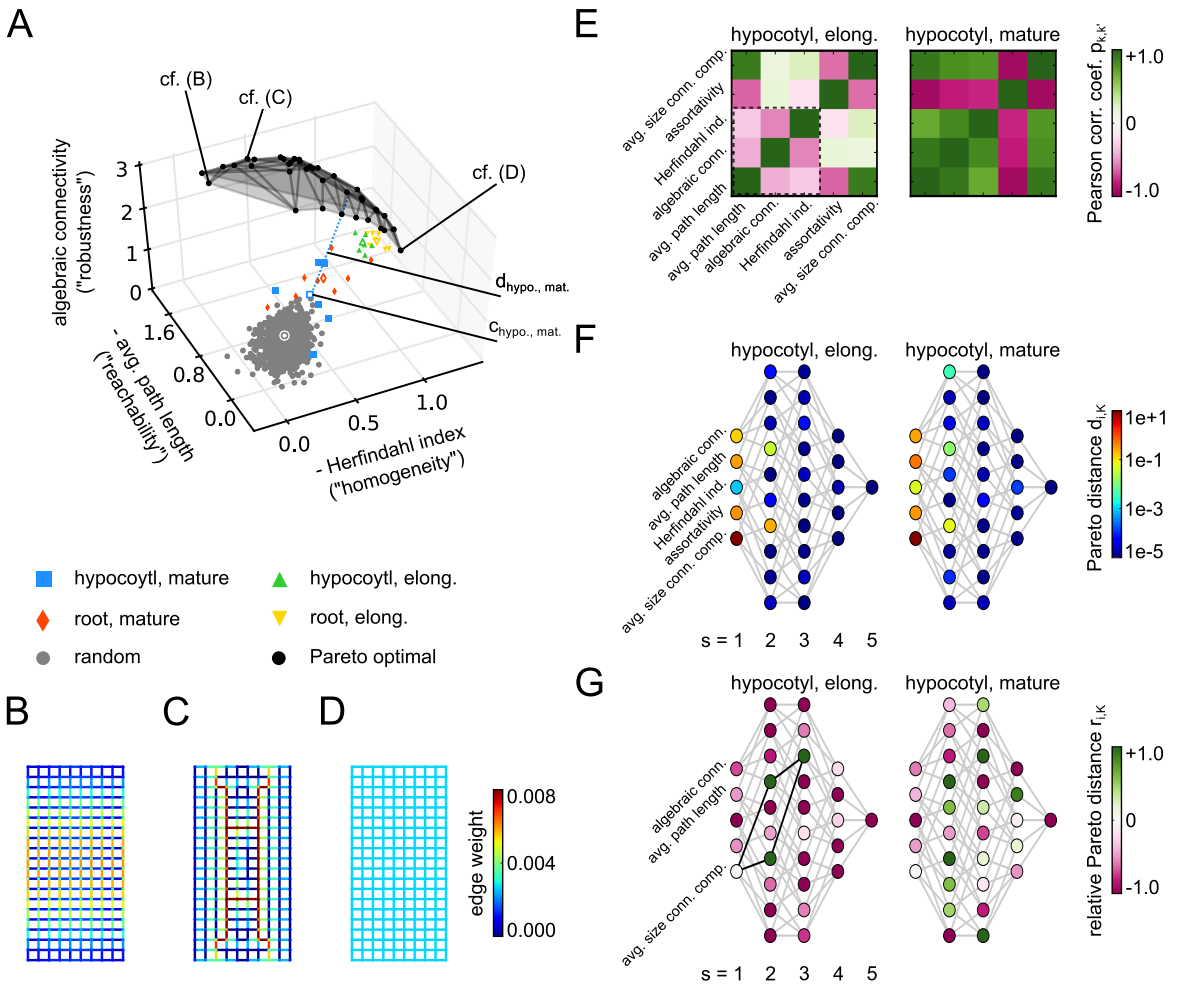


Figure 4.3.2: **Optimality and trade-offs in plant actin cytoskeletal organization.** **(A)** Three-dimensional section through the space of actin network properties and Pareto frontier (black surface) for average path length, Herfindahl index, and algebraic connectivity. Actin network properties from elongating and mature hypocotyl and root cells, respectively, and ensemble of network with edge weights drawn from a random and uniform distribution (gray dots; cf. Materials and Methods). Centroids are given by open symbols and the distance to the Pareto frontier for elongating hypocotyl cells is indicated by the blue dotted line. **(B)** Network with color-coded edge weights and highest algebraic connectivity. **(C)** Network with smallest average path length. **(D)** Network with smallest Herfindahl index. **(E)** Heat maps of Pearson correlation coefficients $p_{k,k'}$ between network properties k and k' for elongating (left panel) and mature hypocotyl cells (right panel). **(F)** For each subset $\mathcal{K}$ of the five studied network properties, we computed the distance $d_{i,\mathcal{K}}$ between the Pareto frontier and the centroid of actin network properties for different cells types i, i.e., elongating (left panel) and mature hypocotyl cells (right panel). The vertical layers of nodes represent the s-element subsets of network properties, node colors reflect the Euclidean distance of the respective centroid to the Pareto frontier, and edges indicate the hierarchical structure of the subsets. **(G)** Same graphs as in panel (E), with node colors representing relative distances to the Pareto frontier $r_{i,\mathcal{K}}$, i.e., the relative distance to the Pareto frontier in comparison to that of the centroid of randomized networks. For actin networks in elongating hypocotyl cells (left panel), only four subsets of network properties, all containing the average size of the connected components, yielded larger distances from the Pareto frontier than expected by chance (cf. green circles and black edges). In contrast, actin networks in mature hypocotyl cells (right panel) commonly showed distances from the Pareto frontier that are comparable to or larger than those of randomized networks.

Euclidean distances $d_{i,\mathcal{K}}$ between the centroid $c_{i,\mathcal{K}}$ of the networks and the Pareto frontier for all possible subsets $\mathcal{K}$ of network properties (Fig. 4.3.2F; centroids are indicated by open symbols and given by $c_{i,\mathcal{K}} = \left(\mathrm{E}_j\left[v_{i,j,k}\right]\right)_{k\in\mathcal{K}}$, where E_j is the average across different cell j of a given cell type i). We found that the studied actin networks were close to the Pareto frontier with respect to the transport-related properties employed. In particular, the networks of elongating hypocotyls were closer to the Pareto frontier than those of mature hypocotyl cells. The same trends like for the elongating and mature hypocotyl cells were observed for the respective root cell types. We note that the Euclidean distance of a centroid to the Pareto frontier with respect to a set of network properties is always as large as the minimum Euclidean distance of the centroid with respect to a subset of the properties due to geometrical considerations (i.e., in the higher dimensional space of the superset centroid, the Pareto frontier is still reachable along the lower dimensional paths of its subset centroids). Therefore, to assess the distances of actin network centroids to the Pareto frontier without this dimensionality bias for a subset of network properties $\mathcal{K}$, we studied the the relative distances $r_{i,\mathcal{K}} = \left(d_{i,\mathcal{K}} - d_{\mathrm{random},\mathcal{K}}\right)/d_{\mathrm{random},\mathcal{K}}$, where $d_{\mathrm{random},\mathcal{K}}$ is the distance of the centroid of randomized networks to the Pareto frontier (Fig. 4.3.2G). We found that for the elongating hypocotyl cells, the centroid of the actin networks was indeed closer to the Pareto frontier than expected by chance for all subsets of network properties, except for four subsets of properties that contained the average size of the connected components. This indicates that elongating hypocotyl cells to not optimize the average size of the connected components, which is in agreement with its positive correlations with most of the other network properties (cf. Fig. 4.3.2E). In contrast, for many subsets of network properties, the centroid for actin networks from mature hypocotyl cells was farther from the Pareto frontier than expected by chance. This is also compatible with the positive correlations between the average size of the connected components and other network properties (cf. Fig. 4.3.2E) that suggest the absence of Pareto-optimality induced trade-offs between these properties. Again, similar trends were observed for root cells. Overall, mature cells do not seem to optimize multiple network properties, or, at least, not those associated to transport. However, elongating cells in both hypocotyls and roots accommodate actin cytoskeletal networks that are multi-functional and near-optimal with respect to various transport-related properties.

4.3.4 Discussion

The work presented here provides quantitative evidence for cell-type specific organization and optimality of cytoskeletal networks, which have been previously hypothesized [Hussey et al., 2006; Staiger et al., 2009; Breuer et al., 2014] or demonstrated for specific functions using only theoretical models [Recho et al., 2014; Ando et al., 2015].

By employing an established and robust network representation of the cortical actin cytoskeleton in living plant cells, we showed that elongating and mature cells as well as hypocotyl and root cells may be identified by their specific actin structures that differ in particular in their facilitation of transport processes and their degree of fragmentation. Although the employed grid-based approach ignores details of actin filament distribution, it captures relevant features of actin organization and may be readily applied to two- and three-dimensional image data [Breuer et al., 2014; Breuer and Nikoloski, 2014]. Another advantage of the grid-based approach, as compared to segmentation-based methods [Baumgarten and Hauser, 2012; Obara et al., 2012b; Breuer et al., 2015], is its full coverage of the cellular region of interest; this coverage yields net-

works that are comparable across cells with different actin distributions, e.g., irrespective of regions without actin. Moreover, it has been shown that the behavior of the studied network properties persists with different grid types [Breuer et al., 2014; Breuer and Nikoloski, 2014]; hence, we expect that our current findings are likewise independent of the grid type, although the validity of this claim remains to be shown in additional analyses.

Moreover, we demonstrated that the actin network, at least in elongating cells, is near-optimal with respect to various transport-related properties. An exception was the average size of the connected components, which reflects the connectedness of the network and whose distance from the Pareto frontier was comparable to that of randomized networks, suggesting that the actin cytoskeleton *in vivo* is readily fragmented into several components. The deviations from optimality of the actin cytoskeleton in mature cells with respect to the studied network properties may be due to three reasons: First, our studied network properties are suitable to capture the transport functionality of the cytoskeleton, but they do not integrate mechanical aspects of cytoskeletal organization which may constitute major objectives in mature cells [Hussey et al., 2006; Sampathkumar et al., 2011]. Second, the present image data focused on a part of the cortical cytoskeleton at the front side of the cell. However, complete understanding of cytoskeletal functionality and optimality may require incorporation of the cylindrical geometry of the cytoskeleton [Breuer et al., 2015]. Third, even the strongly bundled actin cytoskeleton in mature plant cells is dynamic and rearranges on a minute time scale. Such dynamic rearrangements may affect both the effective velocity of cargo transport as well as the cell coverage. Therefore, a full understanding of cellular transport optimality needs to consider the dynamics of cytoskeletal networks. However, to our knowledge, no such framework is available yet.

Finally, we employed sophisticated but heuristic algorithms for high-dimensional, non-linear, and multi-objective optimization. Consequently, the obtained Pareto frontier may (partially) represent local instead of global maxima. Yet, optimization of some of the studied network properties may be solved, or approximated, via semidefinite programming [Jensen and Barnes, 1980; Schrijver, 1998; Sun et al., 2006; Boyd, 2006]. In a continuation of this work, we will therefore solve these semidefinite programs that are convex and hence guaranteed to yield global optima for the network properties or, at least, their approximations (cf. Appendix 6.2.1 for the semidefinite programs).

In conclusion, our work provides quantitative evidence for the common hypotheses of cell-type specific and (near-)optimal organization of the actin cytoskeleton. Additionally, by complementing our combination of experimental microscopy data and theoretical network models by mechanical measurements, we are confident that the established approach may yield comprehensive insights into cytoskeletal organization organization.

4.3.5 Materials and Methods

Recordings of plant actin cytoskeletons from different cell types

We used Arabidopsis Columbia-0 35S:FABD2-GFP-labeled seedlings to study the actin cytoskeleton [Sheahan et al., 2004; Sampathkumar et al., 2013]. To capture fine actin filaments and bundles and their dynamics, as well as to minimize photo bleaching, we employed a spinning-disk confocal microscope with a spatial

resolution of $0.133\,\text{pixel}^{-1}\,\mu\text{m}$ [Sampathkumar et al., 2011]. Details of sample preparation and experimental setup are given in [Sampathkumar et al., 2011; Breuer et al., 2015]. Images were taken as z-stacks with 100 nm steps over 10 slices that were averaged to capture the cortical actin cytoskeleton. Four types of cells were investigated, and for each type we imaged at least 6 cells from 6 different plants: In 4-day-old hypocotyls, we imaged young, elongating cells and mature, non-elongating cells at the top (apical) and the bottom (basical) end of the hypocotyl, respectively. Similarly, in 5-day-old seedlings grown under 16 h light / 8 h dark conditions, we imaged cells in the elongating and the mature region of the roots, respectively.

Extraction of weighted networks from cytoskeletal image data

To investigate the network structure of the plant actin cytoskeleton in different cell types, we employed a robust method for the extraction of weighted networks from gray-scale image data established in [Breuer et al., 2014; Breuer and Nikoloski, 2014]. First, for a given recording, we employed the Fiji-StackReg stack registration algorithm to correct the potential drift of the seedlings under the microscope [Thévenaz et al., 1998]. We normalized the mean intensity of each frame to compensate photobleaching. We applied the Fiji-BackgroundSubtraction rolling ball filter with radius of 50 pixels to improve the signal-to-noise ratio [Sternberg, 1983]. We cropped a rectangular cellular region of interest. We then placed a rectangular grid of 30×15 nodes over the region of interest and, for each grid-edge, determined the edge weight by convolving a two-dimensional Gaussian kernel with a standard deviation of 2 pixels with the original image to capture the underlying intensity [Breuer et al., 2014; Breuer and Nikoloski, 2014]. Thus, the actin cytoskeleton was represented by a weighted, undirected network $G = (\mathcal{N}, \mathcal{E})$ of $N = |\mathcal{N}|$ nodes with positions $r_n \in \mathbb{R}^2$, $n \in \mathcal{N}$, and $E = |\mathcal{E}|$ edges with edge weights w_e, $e = (n,m) \equiv (m,n) \in \mathcal{E}$. The sum of edge weights was normalized to one $\sum_{e\in\mathcal{E}} w_e = 1$.

Studied network properties

We studied various seminal network properties to quantify actin cytoskeletal organization in different cell types [West, 2001; Bornholdt and Schuster, 2002; Newman, 2009]. To capture the heterogeneity of the actin filament distribution across the cell, we computed the Herfindahl index of the edge weights,

$$v_H = \sum_{e\in\mathcal{E}} w_e^2, \tag{4.3.1}$$

whereby values close to zero and one indicate homogeneous, uniform and heterogeneous, strongly bundled distributions of actin filaments across the cell, respectively. As a related measure of actin filament clustering, we computed the degree assortativity

$$v_A = \frac{1}{2E} \sum_{n\in\mathcal{N}} \sum_{m\in\mathcal{N}} \left(w_{e=(n,m)} - \frac{d_n d_m}{2E} \right) d_n d_m, \tag{4.3.2}$$

where $d_n := \sum_{m\in\mathcal{N}} w_{e=(n,m)}$ is the node degree and positive values indicate preferential connection of nodes of similar degrees. Next, as a measure for the reachability of nodes in a given network, we computed the

average minimum path length,

$$v_L = \frac{1}{2N(N-1)} \sum_{n \in \mathcal{N}} \sum_{\substack{m \in \mathcal{N} \\ m > n}} L_{n,m}, \tag{4.3.3}$$

where $L_{n,m}$ is the shortest path length between nodes n and m, the length of an edge e is given by its inverse weight, $l_e := w_e^{-1}$, and small average path lengths suggest efficient connections between all nodes. Similarly, to quantify the connectedness of a given network, we computed the average maximum flow,

$$v_F = \frac{1}{2N(N-1)} \sum_{n \in \mathcal{N}} \sum_{\substack{m \in \mathcal{N} \\ m > n}} F_{n,m}, \tag{4.3.4}$$

where $F_{n,m}$ is the maximum flow between nodes n and m and large values indicate strong connectedness of the network. Moreover, we computed a related measure for the robustness of a network against disruptions, given by the algebraic connectivity,

$$v_C = \lambda_2, \tag{4.3.5}$$

which is the second smallest eigenvalue of the graph Laplacian $\mathscr{L}$ and reflects how well-knit the network is,

$$\mathscr{L}_{n,m} = \begin{cases} d_n & \text{, if } n = m \\ -w_{e(=n,m)} & \text{, if } e \in \mathscr{E} \\ 0 & \text{, otherwise.} \end{cases} \tag{4.3.6}$$

Finally, as a measure for the connectedness of a network, we computed the average size of the connected components after removal of edges with weights below the 50th percentile,

$$v_S = \frac{1}{F} \sum_{f \in \mathscr{F}} N_f, \tag{4.3.7}$$

where N_f is the number of nodes in component $f \in \mathscr{F}$ and $F = |\mathscr{F}|$ is the number of connected components.

Multi-objective optimization of weighted networks

Since the optimization of several of the studied network properties cannot be formulated, or conveniently approximated, as a convex problem, we applied a heuristic approach for solving high-dimensional and non-linear optimization problems. We used the well-established Broyden-Fletcher-Goldfarb-Shanno algorithm with limited memory requirements and bound constraints on the variables (L-BFGS-B; [Byrd et al., 1995; Zhu et al., 1997]) which is a modified Newton method that estimates the local inverse Hessian matrix to find

local maxima, i.e.,

$$\max \quad c_H \cdot \nu_H^{'} + c_L \cdot \nu_L^{'} + c_F \cdot \nu_F^{'} + c_C \cdot \nu_C^{'}, \tag{4.3.8}$$

$$\text{subject to} \quad \sum_{e \in \mathscr{E}} w_e = 1 \tag{4.3.9}$$

where we realized different objective functions by varying the coefficients c_i, $i \in \mathscr{I} = \{H, A, L, F, C, S\}$, in 11 equidistant steps in the unit interval, such that $\sum_{i \in \mathscr{I}} c_i = 1$. Moreover, the second line normalizes the sum of edges weights to one. Since the values of the different network properties may differ by several orders of magnitude, we rescaled them in the objective function via

$$\nu_i^{'} = \frac{\nu_i^* - \nu_{i,5}}{\nu_{i,95} - \nu_{i,5}}, \tag{4.3.10}$$

where ν_i^* is the original value of the network property. Moreover, $\nu_{i,5}$ and $\nu_{i,95}$ were given the 5th and 95th percentile of network property i, respectively, when computed for an ensemble of 10^4 networks with edge weights w_e drawn from a uniform distribution over the open standard $(E-1)$-simplex, guaranteeing that $\sum_{e \in \mathscr{E}} w_e = 1$. As initial edge weights, we drew w_e from the same uniform distribution, and repeated the procedure 10 times for each objective. We manually selected a step size of 10^{-8} for the approximation of the Hessian matrix and, for each multi-objective optimization problem, we ran the algorithm for up to 10^6 function evaluations or until convergence was achieved, i.e., the value of the objective function changed by less than 10^{-6} percent. Results of the optimization procedure are shown in Fig. 4.3.2.

4.3.6 Acknowledgements

Support: We thank Anna Stief for valuable comments on the manuscript.

Funding: D.B. acknowledges funding from an International Max Planck Research School scholarship. Y.Z., S.P. and Z.N. were funded by the Max Planck Society. S.P. was funded by a R@MAP Professorship at University of Melbourne and is grateful for a Dyason travel grant. S.P. and Z.N. acknowledge an IRRTF (RNC) grant that aided in completing this work.

Contributions: D.B. and Z.N. designed the study; Y.Z. recorded the data; D.B. implemented the method and analyzed the data; D.B., Y.Z., S.P. and Z.N. interpreted the results and wrote the manuscript.

4.4 System-wide organization of the actin cytoskeleton drives organelle transport in plant cells

Publication: *submitted (Sci. Adv., 16.12.2015)*

Authors: David Breuer[1,2,*], Alexander Ivakov[3,4], Jacqueline Nowak[1,2,4], Staffan Persson[3,4,5], Zoran Nikoloski[1]

Affiliations: [1]Systems Biology and Mathematical Modeling, Max Planck Institute of Molecular Plant Physiology, Am Muehlenberg 1, 14476 Potsdam, Germany

[2]Institute of Biochemistry and Biology, University of Potsdam, Karl-Liebknecht-Straße 24-25, 14476 Potsdam, Germany

[3]Max Planck Institute of Molecular Plant Physiology, Am Muehlenberg 1, 14476 Potsdam, Germany

[4]School of Biosciences, University of Melbourne, Grattan Street, Parkville, Victoria 3010, Australia

[5]ARC Centre of Excellence in Plant Cell Walls, School of Botany, University of Melbourne, Grattan Street, Parkville, Victoria 3010, Australia

Contact: *breuer@mpimp-golm.mpg.de

4.4.1 Abstract

The actin cytoskeleton provides an essential intracellular filamentous structure in organisms from all kingdoms of life. In plant cells, the actin cytoskeleton constitutes a dynamic network that underpins myosin-based vesicle trafficking and cytoplasmic streaming. Yet, the design principles and system-level properties of actin-based cellular trafficking remain tenuous, largely due to the inability to quantify key features of the actin cytoskeleton. Here, we developed an automated image-based and network-driven framework to accurately segment and quantify the actin cytoskeleton structure and Golgi transport flow. We show that the actin cytoskeleton is optimized for efficient transport. Furthermore, we demonstrate that Golgi transport may be accurately predicted from actin network properties alone. In particular, we show that Golgi transport is determined not only by local distribution of actin filaments, but also by the global cellular architecture of the actin cytoskeleton. Our data also suggest that the previously observed erratic movement of Golgi is a stable cellular phenomenon that may optimize distribution efficiency of cell material, analogous to movement patterns of foraging animals. Our findings thus provide quantitative evidence for whole-cell coordination of cellular transport in plant cells. Our framework can be readily applied to investigate cytoskeletal transport in other scenarios, and paves the way for quantitative network-driven understanding of cellular transport.

Keywords: cellular transport, actin, cytoskeleton, Golgi, systems biology, image processing, networks

4.4.2 Introduction

The cell interior is a heterogeneous and crowded space comprising a large range of molecules and organelles [Luby-Phelps, 2000; Ellis, 2001]. Since diffusion through this complex environment is not sufficient to match varying demands for cell maintenance and growth, intricate cellular transport schemes have evolved [Geisler et al., 2008; Goldstein et al., 2008; Wightman and Turner, 2010; Goldstein and van de Meent, 2015; Wang and Hussey, 2015]. Transport of cellular components across large distances relies substantially on the cytoskeleton [Shafrir et al., 2000; Kim et al., 2005; Cheung and de Vries, 2008; Bálint et al., 2013; Wang and Hussey, 2015]. Moreover, in plant cells, many organelles move rapidly due to actomyosin-based cytoplasmic streaming [Volkmann and Baluska, 1999; Akkerman et al., 2011; Peremyslov et al., 2013]. For instance, Golgi transport relies on the acto-myosin system, and an impaired actin cytoskeleton leads to Golgi aggregation and reduced secretion and endocytosis [Nebenführ et al., 1999; Avisar et al., 2008a; Akkerman et al., 2011; Sampathkumar et al., 2013]. While many molecular features of actin-based transport in plant cells have been elucidated (see [Staiger et al., 2000; Shimmen and Yokota, 2004] for reviews), quantitative measures of the structure of the actin cytoskeleton, and how this structure relates to organelle transport, remain elusive. This is largely due to the difficulties in accurately segmenting the actin cytoskeleton and organelle movement, in particular in growing plant cells.

Theoretical models have been used to analyze the interplay between cytoplasmic streaming and actin organization, demonstrating the emergence of self-organized, rotational streaming patterns [Woodhouse and Goldstein, 2013; Goldstein and van de Meent, 2015]. However, these studies neglected the discrete, filamentous structure of the cytoskeleton. Investigations that have considered discrete cytoskeletal structures revealed different regimes of transport depending on the contribution from diffusion or motor-protein driven transport along random networks of segments [Neri et al., 2013], the impact of motor protein movements on cytoplasm in lattice networks [Houtman et al., 2007], and the effect of length, orientation and polarity of random filament segments on average transport rates [Ando et al., 2015]. Nevertheless, these studies did not incorporate biological data of the cytoskeletal structures.

The three-dimensional structures and rapid dynamics of the actin cytoskeleton coupled with imaging limitations in living plant cells have impeded accurate quantitative measures, which constitute the basis for understanding cytoskeletal transport. Recent reports highlight that plant cytoskeletal networks may support efficient transport processes [Breuer et al., 2014]. However, this finding is based on a lattice approximation of the actin structure that disregards details of actin filament (AF) distribution, and on the assumption that bundled actin supports high transport rates. Moreover, it has been shown experimentally that organelle movement depends on the local actin structures [Akkerman et al., 2011]. Nevertheless, these findings distinguish between only two regions, i.e. fine AFs and bundled actin, thus neglecting the discrete filamentous nature and varying alignment of the local network of AFs. Furthermore, by limiting analyses to local AF organization, the system-wide view of organelle transport remains elusive.

Here, we developed a network-based framework that accurately segments the actin cytoskeleton in growing plant cells and combined it with an automated tracking of Golgi transport. This approach allows us to analyze the four aspects of the actin cytoskeletal transport system, including its structure, design principles, dynamics and control [Kitano, 2002]. By relating these aspects to cytoskeletal transport, we find that the actin cytoskeleton maintains properties supporting efficient transport over time, despite rapid reorganization.

We also show that Golgi wiggling behavior is reminiscent of optimized search strategies that may guarantee efficient uptake and deposition of Golgi-related cell material. In addition, we demonstrate that features of Golgi transport can be predicted by properties of the system-wide organization of the actin cytoskeleton. Altogether, our framework opens up a systems perspective to dissect and understand the transport functionality of the actin cytoskeleton.

4.4.3 Results

A pipeline to extract and represent the actin cytoskeleton as a network

Since the actin cytoskeleton is composed of discrete and interconnected filaments, it can be represented in a network-based framework [Gardel et al., 2004; Breuer et al., 2014; Banerjee and Park, 2015]. In such a framework, the nodes capture crossings or end points of AFs and network edges capture AF segments. The AFs may align and form bundles of varying thickness [Bartles, 2000; Thomas et al., 2009], and the edges can analogously be assigned weights which capture properties of the AF segments (e.g. length or average thickness).

To investigate the organization of the plant actin cytoskeleton, we extracted weighted network representations from three-day-old *Arabidopsis thaliana* (Arabidopsis) hypocotyl cells that expressed FABD-GFP using spinning-disc confocal microscopy data (cf. Fig. 4.4.1F for an overview of the pipeline; cf. Appendix 6.3.1 for video of actin cytoskeleton). The images were corrected for drift and bleaching (Fig. 4.4.1A), cropped to the cellular region of interest and filtered to enhance tube-like structures of the cytoskeleton with a parameter ν_{width} (Fig. 4.4.1B; [Frangi et al., 1998]). AFs were segmented by applying an adaptive median threshold of blocksize ν_{thres} (Fig. 4.4.1C). The binary images were skeletonized to obtain AF center lines and spurious fragments below a pixelsize of ν_{size} or below ν_{int} of the average fragment intensity in the original image were removed (Fig. 4.4.1D). From the resultant images we extracted weighted networks (Fig. 4.4.1E) by identifying the nodes, adding edges between pairs of nodes that were directly connected via the skeleton and assigning different edge weights that reflect the average thickness of an AF segment.

To test if our network-based framework captured relevant biological features of the actin cytoskeleton, we first compared our automated segmentation against contrived images of artificial cytoskeletons of known structure (Fig. 4.4.1G), as well as manually segmented biological cytoskeleton images as a gold standard (Fig. 4.4.1H). Since the accuracy of the network representation relies on four parameters (i.e. ν_{width}, ν_{thres}, ν_{size} and ν_{int}) we performed extensive gauging (cf. [Obara et al., 2012b; Xu et al., 2015]) by varying these parameters in a wide range of values (Figs. 4.4.1I and J) and identified those ensuring best agreement between the center lines of manual and automated segmentations. The quality of the automated segmentation was measured by the Haussdorf distance, i.e. the average minimum distance between pixels of automated and manual segmentations, and parameter gauging yielded an optimal average of $\mathrm{E}[d_{\text{HD}}] = 2.4 \pm 2.1$ pixels (Figs. 4.4.1J and K; [Mayerich et al., 2012; Xu et al., 2015]), comparable to contending approaches (Appendix 6.3.2).

In comparison to similar methods for the extraction of networks from different image sources and systems [Dhondt et al., 2012; Baumgarten and Hauser, 2012; Obara et al., 2012b; Xu et al., 2015] (Appendix 6.3.2),

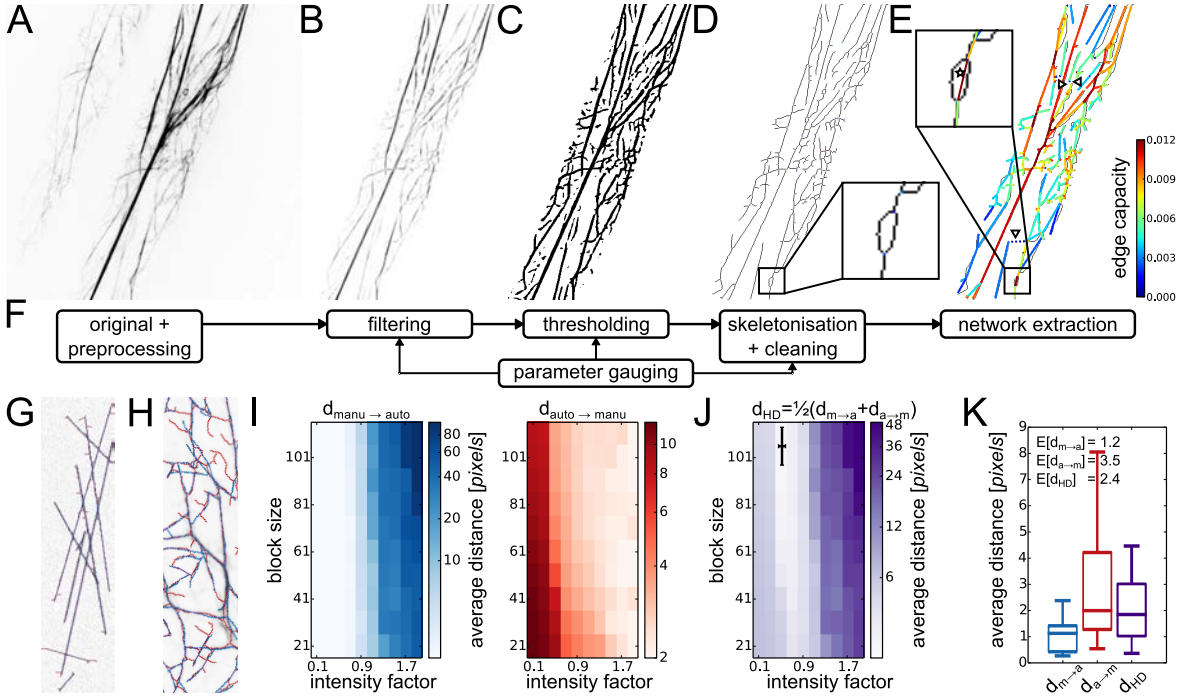


Figure 4.4.1: **Extraction and validation of networks from actin cytoskeletal image data.** **(A)** Grey-scale confocal image of two Arabidopsis hypocotyl cells after registration and background subtraction ("original + preprocessing"). **(B)** Cytoskeleton image with improved signal-to-noise ratio after application of tubeness filter of width $\upsilon_{\text{width}} = 1.8$ ("filtering") and cropping of the largest cell . **(C)** Binary cytoskeleton image after application of adaptive median threshold of blocksize $\upsilon_{\text{thres}} = 101$ ("thresholding"). **(D)** Skeletonized cytoskeletal structures after removal of spurious fragments of small size (i.e. below $\upsilon_{\text{size}} = 27$ pixels) or low intensity (i.e. below $\upsilon_{\text{int}} = 0.5$ of the average fragment intensity in the original image) ("skeletonization + cleaning"). Positions of network nodes are marked by colored pixels (cf. enlarged section). **(E)** Overlay of skeleton image and extracted cytoskeletal network with edges color-coded by their capacity, i.e. the average intensity of the corresponding actin filament per unit length, reflecting the average filament thickness ("network extraction"). Enlarged section shows an edge (cf. star) with high capacity resulting from two contributing filaments. Edges were added to connect the network (cf. dotted lines and triangles). **(F)** Overview of automated pipeline (cf. Materials and Methods) for network-based representation of the actin cytoskeleton incorporating image processing (cf. (A) to (D)), network extraction (cf. (E)) and parameter gauging (cf. (G) to (K)). **(G)** Overlay of contrived filaments (blue) and automated segmentation (red) of an artificial cytoskeleton image (gray). **(H)** Overlay of manual (blue) and automated segmentation (red) of a biological cytoskeleton image (gray). **(I)** Four image processing parameters were varied to determine their optimal values for 20 and 20 images of artificial and biological cytoskeletons, respectively, which were segmented manually for comparison ("parameter gauging"). The quality of segmentation was measured by the average of the smallest distance, $d_{\text{manu}\to\text{auto}}$, from the pixels of the manual segmentation to those of the automated segmentation (blue) and vice versa, $d_{\text{auto}\to\text{manu}}$ (red). Sections of the parameter space, averaged over all 40 studied images for fixed $(\upsilon_{\text{width}}, \upsilon_{\text{int}}) = (1.8, 101)$ and varying blocksize υ_{thres} and size threshold υ_{size}. **(J)** Minimization of the Haussdorf distance $d_{\text{HD}} = \frac{1}{2}\left(d_{\text{manu}\to\text{auto}} + d_{\text{auto}\to\text{manu}}\right)$ (purple) to avoid both over- and undersegmentation, averaged over all 40 studied images, yielded $\left(\upsilon^*_{\text{width}}, \upsilon^*_{\text{thres}}, \upsilon^*_{\text{size}}, \upsilon^*_{\text{int}}\right) = (1.8, 101, 27, 0.50) \pm (0.2, 8.0, 8.9, 0.06)$ (cf. errorbars). **(K)** Distribution of average distances between manual and automated segmentations for the optimal parameters with $\mathrm{E}\left[d_{\text{manu}\to\text{auto}}\right] = 1.2 \pm 0.9$ pixels, $\mathrm{E}\left[d_{\text{auto}\to\text{manu}}\right] = 3.5 \pm 3.2$ pixels and $\mathrm{E}\left[d_{\text{HD}}\right] = 2.4 \pm 2.1$ pixels , respectively.

our approach is fully automated and copes with low signal-to-noise ratios while providing an intensity-based weighted network representation necessary for studying cytoskeletal transport. We note that our current extraction procedure captures only two-dimensional networks, despite the three-dimensional structure of the cytoskeleton [Racine et al., 2007; Breuer et al., 2014; Xu et al., 2015]. However, our focus on two-dimensional networks is justified by the cylindrical shell geometry of the cortical cytoskeleton [Clark et al., 2013; Breuer et al., 2014] as well as the size of the transported Golgi, which may bridge gaps between cortical AFs that are not resolved in two-dimensional confocal images [Boevink et al., 1998; Dupree and Sherrier, 1998; Nebenführ and Staehelin, 2001]. Nevertheless, our network framework can be directly extended to three-dimensional image data, at the cost of more advanced techniques of image processing, lower temporal resolution and/or lower signal-to-noise ratios towards the distant side of the cell. Thus, our approach yields an accurate and mathematically powerful network representation of the cytoskeleton in growing plant cells from image data.

The network representations capture biologically relevant features of the actin cytoskeleton

To assure that our framework captures known changes in the actin cytoskeleton, we determined differences in cytoskeletal organization between cells treated with Latrunculin B (LatB; Fig. 4.4.2A), a drug that inhibits actin polymerization [Yarmola et al., 2000; Wakatsuki et al., 2001], and control cells (Fig. 4.4.2B; seven cells from seven different seedlings per treatment).

To quantify actin network phenotypes, we computed the average fraction of nodes per connected component after removal of edges with capacities below the 50th percentile as a measure of connectedness (Fig. 4.4.2D). The connectedness was higher in networks of control than of LatB-treated cells (Fig. 4.4.2E; independent two-sample t-test p-value, denoted by p_t, $p_t < 10^{-50}$), indicating that large connected patches of AFs were absent in LatB-treated cells, consistent with visual inspection. Similarly, the average edge capacity was higher in control than in LatB-treated cells ($p_t < 10^{-38}$), reflecting a reduction in actin bundling in the LatB-treated cells. Finally, these findings were corroborated by the assortativity, which quantifies whether two adjacent nodes are of similar degree and reflects the network heterogeneity. We again found stronger heterogeneity for control than for LatB-treated cells ($p_t < 10^{-50}$), suggesting regions of bundled actin that are surrounded by AFs in the control cells.

To further assess the filamentous structure of the actin cytoskeleton, we compared the arc length of filament segments to their Euclidean length and found a strong correlation (Fig. 4.4.2F; Pearson correlation coefficient $c_P = 0.998$ and p-value, denoted by p_P, $p_P < 10^{-50}$). Consequently, filament bending, i.e. the ratio of the two lengths [Staiger et al., 2009], was small, $\mathrm{E}[B] = 1.2 \pm 0.2$, in particular for long filament segments (Fig. 4.4.2F, inset). This limited bending of longer filament segments is plausible since actin bundles, typically resulting in longer filament segments, exhibit higher stiffness as compared to AFs [Claessens et al., 2006]. Furthermore, we found that filament segments were preferentially oriented in parallel to the cell axis in control cells, while such preferred orientation was absent in LatB-treated cells (Fig. 4.4.2G).

The changes in actin organization after treatment with the actin-disrupting drug LatB have been previously only described qualitatively [Wakatsuki et al., 2001], or in an approximative grid-based representation of the cytoskeleton [Breuer et al., 2014]. Here, we confirmed these findings and gained additional insights into key aspects of cytoskeletal organization, i.e. filament bending and orientation. Therefore, our results

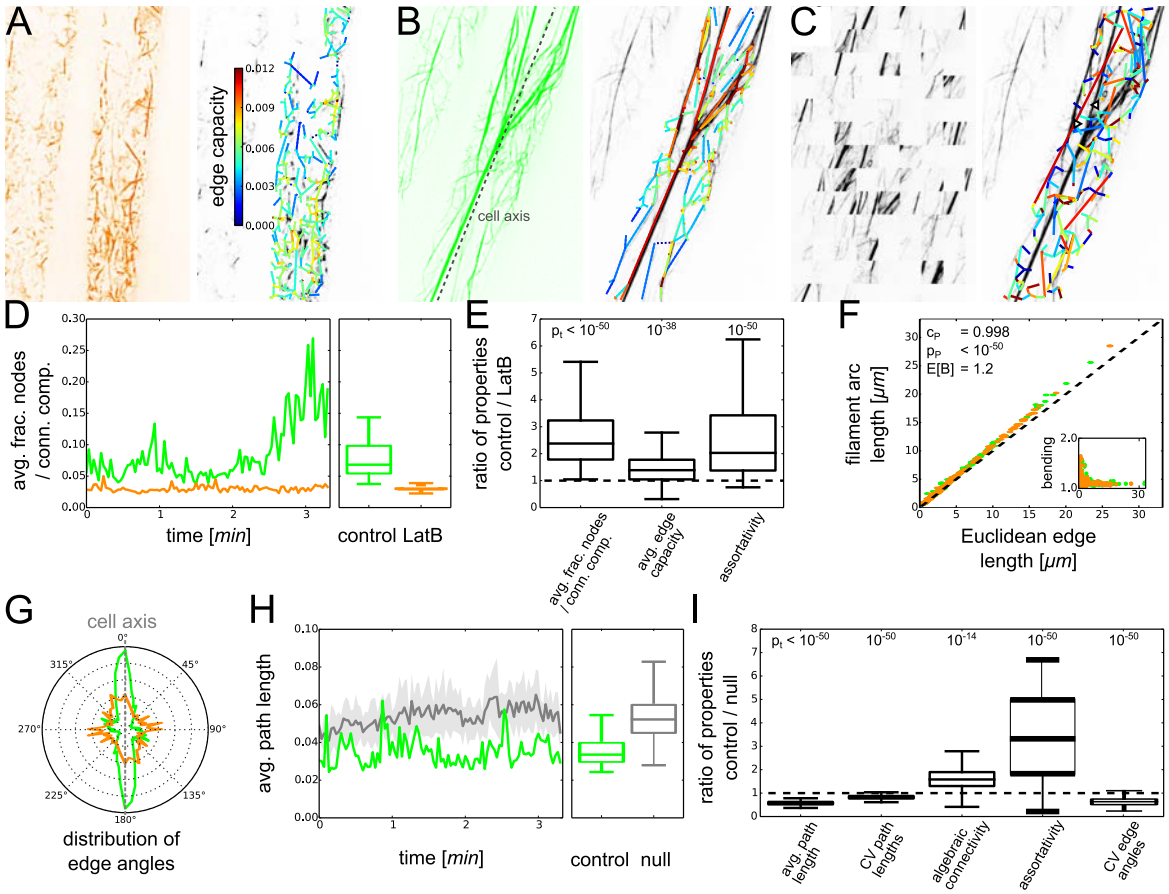


Figure 4.4.2: **Phenotyping of cytoskeletal structures using the extracted networks captures biological signals and reveals transport efficiency.** Results for cells of plants treated with the actin-disrupting drug LatB (orange), untreated control cells (green) and ensembles of 20 randomized networks (gray) employed as a null model to assess the biological relevance of different network properties. **(A)** Cellular recording (left panel; orange) and extracted actin network of a LatB-treated cell (right panel) with edge colors representing edge capacities (cf. Fig. 4.4.1E). **(B)** Cellular recording (left panel; green) and extracted actin network of an untreated control cell (right panel). **(C)** Artistic interpretation of the randomization procedure (left panel; gray) and a randomized network (right panel) of the control cell shown in (B) with occasional edge crossings (cf. e.g. triangles). **(D)** Caption continued on next page.

(D) Time series and boxplots of the average fraction of nodes per connected component after removal of edges with capacities below the 50th percentile ("connectedness") for a control and a LatB-treated cell. **(E)** Ratios of different properties of networks extracted from seven control and seven LatB-treated cells. The average fraction of nodes per connected components ("connectedness"), the average edge capacity ("bundling") and the degree assortativity ("heterogeneity") were significantly higher for control than for LatB-treated cells (independent two-sample t-test p-values $p_t < 0.05$ were considered significant). **(F)** Scatter plot of the arc length $a_{e,F}$ of the filament segments versus the Euclidean length $a_{e,E}$ of the corresponding edges e showed strong correlation for control and LatB-treated cells (Pearson correlation coefficient $c_P = 0.998$ and p-value $p_P < 10^{-50}$). Inset displays relative lengths $B = a_{e,E}^{-1} a_{e,F}$ ("bending") with an average of $\mathrm{E}[B] = 1.2 \pm 0.2$. **(G)** Distribution of the edge angles relative to the cell axis showed a prevalence of AFs parallel to the cell axis in the control cells but not in LatB-treated cells. **(H)** Time series and boxplots of the average path length ("reachability") for one control cell (green) and 20 randomized networks of the first null model for each time step (gray; mean ± standard deviation). **(I)** The extracted actin networks of the seven control cells showed significantly lower average path lengths ("reachability"), CV of the path lengths ("dispersal") and CV of the edge angles ("contortion") than their counterparts from the first null model. The algebraic connectivities ("robustness") and assortativities ("heterogeneity") of the actin networks were higher than expected from the first null model. Findings persist for the second null model (Appendix 6.3.4).

demonstrate that the extracted network representations of the actin cytoskeleton enable fully automated and unbiased phenotyping of cytoskeletal structures.

The actin cytoskeleton supports efficient transport

A major function of the actin cytoskeleton is to mediate transport of a range of organelles and compartments. To assess the efficiency of the actin networks in terms of transport, we computed a number of seminal transport-related network properties and compared them against ensembles of two types of randomized null model networks (cf. Fig. 4.4.2C for example network of first null model; cf. Appendix 6.3.4 for second null model).

We computed the average path length [West, 2001], which reflects the reachability of a network, and compared this against the ensemble of networks from the first null model (Fig. 4.4.2H). We found that the average path length of the extracted networks was smaller than that of the null model networks (Fig. 4.4.2I; $p_t < 10^{-50}$). This difference indicates that the actin cytoskeleton is tuned towards short path lengths. Similarly, the coefficients of variation (CVs) of the shortest path lengths in the extracted networks were smaller than expected from the null model networks ($p_t < 10^{-50}$), indicating that also fluctuations in the path length between any two nodes are maintained at a low level. Another classical transport-related network property is the algebraic connectivity [West, 2001], which reflects the redundancy of paths between any two nodes and thus captures the robustness of the transportation network against disruptions. The algebraic connectivity of the extracted networks was higher than expected by chance ($p_t < 10^{-14}$). In contrast, the LatB-disrupted actin cytoskeletons did not show any significant differences in their transport-related network properties as compared to the null model networks (Appendix 6.3.4). This supports the hypothesis that transport efficiency is a biological design principle of the actin cytoskeleton [Breuer et al., 2014].

To investigate the structural origin of this transport efficiency, we reconsidered the assortativity of the cytoskeleton and found that it was higher in the extracted networks than expected from the null model networks ($p_t < 10^{-50}$). Similarly, we found that the CV of the angles between AF segments and the major cell

axis was smaller in the extracted networks than expected by chance ($p_t < 10^{-50}$). Together with the observed transport efficiency, these findings suggest that the formation of connected patches of aligned actin bundles is a functionally relevant feature of the cytoskeleton.

To assure that our results were robust, we used an additional and more restricted null model, which only shuffles edge weights. While the first null model is more flexible and explores a larger space of random networks, the second null model excludes potential artifacts in the first null model that could arise from an increased number of edge crossings or a more homogeneous distribution of node positions as compared to the extracted networks (Appendix 6.3.4). Our findings from the first model were consistently confirmed by the second null model. Hence, differences in the studied network properties of extracted and null model networks are not an artifact of the randomization procedures.

Another potential issue, shared by all current approaches that extract transport-related networks from image data, is the unknown directionality of the edges. Our analyses of cytoskeletal transport capacity rely on the assumption of bi-directional transport along edges. Although individual AFs usually allow uni-directional movement of motor proteins only, bi-directional movement may occur due to fluctuations, different orientations of filaments in actin bundles or cytoplasmic streaming [Langford, 1995; Badoual et al., 2002; Lee et al., 2004]. Indeed, our data showed that only a small fraction of around 15% of actin edges showed predominantly uni-directional transport, irrespective of the actin bundle thickness (cf. Fig. 4.4.3 below and Appendix 6.3.6), justifying the assumption of bi-directional transport. Our analyses therefore indicate that transport efficiency is a central design principle of the actin cytoskeleton in growing plant cells.

Automated quantification of Golgi movement

To quantitatively investigate actin-based cellular transport, we studied the flow of Golgi bodies along the actin cytoskeleton in three-day-old FABD-GFP and tdTomato-CesA6 (tdT-CesA6) dual-labeled hypocotyl cells (cf. Appendix 6.3.1 for video of Golgi flow). The tdT-CesA6 marker was used as proxy for Golgi movement as the cellulose synthase labels Golgi bodies [Paredez et al., 2006; Luo et al., 2015]. We analyzed the flow of Golgi through automated tracking (Fig. 4.4.3A; [Jaqaman et al., 2008; Schindelin et al., 2012]) in image series from control and LatB-treated cells (Figs. 4.4.3B and C).

We found that average Golgi bodies moved with a velocity of $\mathrm{E}[v] = 0.36 \pm 0.26\,\mathrm{s}^{-1}\mu\mathrm{m}$ in control cells (Fig. 4.4.3D), which is higher than $\mathrm{E}[v] = 0.04 \pm 0.03\,\mathrm{s}^{-1}\mu\mathrm{m}$ in LatB-treated cells, consistent with previous studies [Nebenführ et al., 1999; Hawes and Satiat-Jeunemaitre, 2005; Akkerman et al., 2011]. The Golgi movement was predominantly parallel to the major cell axis in control cells while such preferential orientation was absent in LatB-treated cells (Fig. 4.4.3F). The preferred orientation of movement, therefore, correlates with the orientation of long actin bundles (cf. Fig. 4.4.2). Thus, our automated tracking captures known features of Golgi movement and may therefore be suitable for further, more detailed analyses of Golgi behavior.

Golgi bodies exhibit wiggling, which does not change over time or with distance to the actin cytoskeleton

The movement of Golgi bodies is characterized as saltatory or stop-and-go [Nebenführ et al., 1999; Nebenführ and Staehelin, 2001], whereby Golgi switch between periods of directed movement and undirected

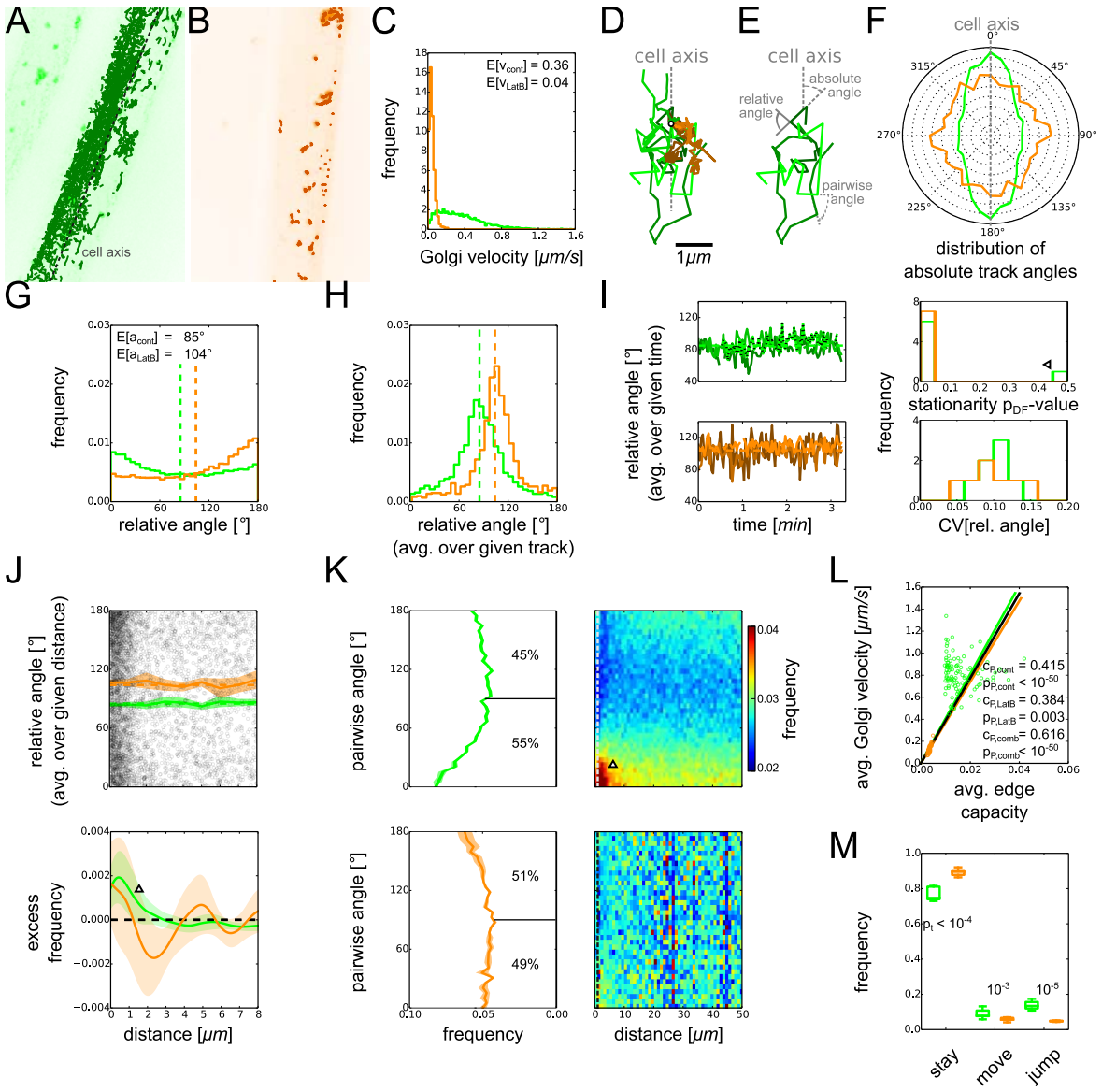


Figure 4.4.3: **Quantification of Golgi flow and movement along individual actin filaments.** Comparison of Golgi dynamics in cells of plants treated with the actin-disrupting drug LatB (orange) and untreated control plants (green). **(A)** Overlay of cellular recording of Golgi (green) and Golgi tracks throughout the recording from a control cell (dark green). **(B)** Overlay of cellular recording of Golgi (orange) and Golgi tracks throughout the recording from LatB-treated cell (dark orange). **(C)** Caption continued on next page.

(C) Distributions of Golgi velocities showed lower average velocities in LatB-treated cells, $\mathrm{E}[v] = 0.04 \pm 0.03\,\mathrm{s}^{-1}\mu\mathrm{m}$, than in untreated control cells, $\mathrm{E}[v] = 0.36 \pm 0.26\,\mathrm{s}^{-1}\mu\mathrm{m}$. **(D)** Examples of Golgi tracks from control and LatB-treated cells, with starting point shifted to the origin (cf. white dot) and rotated major cell axis (cf. gray dashed line). **(E)** Schematic of different angles used to study movement of Golgi bodies: The absolute angle measures the angle between a Golgi track segment and the major cell axis (cf. dashed gray angle). The relative angle measures the angle between two consecutive segments of a given Golgi track (cf. solid gray angle). The pairwise angle measures the angle between two segments of two different Golgi tracks at a given time step (cf. dotted gray angle). **(F)** Distributions of absolute angles, i.e. angles of Golgi track segments with respect to the major cell axis (cf. (A)), showed a preferential movement of Golgi parallel to the cell axis in control cells while such preference was absent in LatB-treated cells. **(G)** Distributions of relative angles, i.e. angles between consecutive Golgi track segments, were broad with averages $\mathrm{E}[a] = 85 \pm 55°$ and $\mathrm{E}[a] = 104 \pm 55°$ for control and LatB-treated cells, respectively (cf. dashed lines). **(H)** Relative angles, averaged over a given track, showed unimodal distributions with peaks around 85° and 104° for control and LatB-treated cells, respectively (cf. dashed lines). **(I)** Time series of average relative angles, averaged over a given time step, for each of the studied control and LatB-treated cells (left panels). All time series were stationary, i.e. they showed no increasing or decreasing trend, as judged from augmented Dickey-Fuller unit root test p-values $p_{DF} < 0.05$ (upper right panel), except one (cf. triangle and cf. dotted black line in upper left panel). The fluctuations of the time series were small, with $\mathrm{E}[\mathrm{CV}[a]] = 0.10 \pm 0.03$ (lower right panel). **(J)** Distribution of relative angles in dependence of distance of the track segments from the cytoskeleton (upper; circles show results for control cells). The average relative angles were largely independent of the distance from the cytoskeleton (upper panel; solid lines and shaded areas show mean ± standard deviation). In particular, the average relative angles matched the overall average relative angles 85° and 104° for control and LatB-treated cells, respectively (upper panel; dashed lines and cf. (G)). The frequency of Golgi at a given distance from the cytoskeleton showed an accumulation of Golgi at small distances up to $2\,\mu$m for both control and LatB-treated cells (lower panel; solid lines and cf. triangle) when compared to a null model in which Golgi were randomly and uniformly distributed across the cell area (lower panel; dashed black line). **(K)** Heat maps of the distributions of pairwise angles, i.e. angles between two segments of two different Golgi tracks at a given time step, in dependence of the spatial separation of the two track segments (right panels). Distributions of pairwise angles for small spatial separations below $1\,\mu$m (left panels and cf. dashed black lines in right panels). In the control cells, there was a tendency towards parallel or antiparallel movement of Golgi (upper right panel) and for distances below $10\,\mu$m, a large fraction of Golgi tracks showed parallel alignment (cf. triangle). Nevertheless, even at small spatial separations below $1\,\mu$m, a fraction of 45% of Golgi pairs exhibited antiparallel movement with angles above 90° (upper left panel). In the LatB-treated cells (lower panels), no preferred alignment was observed, irrespective of the spatial separation of the track segments. **(L)** Scatter plot of actin bundling, measured by the average edge capacity, and the average Golgi velocity showed positive correlations (Pearson correlation p-values $p_P < 0.05$) for control (green) and LatB-treated cells (orange) as well as the combined data (black). **(M)** Boxplots of the fractions of Golgi in control and LatB-treated cells which stayed close to the same filament segment within one time step (referred to as "stay"), those that likely moved to a different segment of the same filament ("move"), and those that moved to a different, non-adjacent filament ("jump"). Fraction of stationary Golgi was higher for LatB-treated than for control cells (independent two-sample t-test p-values $p_t < 0.05$ were considered significant).

"wiggling" behavior (Fig 4.4.3E). While it has been suggested that Golgi wiggling is not specific to individual Golgi bodies [Akkerman et al., 2011], it is yet unclear whether Golgi wiggling changes over time or depends on the distance of the Golgi from the actin cytoskeleton. To quantify these characteristics, we computed the angles between consecutive Golgi track segments (referred to as "relative angles"; cf. 4.4.3E) and refer to movement with relative angles above 90° as wiggling behavior.

The observed distributions of relative angles across the studied cells were broad, demonstrating that both largely uni-directional movement and wiggling behavior were present (Fig. 4.4.3G). For LatB-treated cells, the average relative angle $\mathrm{E}[a] = 104 \pm 55°$ was larger than for control cells with $\mathrm{E}[a] = 85 \pm 55°$ and wiggling was thus more common. To confirm that wiggling is not simply a feature of individual Golgi [Akkerman et al., 2011], we computed the relative angles averaged across a given track (Fig. 4.4.3H). The resulting distributions of angles peaked at around 90° for both untreated and LatB-treated cells, showing that the majority of Golgi tracks contained both periods of directed movement and wiggling behavior. Hence, our findings confirm that wiggling behavior is not specific to individual Golgi.

To test whether the prevalence of Golgi wiggling changes over time, we calculated the distribution of average relative angles over time (Fig. 4.4.3I, left). We found that Golgi motility did not change during the course of the time series, (Fig. 4.4.3I, upper right; only one of the studied cells showed a slight increase in Golgi wiggling over time). Moreover, the prevalence of Golgi wiggling showed only very minor fluctuations within and across time series (Fig. 4.4.3I, lower right; $\mathrm{E}[\mathrm{CV}] = 0.10 \pm 0.03$), indicating that a constant fraction of Golgi exhibited wiggling behavior over time. Our data therefore suggest that Golgi wiggling is a common and stable cellular phenomenon.

To study the effect of the distance between the actin cytoskeleton and Golgi in the context of Golgi wiggling, we computed the relative angles between consecutive track segments at a given distance from the cytoskeletal center line (Fig. 4.4.3J, upper). The frequency of Golgi was dependent on the distance to the AFs (Fig. 4.4.3J, lower), and we found that the Golgi density was highest within $2\,\mu$m from the AF center lines. Surprisingly, the prevalence of Golgi wiggling did not depend on the distance from the AFs. Furthermore, when considering our weighted network representation of the actin cytoskeleton, we found that Golgi wiggling in the vicinity of a given edge did not depend on the edge thickness (Appendix 6.3.5).

We note that the Golgi wiggling resembles the searching behavior of foraging animals [Edwards et al., 2007; Benhamou, 2007] or microbial motion [Matthäus et al., 2009] that has been suggested to optimize search efficiency [Viswanathan et al., 1999; Humphries et al., 2012; Humphries and Sims, 2014]. It is plausible that the Golgi, and the loosely associated trans-Golgi network/early endosomes, might scan the plasma membrane, endoplasmatic reticulum (ER) or other compartments for areas that need to exchange material. Assuming that these sites are not globally coordinated by the cell, the switching of Golgi between directed movement and wiggling behavior may therefore provide an efficient search strategy in this setting. This is compatible with proposed models of Golgi movement [Nebenführ and Staehelin, 2001; Hawes, 2004], such as the "vacuum cleaner" model (Golgi move through the cell and pick up products from the ER) or the "recruitment" model (Golgi pause in vicinity of active ER sites to facilitate trafficking). Our findings thus suggest a connection between Golgi wiggling and the optimization of uptake and delivery of Golgi-related material throughout the cell.

Golgi movement is largely dependent on direct, motor-based transport along actin bundles

Despite recent studies, it remains unclear whether Golgi bodies are transported through the cell by direct interactions with motor proteins or indirectly via cytoplasmic streaming [Buchnik et al., 2014; Cai et al., 2015]. The two scenarios are reflected in the behavior of neighboring Golgi tracks. Movement of close-by Golgi tracks in the same direction is indicative of indirect bulk flow. In contrast, movement in opposite directions suggests direct actomyosin-based transport of Golgi. To distinguish these two cases, we measured the angles between any two segments of different Golgi tracks within the same time step (referred to as "pairwise angles") in dependence of their spatial separation. For the LatB-treated cells, the frequency of a given pairwise angle of Golgi movement correlated with neither the angle nor the spatial separation of the two Golgi track segments (Fig. 4.4.3K, lower right). In contrast, the Golgi movement displayed mainly parallel or antiparallel trajectories in control cells (Fig. 4.4.3K, upper right), consistent with Golgi movement occurring preferentially along the major cell axis (cf. Fig. 4.4.3F). Even for small distances below 1 μm between different Golgi, a substantial fraction of Golgi trajectories showed antiparallel movement at a given time (Fig. 4.4.3K, left; 45% of $n \approx 5000$ pairwise angles across the studied cells and imaging periods). As the low Reynolds numbers of the cytoplasm favor laminar streaming [Luby-Phelps, 2000; Lew, 2005], this antiparallel movement of close-by Golgi contradicts the assumption of indirect, cytoplasmic-streaming induced movement and instead supports myosin-based transport of a substantial fraction of Golgi bodies.

Regions of bundled actin may lead to higher average velocities of Golgi movement in root epidermal cells [Akkerman et al., 2011]. To test if our network-based framework supports these findings in hypocotyl cells, we computed the average Golgi velocity and compared it to the overall actin bundling in the cell (Fig. 4.4.3L) as measured by the average edge capacity (cf. Fig. 4.4.2E). Indeed, actin bundling showed a high correlation with Golgi velocities for both control and LatB-treated cells (Pearson correlation coefficients $c_P \gtrsim 0.4$ and p-values $p_P < 0.05$).

Reasons for this correlation are manifold: Thick bundles are typically surrounded by fewer AFs that might slow down the Golgi (cf. Figs. 4.4.1 and 4.4.2I). The high rigidity of bundles increases the run length of motor proteins [Berger et al., 2015], which may be further extended through binding of multiple motor proteins [Klumpp and Lipowsky, 2005]. Furthermore, the varying orientations in an array of fine AFs are suggested to counteract cooperative movement of cargo [Akkerman et al., 2011]. Thus, the average velocity of Golgi in hypocotyl cells is determined by the prevalence of actin bundles. However, studying overall actin bundling and average Golgi velocities does not consider the identity of individual AFs and the potential movement of Golgi along and between filaments.

The rate of cytoplasmic streaming controls Golgi redistribution among actin bundles

To assess Golgi behavior in the context of the detailed AF structure, we divided Golgi movement into three classes (Fig. 4.4.3M; Appendix 6.3.5): 1. Golgi that maintained positions along an edge in the actin network between consecutive time steps (around 80% and 90% of $n \approx 40000$ and $n \approx 25000$ Golgi track segments for control and LatB-treated cells, respectively). 2. Golgi that moved to a different edge along a path with angles smaller than 90° (around 10% and 5% for control and LatB-treated cells, respectively). Since individual AFs, and especially bundles, rarely exhibit strong bending [Claessens et al., 2006; Staiger et al., 2009], the Golgi in

this class likely moved along the same AF between the given time steps. 3. Golgi that moved to an altogether different AF (around 10% and 5% for control and LatB-treated cells, respectively). For the LatB-treated cells, the fraction of Golgi in the first class was higher than for the control cells ($p_t < 10^{-50}$). This increase is compatible with the general reduction of Golgi movement in LatB-treated cells (cf. Fig. 4.4.3D). As expected, also within a given control or LatB-treated cell, faster Golgi were more likely to reach different AFs (Appendix 6.3.5). Our data thus suggest that switching of Golgi to adjacent AFs is myosin-dependent, while switching to non-adjacent AFs is due to cytoplasmic streaming that may carry the Golgi over large distances.

Switching of cargo between different, intersecting filaments has been previously shown for organelles tracking along microtubules in animal cells [Ross et al., 2008; Bálint et al., 2013; Osunbayo et al., 2015]. However, these studies focused on movement and switching of cargo at filament intersections and did not investigate switching of cargo to non-adjacent filaments.

Local and global actin network architecture may be used to predict direction and velocity of Golgi movement

Our previous analyses assumed that the capacity of a given actin network edge, i.e. its average thickness, reflects its potential to transport cellular cargo (cf. Fig. 4.4.2). To test this hypothesis, we studied the Golgi flow on two levels: First, we computed pairwise correlations between the properties of Golgi flow and actin structures, as modeled by our extracted networks. Second, we combined different edge properties of the actin networks to predict features of Golgi flow (e.g. direction and velocity) using a multiple linear regression approach [Freedman, 2009].

We first extracted the capacity-weighted actin networks from FABD-GFP and tdT-CesA6 dual-labeled seedlings (Figs. 4.4.4A and B). We determined the local edge capacities and global edge properties that incorporated information about the importance of any given edge in the network context, i.e. edge degree, edge page rank as well as edge path and flow betweenness (Appendix 6.3.3). Edge path and flow betweenness measure the number of shortest paths and the total amount of flow that traverse a given edge when solving the short path and the maximum flow problem, respectively [Freeman, 1977; Brandes, 2008; Newman, 2009]. The edge page rank (Fig. 4.4.4C; [Brin and Page, 1998; Newman, 2009]; hereafter referred to as edge rank) quantifies the expected frequency of cargo at a given edge when the cargo is assumed to randomly switch between edges. The edge degree is given by the sum of adjacent edge capacities.

In parallel, we used the tdT-CesA6 image data to automatically track Golgi movement (Fig. 4.4.4E). From the resulting track segments at each time step, we constructed an auxiliary network that we refer to as Golgi flow network: We copied the structure of the actin network at the given time step and assigned new edge weights in the Golgi flow network according to various features of Golgi movement in the vicinity of the respective edge. To this end, we weighted the Golgi flow network edges by the number of Golgi track segments in the vicinity (Fig. 4.4.4F), by the average intensity (Fig. 4.4.4G) or by the direction and velocity of close-by Golgi (Fig. 4.4.4H).

To investigate the relationship between actin structure and Golgi dynamics, we first computed the correlation between the determined edge properties of actin and Golgi flow networks. For instance, we studied the dependence of the Golgi direction and velocity on the actin edge rank. The correlation between the two

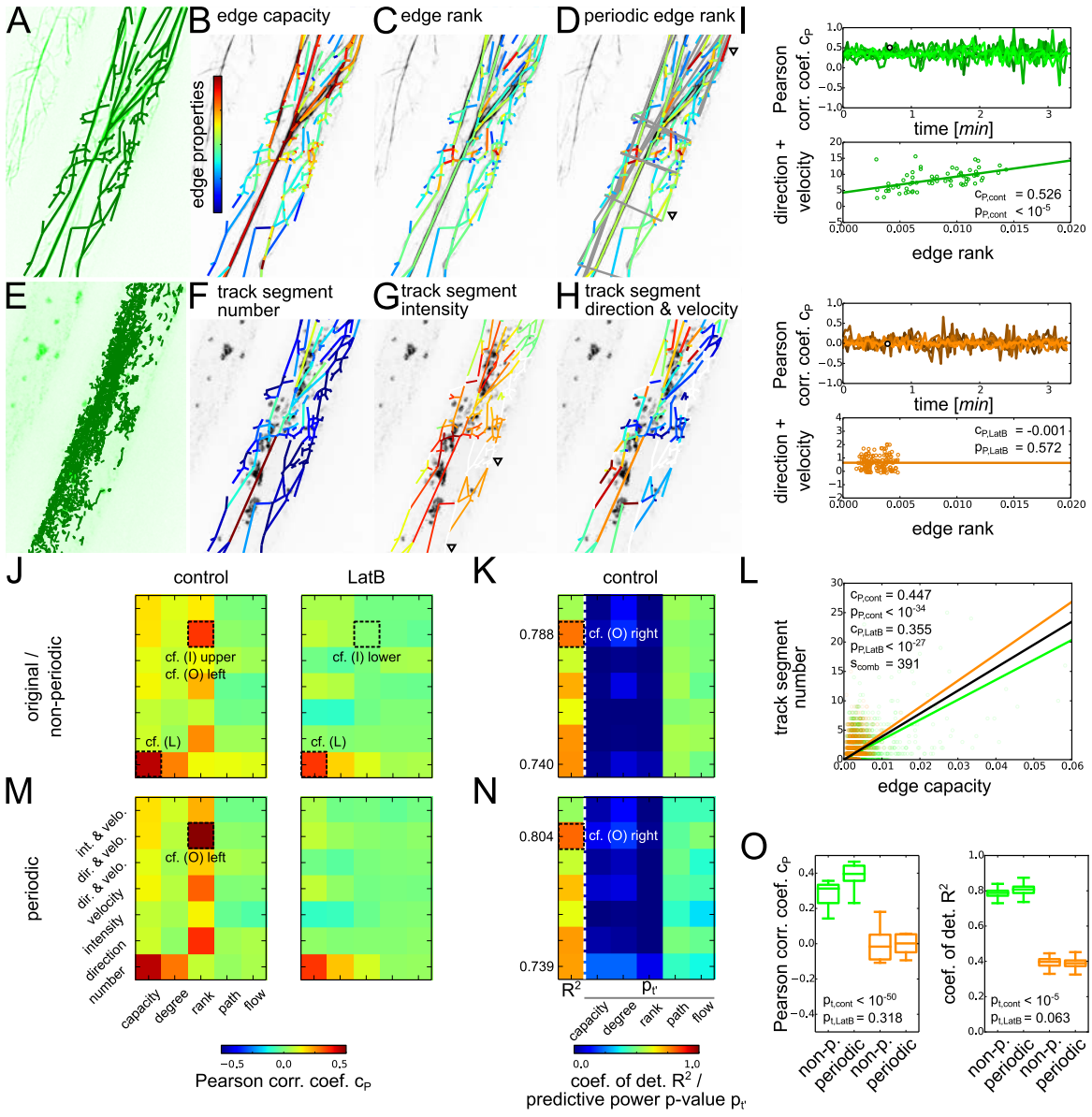


Figure 4.4.4: **Coordination and prediction of Golgi dynamics by whole-cell organization of actin cytoskeleton. (A)** Overlay of cellular recording of actin cytoskeleton (green) and extracted actin network (dark green). **(B)** Extracted actin network with edge colors representing their capacities, i.e. average thicknesses. **(C)** Extracted actin network with edge colors representing their edge ranks. **(D)** Caption continued on next page.

(D) Extracted actin network with cylindrical periodic boundary conditions. Edge colors represent their edge rank and gray edges indicate network edges added to implement periodic boundary conditions. Edge ranks in the periodic and non-periodic networks generally differ (cf. (C) and triangles). **(E)** Overlay of cellular recording of Golgi (green) and Golgi tracks throughout the recording (dark green). **(F)** Golgi flow network with edge colors representing the numbers of close-by Golgi track segments, i.e. track segments with starting points within 20 pixels from a given edge. **(G)** Golgi flow network with edge colors representing average intensity of close-by Golgi track segments. **(H)** Golgi flow network with edge colors representing the direction and velocity of close-by Golgi track segments. **(I)** Time series of Pearson correlation coefficients c_P between Golgi track segment direction and velocity and actin edge rank (upper panel) and exemplary scatter plots for one time point for a control and a LatB-treated cell, respectively (lower panel; white dots in upper panels indicate cells and time steps used for illustration of correlations). Across the studied control cells, the average correlation coefficient was $c_P = 0.354$ while for the studied LatB-treated cells, no significant correlation was observed with $c_P = 0.008$. **(J)** Heat maps of Pearson correlation coefficients c_P between different edge properties of original, non-periodic actin and Golgi flow networks for control (left panel) and LatB-treated cells (right panel). For control cells and LatB-treated cells, there was a significant correlation between the actin edge capacity and the Golgi number (cf. also (L)). The significant correlations between the actin edge rank and several direction- and velocity-related Golgi properties for the control cells were absent for the LatB-treated cells (cf. e.g. (I)). **(K)** Heat map of coefficients of determination R^2 for multiple linear regressions of the Golgi flow network edge properties (first column) and the respective predictive power p-values $p_{t'}$ of the original, non-periodic actin edge properties that were used as predictors (last five columns; one-sample two-sided t-tests p-values $p_{t'}$ measure whether inclusion of the respective predictors improves the prediction and $p_{t'} < 0.05$ were considered significant) for the control cells. As for the correlation coefficients (cf. (J)), the interdependence between actin edge properties and the direction- and velocity-related Golgi properties was highest ($R^2 > 0.7$). In the multiple linear regressions, the actin capacity, degree and rank were more reliable predictors ($p_{t'} < 0.05$) than edge flow and path betweenness ($p_{t'} \geq 0.05$). **(L)** Scatter plots of the number of Golgi less than 20 pixels from a given edge in dependence of the edge capacity showed positive correlations with $c_P = 0.447$ for control and $c_P = 0.355$ for LatB-treated cells. The slope of a linear regression for the combined data was $s = 391$. **(M)** Heat maps of Pearson correlation coefficients c_P between different edge properties of periodic actin and Golgi flow networks for control (left panel) and LatB-treated cells (right panel). In particular, some correlations involving the direction and velocity of Golgi were higher for the periodic than for the non-periodic networks (cf. (O)). **(N)** Heat map of coefficients of determination R^2 for multiple linear regressions of the Golgi flow network edge properties (first column) and the respective predictive power p-values $p_{t'}$ of the periodic actin edge properties used as predictors (last five columns) for the control cells. Analogue to the correlation-based analyses, the coefficients of determination especially of the direction- and velocity-related Golgi properties were higher for the periodic than for the non-periodic networks (cf. (O)). **(O)** Boxplot of Pearson correlation coefficients between Golgi direction and velocity and actin edge rank without and with periodic boundary conditions, respectively (left panel; independent two-sample t-tests p-values $p_t < 0.05$ were considered significant). Boxplot of coefficients of determination for predictions of Golgi direction and velocity without and with periodic boundary conditions, respectively (right panel). The correlation as well as the prediction quality for actin networks incorporating periodic boundary conditions were significantly higher than for networks without periodic boundary conditions for the control cells ($p_t < 10^{-50}$), but not for the LatB-treated cells ($p_t \geq 0.05$).

properties varied over time and across cells (Fig. 4.4.4I, upper). It was typically significant for control cells, while no significant correlation was found for the LatB-treated cells (Fig. 4.4.4I, lower; average Pearson correlation coefficients across all studied cells were $c_P = 0.354$ and $c_P = 0.008$ for control and LatB-treated cells, respectively, cf. Fig. 4.4.4J). These findings are compatible with the severely reduced flow (cf. Fig. 4.4.3D) and increased wiggling behavior of Golgi (cf. Fig. 4.4.3G) in LatB-treated cells.

In a more systematic approach, we evaluated the correlations between all pairs of edge properties, averaged across the studied cells and time points (Fig. 4.4.4J). Some pairs of properties, such as Golgi direction and velocity and actin edge rank discussed above, were correlated for the control cells ($|c_p| > 0.2$) but not for the LatB-treated cells ($|c_p| \leq 0.2$; cf. Fig. 4.4.4I). Only the number of Golgi close to a given edge was correlated with the respective edge capacity and edge degree for both control and LatB-treated cells (Fig. 4.4.4L). These findings show that although the flow of Golgi is severely altered by the LatB treatment, Golgi may still agglomerate in the vicinity of the actin stubs (cf. Fig. 4.4.3J, lower). However, for most pairs of actin and Golgi flow network edges properties, there was no or only weak correlation (Pearson correlation coefficients $|c_p| \leq 0.2$). Hence, while in particular edge flow and path betweenness have been suggested to predict transport in real-world networks [Borgatti, 2005; Kurant and Thiran, 2006; Jiang et al., 2008; Kazerani and Winter, 2009], they are not predictive of Golgi transport along the actin cytoskeleton in plant cells. The latter may be due to different transport requirements in particular regions of the cell, especially during hypocotyl elongation growth [Gendreau et al., 1997; Geisler et al., 2008].

Since pairwise correlations were only of moderate value, we secondly used multiple linear regression to see if certain aspects of Golgi flow could be predicted from a combination of actin edge properties. Indeed, the number of Golgi close to an actin edge (Fig. 4.4.4K; coefficient of determination $R^2 = 0.74$) and the Golgi direction and velocity ($R^2 = 0.79$) were accurately predicted from the edges properties of the underlying actin cytoskeletal network. Moreover, edge capacity, edge degree and edge rank of the actin network had higher predictive power (one-sample two-sided t-tests p-values, denoted by $p_{t'}$, $p_{t'} < 0.05$ for most Golgi flow properties) than the edge path and flow betweenness ($p_{t'} \geq 0.05$).

As edge capacity and edge degree reflect (semi-)local actin bundling, their observed high predictive power supports the finding that actin bundling is correlated with Golgi density and velocity (cf. Figs. 4.4.3J and L; [Akkerman et al., 2011]). As indicated above, the edge rank measures the (global) importance of an edge in the network context, and corresponds to the probability that an agent that randomly traverses the network is found at the given edge [Pinski and Narin, 1976; Brin and Page, 1998]. In our actin cytoskeletal transportation network, the edge rank thus models cargo that randomly switches between adjacent filament segments whereby thicker filaments are frequented with higher probabilities. Therefore, the system-wide organization of the actin cytoskeleton shapes, and may be used to predict, the dynamic flow of Golgi.

Incorporation of cylindrical geometry of the actin network enhances Golgi flow predictions

The two-dimensional rendering of images captures only a part of the epidermal actin cytoskeleton. This restriction introduces boundaries and, hence, bias in the extracted network as the cortical actin cytoskeleton follows the near-cylindrical shape of the growing hypocotyl cells. Since imaging three-dimensional time series of the actin cytoskeleton is intrinsically challenging, we modeled the cylindrical geometry of the cor-

tical cytoskeleton by periodically extending the original, two-dimensional extracted network (Fig. 4.4.4D). We refer to these networks as periodic, and the ones with boundaries as non-periodic.

We recomputed the edge betweenness, and repeated both our correlation-based (Fig. 4.4.4M) and regression-based analyses, for the periodic networks (Fig. 4.4.4N). For the control cells, Golgi velocity and direction showed highest correlation with, and were most accurately predicted by, the global edge rank rather than other local and global edge measures of the actin network. These data are in agreement with the non-periodic networks above. Intriguingly, the correlation between actin edge rank and Golgi direction and velocity, as well as the prediction of the latter, was improved for the periodic networks as compared to the non-periodic ones (Fig. 4.4.4O; independent two-sample t-test p-values $p_t < 0.05$). For the LatB-treated cells, no differences in the correlation coefficients ($p_t = 0.318$), nor in the coefficients of determination ($p_t = 0.063$), were observed between the periodic and non-periodic networks.

Our periodic boundary conditions conform to parsimony by assuming identical actin structures at the two sides. Although it is possible that the actin cytoskeleton at the distant periclinal side of the cell may differ from the epidermal periclinal side, e.g. due to different mechanical forces inside the hypocotyl, three-dimensional imaging will be necessary to resolve such differences. However, imaging with sufficient spatio-temporal resolution, and without substantial photobleaching, to accurately capture the fast dynamics of actin rearrangement and Golgi movement in distant parts of the cell introduces major limitations. In addition, absorption and scattering of light by plant cell features will result in images of reduced quality and may not resolve fine AFs. Therefore, our implementation of periodic boundary conditions appears reasonable until high-quality data of the complete cortical actin cytoskeleton become available. Taken together, Golgi transport is not merely determined by the local structure of the cortical cytoskeleton, but depends on larger architectural contexts, as well as the cylindrical geometry of the hypocotyl cell.

4.4.4 Concluding remarks

Advances in determining the genetic and molecular basis of cytoskeletal transport have not yet been matched by a systems perspective. To address this gap, we introduced an accurate image-based network representation of the actin cytoskeleton to facilitate automated and unbiased quantification of cytoskeletal phenotypes. We used this framework to investigate the relation between the properties of the actin cytoskeleton and features of Golgi transport in Arabidopsis hypocotyl cells.

We found that the actin cytoskeleton supports efficient transport processes. Through automated tracking of Golgi bodies, we quantified contributions of myosin- and cytoplasmic-streaming based transport. We demonstrated that Golgi wiggling is a prominent biological phenomenon that is reminiscent of optimized search strategies that may guarantee efficient uptake and/or delivery of Golgi-related cell material. Furthermore, we found that Golgi dynamics, in particular velocity and directionality of movement, can be predicted from the structure of the actin network. These data highlight the importance of objectively quantifying the actin cytoskeleton organization and its cylindrical geometry to obtain a systems view of cellular transport.

Despite the diversity of cellular transport processes, our approach of integrating cytoskeletal network structures with tracking data of organelles is directly transferable to various biological systems and functions: In

plants, transport of mitochondria [Akkerman et al., 2011; Wang and Hussey, 2015] or photodamage avoidance movement of chloroplasts [Kasahara et al., 2002] represent interesting test grounds. In animals, it has been shown that cytoplasmic streaming in *Drosophila melanogaster* oocytes is related to structural features of microtubules [Ganguly et al., 2012], and that transport of lysosomes in monkey kidney cells depends on microtubules and microtubule cross-overs [Bálint et al., 2013]. While these are interesting local correlations of cytoskeletal features and organelle transport, we expect broader, system-level understanding of these processes by the application of approaches such as ours.

Our automated framework paves the way for quantitative descriptions of changes in the actin cytoskeleton and trafficking characteristics in, for example, large-scale chemical and genetic screens. This should certainly improve our understanding of how the actin cytoskeleton impacts on organelle transport. Moreover, our prediction of organelle transport from properties of the actin network indicates that network-based models may be used to reverse-engineer or predict potential uptake or deposition sites of Golgi-related material across the cell. With advances in imaging technologies, our framework can also be readily used to investigate how differences in cellular transport efficiency between cell types relate to mechanical properties or growth phenotypes. Altogether, the presented combination of experimental imaging techniques and theoretical network-based analyses provides an important step towards a systems understanding of cytoskeletal organization and transport dynamics and, ultimately, control of cytoskeleton-based transport.

4.4.5 Materials and Methods

Plant material and experimental setup

We used Arabidopsis Columbia-0 35S:FABD-GFP and pCesA6:tdT-CesA6 dual-labeled seedlings [Sheahan et al., 2004; Sampathkumar et al., 2013] to study actin cytoskeleton and Golgi bodies. The seedlings were surface sterilized (ethanol), stratified for 2 days at 4°C and germinated on MS agar plates (1X Murashige and Skoog salts, 8 L^{-1}g agar, 1X B5 vitamins, and 10.8 g 8 L^{-1}g sugar). All plants were grown in the dark on vertical plates at 21.8°C for 3 days. For drug treatment, seedlings were floated on distilled water containing 150 μM LatB, and a set of control seedlings on pure water in 6-well plates. The seedlings were incubated in the dark with gentle shaking for 4 hours before imaging. To immobilize the seedlings and to avoid mechanical damage, they were mounted between a cover glass and a 1 mm thick 1% agar pad affixed on a circular cover slip. A spinning-disk confocal microscope was used to capture rapid changes and to minimize bleaching, yielding images with a spatial resolution of 0.133 pixel^{-1} μm [Sampathkumar et al., 2011]. Exposure times were 400 ms for FABD-GFP and 300 ms for tdT-CesA6 with a time interval of 2 s between subsequent actin and Golgi images, respectively. The cells were recorded for at least 100 frames, i.e. about 3 min. Only seedlings expressing both fluorescent markers were used for further analyses. Here, seven recordings of different control and LatB-treated cells were analyzed, respectively.

Image preprocessing of actin and Golgi recordings

We preprocessed the confocal recordings using the image processing package Fiji [Schindelin et al., 2012] (cf. Fig. 4.4.1A): We corrected the potential drift of the seedlings under the microscope by applying the

Fiji-StackReg stack registration algorithm to the image series [Thévenaz et al., 1998]. To enable simultaneous registration of the dual-labeled plant recordings, we merged actin and Golgi recordings from one cell as different color channels and split the channels after registration. We compensated photobleaching by normalizing the mean intensity of each frame. We improved the signal-to-noise ratio by using the Fiji-BackgroundSubtraction rolling ball filter with radius of 50 pixels [Sternberg, 1983]. To automatically determine the cellular region of interest for each cell, we performed a maximum projection of the registered actin recordings, applied a Gaussian filter with a standard deviation of 20 pixels, thresholded the image with a global Otsu threshold [Otsu, 1975] and selected the largest connected component as the region of interest for both actin and Golgi images. Moreover, to identify the cell axis, we skeletonized the binary representation of the cellular region of interest and selected the pixels along 20% and 80% of the length of the resulting center line to compute the angle γ of rotation of the cell (cf. Fig. 4.4.2G).

Extraction of actin networks from image data

From the preprocessed image data, we represented the actin cytoskeleton as a network through a custom procedure which has been developed and implemented using Python [Van Rossum and Drake, 2011], SciPy [Olivier et al., 2002], the SciKit image processing toolbox [van der Walt et al., 2014] and NetworkX [Hagberg et al., 2008] (cf. Fig. 4.4.1 for illustration and http://mathbiol.mpimp-golm.mpg.de/CytoSeg/ for the open-source code of the implemented network extraction procedure): First, to obtain the filamentous actin skeleton, we applied a two-dimensional tubeness filter to each frame of the preprocessed actin images to enhance the signal of the filamentous structures of width ν_{width} [Sato et al., 1998] (cf. Fig. 4.4.1B). Next, we obtained binary images by applying an adaptive median threshold with blocksize ν_{thres} (cf. Fig. 4.4.1C). We determined the center lines of the actin structures by skeletonizing the thresholded image [Haralick et al., 1987], i.e. the skeleton and the background are given in a binary representation by 1- and 0-pixels, respectively. Then, we removed all spurious connected components of less than ν_{size} pixels in size as well as those whose average intensity in the original actin image was below ν_{int} of the average component intensity (cf. Fig. 4.4.1D). The image processing parameters ν_{width}, ν_{thres}, ν_{size} and ν_{int} were determined in a gauging step by comparison of automated segmentation to manually obtained gold standard segmentations (cf. Fig. 4.4.1J and below).

Second, for each skeletonized, binary actin image, we identified the nodes of the network as crossings or endpoints of filaments by checking the 3×3 pixels^2 neighborhood for each pixel and assigning a node if the center pixel was 1 and the outer ring of the neighborhood contained exactly one or more than two connected 1-pixels. By ignoring potential nodes with zero or two connected 1-pixels in the outer neighborhood we excluded disconnected pixels and pixels in the middle of a filament, respectively. In cases where several nodes touched, we removed all except for the one whose position coincided with the center of mass of the touching node pixels. Then, we labeled the node pixels and created an auxiliary image for which the background was set to -1, the filaments were set to 0 and the N nodes were labeled sequentially from 1 to N (cf. Fig. 4.4.1E).

Third, we constructed a weighted network by starting from an empty multigraph $G = (\mathcal{N}, \mathcal{E})$ with $N = |\mathcal{N}|$ nodes at positions $x_n \in \mathbb{R}^2$, $n \in \mathcal{N}$, and initially $E = |\mathcal{E}| = 0$ edges. We iteratively propagated the node labels to the eight neighboring pixels, given that these pixels were part of a filament, and aborted the iteration when

no filament pixels were left. Then, for each two neighboring pixels with different labels, $n \neq m$, $n, m \in \mathcal{N}$, we added an edge $e = (n, m)$ to the multigraph G. To obtain further information about the edges, we computed different properties parallel to the propagation of node labels: We defined the arc length $a_{e,F}$ of an edge e as the arc length of the corresponding filament segment. We computed it by creating another auxiliary image of zeros and, when propagating a node label, setting the value at the new position to the value of the old position of that array and adding $\sqrt{2}$ or 1 depending on whether the propagation step was diagonal or not. The arc length $a_{e,F}$ of an edge e was then given by the sum of values of the two neighboring pixels which were used to create the edge. Similarly, we measured the weight $a_{e,w}$ of an edge by using another auxiliary image and propagating the intensity of the original actin image, filtered with a Gaussian kernel with a standard deviation of 2 pixels, along the filaments. Again, $a_{e,w}$ was given by the sum of the values of the two different neighboring, propagated node pixels. Due to the Gaussian filtering, $a_{e,w}$ is a measure for the summed intensity in the neighborhood of the filament and, by construction, increases with the length of the filament. We therefore derived the capacity $a_{e,c}$ of an edge as the ratio of its weight and arc length, $a_{e,c} = a_{e,F}^{-1} a_{e,w}$, $a_{e,F} \geq 1$ by construction. The edge capacity reflects the average thickness of the filament segment and was used as a measure for the average potential amount of cargo that may traverse the edge (cf. Fig. 4.4.2). For the computation of shortest paths in the network, we further defined the length $a_{e,l}$ of an edge as the inverse of the capacity, $a_{e,l} = a_{e,c}^{-1}$. We normalized the edge capacities and lengths of a given network to one for better comparability. Additionally, we calculated the Euclidean edge length $a_{e,E}$ directly from the node positions and the bending of a filament segment was measured by the ratio of arc and Euclidean length, $a_{e,B} \equiv B = a_{e,E}^{-1} a_{e,F}$.

In general, the extracted network is a multigraph (e.g. two curved filaments may cross twice, leading to two edges between the same pair of nodes; cf. Fig. 4.4.1E). For simplicity, we projected the multigraph onto a simple graph by summing the multiple edge capacities $a_{e,c}$ and taking the minimum of the remaining multiple edge properties. This is justified by our assumption of current-like flow along the filaments which is additive in $a_{e,c}$, and our calculation of shortest paths which favor the minimum $a_{e,l}$ in the case of multiple edges. Since the network is not guaranteed to be connected, we simplified analyses by adding edges of minimum total Euclidean length to obtain a connected network. Their edge weights were given by the average intensity of the Gaussian filtered image along a connecting line of pixels between its two respective nodes, similar to the original edges above (cf. Fig. 4.4.1E). Network extraction procedures similar to ours are reviewed in Appendix 6.3.2.

Last, we derived several higher-level edge properties which do not reflect only the local structure of the cytoskeleton but capture global features of its organization (Appendix 6.3.3). As a simple measure of the importance of an edge e in the network context, we computed its degree $a_{e,\mathrm{deg}}$ in the line graph LG, i.e. the summed capacity of its adjacent edges. The line graph LG of a graph G has a node for each edge in G and an edge between two nodes if the corresponding edges are adjacent in G. Moreover, we derived the edge page rank $a_{e,\mathrm{rank}}$ of an edge e in the line graph LG [Brin and Page, 1998; Langville and Meyer, 2005]. To capture the importance of an edge e with respect to shortest paths in the network context, we further derived the edge path betweenness $a_{e,\mathrm{path}}$, given by the number of shortest paths between all pairs of nodes which traverse e [Freeman, 1977; Newman, 2009]. Finally, the edge current flow betweenness $a_{e,\mathrm{flow}}$ was computed as the sum of maximum flows through e between all pairs of nodes [Harris and Ross, 1955; Newman, 2009].

Gauging of network extraction parameters

To ensure an accurate network representation of the actin cytoskeleton, we performed extensive gauging of the four imaging parameters υ_{width}, υ_{thres}, υ_{size} and υ_{int}. To this end, we generated 20 contrived images of known, cytoskeleton-like structures and created manual segmentation of 20 biological cytoskeleton images as a gold standard for comparison against the automated segmentation results (Appendix 6.3.2). We note that the extraction of networks from the segmented center lines is deterministic, i.e. identical segmentations result in identical networks, and we therefore focused on comparing the segmented center lines.

We then varied all four parameter in a wide range with ten linear steps each, $\upsilon_{\text{width}} \in \{0.4,\ldots,2.2\}$, $\upsilon_{\text{thres}} \in \{21,\ldots,111\}$, $\upsilon_{\text{size}} \in \{2,\ldots,47\}$ and $\upsilon_{\text{int}} \in \{0.1,\ldots,1.9\}$ (cf. Figs. 4.4.1I and J) for all 40 gauging images. As a measure of agreement between the segmentations, we used the average minimum Euclidean distance between two pixels from the manual to the automated segmentation, $d_{\text{manu}\to\text{auto}}$, and vice versa, $d_{\text{auto}\to\text{manu}}$. Small values for $d_{\text{manu}\to\text{auto}}$ typically favor parameters that detect only thick actin bundles (cf. Fig. 4.4.1I, left), while small values for $d_{\text{auto}\to\text{manu}}$ typically yield parameters that overestimate the prevalence of filamentous actin in the cell (cf. Fig. 4.4.1I, right). Therefore, we minimized the Haussdorf distance $d_{\text{HD}} = \frac{1}{2}\left(d_{\text{manu}\to\text{auto}} + d_{\text{auto}\to\text{manu}}\right)$ to obtain a compromise between over- and undersegmentation (cf. Fig. 4.4.1J).

The optimal parameters and their confidence intervals were determined as follows: We randomly selected 40 of the 40 images, allowing duplicates, stored the parameters that minimized the average d_{HD}, repeated the procedure 500 times, and computed mean and standard deviation for each of the stored parameters (cf. Fig. 4.4.1K). Employing these optimized parameters guaranteed small average distances between the pixels of manual and automated segmentations and, hence, accurate network representations of the cytoskeleton.

Construction of Golgi flow networks from tracking data

To automatically track the movement of Golgi through the cell, we used Fiji-TrackMate to detect the Golgi as particles of 5 pixels in radius and discarded those with quality values below the 80th percentile [Jaqaman et al., 2008; Schindelin et al., 2012]. We then linked the Golgi in different frames using the linear assignment problem tracker with a maximum linkage distance of 24 pixels, a maximum gap-closing distance of 24 pixels and a maximum frame gap number of 4. The detection and tracking parameters were determined manually and while changes in the tracking parameters may alter the full Golgi tracks, the majority of our analyses focused on individual track segments which are more robust against changes in parameters. The tracking results enable detailed analysis of Golgi dynamics, e.g. over time or under different conditions (cf. Fig. 4.4.3).

Next, we constructed networks from the tracking data, referred to as "Golgi flow networks", for comparison with their actin cytoskeleton counterparts (cf. Fig. 4.4.4). For a given time step, we copied the nodes and edges of the actin network and computed the minimum distance $d_{t,e}$ between the edges e and the center points x_t of the tracking segments which originate in the respective time step. Each edge e of the Golgi flow network was then assigned different weighting factors (Appendix 6.3.3): the number $g_{e,n}$ of track segments which are closer to the considered edge than a cut-off distance of 20 pixels; the average intensity $g_{e,i}$ of close-by Golgi; the average velocity $g_{e,v}$ of close-by Golgi; the average angle $g_{e,d}$ between the edge and the close-by Golgi track segments, and combinations thereof.

A crucial step in the computation of several edge properties of the Golgi flow network involves the scalar product between a segment vector of a Golgi track and an actin network edge vector (i.e. the vector connecting the two edge's nodes). Although the actin filaments may be curved, we showed that the bending of a filament segment is typically very small with $\mathrm{E}[B] = 1.2 \pm 0.2$ (cf. Fig. 4.4.2E), justifying the assumption of straight edge vectors.

Randomization and null models of actin networks

We investigated the structure of the actin networks using a number of seminal and biologically relevant properties, such as assortativity and average path length (Appendix 6.3.3). While some of these properties may be interpreted by themselves (like the sign of the assortativity which provides information about about whether thick actin bundles are grouped together or whether they are intermingled with filaments), a suitable reference is needed to interpret others (like the average path length whose value depends, e.g. on the size of the network and the sum of edge weights).

Therefore, for any given actin network, we introduced two types of null models that randomize certain properties of the network while preserving relevant others (Appendix 6.3.4 and http://mathbiol.mpimp-golm.mpg.de/CytoSeg/ for the open-source code of the implemented randomization procedure). From both null models we generated an ensemble of $R = 20$ randomized networks for comparison with a given extracted network. For the first null model, the actin network was copied, its edges were divided into 10 bins according to their Euclidean length $a_{e,E}$ (bins were given by $[0, 10, 20, 30, 40, 50, 60, 70, 80, 90, \infty]$ pixels), saved in a temporary list (their other properties, like the capacity $a_{e,c}$, remained unchanged) and removed from the network. All nodes were distributed randomly and uniformly across the cell. Then, the longest edge in the temporary list was inserted between two random nodes whose discretized Euclidean distance matched the length bin of the edge. For each edge, 50 pairs of matching nodes were tested and the edge was inserted such that the number of crossings with already placed edges was minimized. We repeated the procedure until all edges were added to the network. Since the randomized network may be disconnected, we connected it by adding edges of minimum total Euclidean length, as discussed for the original networks above. Although this procedure increased the number of edges by a factor of $\mathrm{E}[r] = 1.12 \pm 0.07$ with respect to the original extracted networks, the increase is small, (Appendix 6.3.4). While the randomized networks are not guaranteed to be planar, the fraction of crossing edges is typically small, around $\mathrm{E}[X] = 0.14 \pm 0.11$ (Appendix 6.3.4). The obtained null model networks share several important features with the extracted cytoskeletal networks: Because the distribution of edge capacities is not changed, also their sum is preserved, reflecting the amount of filamentous actin in the cell. In addition, the distribution of edge lengths is identical to that of the actin cytoskeletal network.

We further employed a second, more restricted model which does not increase the number of edges and edge crossings and does not randomize the node positions (Appendix 6.3.4). These null model networks were generated by only shuffling the edge properties of the original network [Breuer et al., 2014; Breuer and Nikoloski, 2014]. The second null model, too, leaves the distribution of edge capacities unchanged.

Periodic boundary conditions for actin networks

To correlate the flow of Golgi with the actin network structures, we employed edge betweenness measures which assess the edges' relevance in their network context (cf. Fig. 4.4.4 and Appendix 6.3.3). Yet, due to the finite imaging domain, these properties are biased because, e.g. edges at the border of the cell will typically participate in fewer shortest paths than those in the center. In a plant cell, however, such bias is absent as the cortical cytoskeleton envelops the cell, forming a network on a cylindrical domain.

Therefore, we modeled the cylindrical geometry of the cortical cytoskeleton by introducing periodic boundary conditions, i.e. we assumed that the cytoskeleton at the back of the cell is identical to its imaged counterpart at the front. We implemented the boundary conditions for arbitrary cell shapes by augmenting the original network (cf. Fig. 4.4.4D and http://mathbiol.mpimp-golm.mpg.de/CytoSeg/ for the open-source code of the implementation of periodic boundary conditions): We started from the cellular region of interest and created an empty graph JG, termed jump network, with nodes and node positions given by the pixels bounding this region. We computed the angle γ of the cell axis as described above. Next, we rotated the node positions by an angle of $-\gamma$ and rounded the new coordinates to integer numbers. We added an edge to JG for each pair of nodes with the same x- or y-coordinates given that the connecting line was fully contained within the region of the cell. These links allow jumps parallel and perpendicular to the cell axis and were therefore assigned edge attributes $a_{e,J} = 1 \equiv a_{e,\parallel}$ for parallel and $a_{e,J} = 10^{-5} \equiv a_{e,\perp}$ for perpendicular jumps. Furthermore, we added edges between all neighboring boundary nodes to the jump network JG. We then coupled the nodes of the original actin network G to the nodes of jump network JG if their Euclidean distance $a_{e,E}$ was smaller than a threshold value of 10 pixels. Finally, we recomputed the shortest path lengths for all pairs of nodes of the original network using a modified Dijkstra's algorithm [Dijkstra, 1959] which increases the auxiliary path lengths by an arbitrarily large number if more than one parallel or perpendicular jump has been executed. Finally, we connected all nodes of the original actin network which were connected by a shortest path of less than 10 pixels in length and involved one parallel and/or perpendicular jump. Thus, the augmented network displays periodic, cylindrical boundary conditions. The properties of the new edges that were needed to assess the organization of the cytoskeletal network were computed as follows: The Euclidean and the filament arc lengths $a_{e,E}$ and $a_{e,F}$ were given by the sums of edge lengths for the respective shortest path, whereby jumps were excluded. The edge weight $a_{e,w}$ was given by the total intensity of the Gaussian filtered actin image, summed along the connecting lines of the path, again excluding jumps. The edge capacity $a_{e,c}$ and length $a_{e,l}$ were derived from these properties as before. For the augmented network with periodic boundary conditions, we recomputed the edge betweenness properties $a_{e,\mathrm{deg}}$, $a_{e,\mathrm{rank}}$, $a_{e,\mathrm{path}}$ and $a_{e,\mathrm{flow}}$. The procedure can be readily applied to realize periodic, cylindrical boundary conditions for cellular networks of arbitrary shapes.

4.4.6 Acknowledgements

Support: We thank Joachim Selbig for helpful comments on an early version of the manuscript. We thank Anna Stief for thorough comments on the manuscript.

Funding: D.B. acknowledges funding from an International Max Planck Research School scholarship. S.P. and Z.N. were funded by the Max Planck Society. S.P. was funded by a R@MAP Professorship

at University of Melbourne and is grateful for a Dyason travel grant. S.P. and Z.N. acknowledge an IRRTF (RNC) grant that aided in completing this work.

Contributions: D.B., A.I., Z.N. and S.P. designed the project; A.I. recorded the data; D.B. and J.N. implemented the methods; D.B. analyzed the data; D.B., A.I., J.N., S.P. and Z.N. wrote the manuscript.

4.5 DeFiNe: an optimization-based method for robust disentangling of filamentous networks

Publication: *Sci. Rep., 2015, 5:18267*

Authors: David Breuer[1,2,*], Zoran Nikoloski[1]

Affiliations: [1]Systems Biology and Mathematical Modeling, Max Planck Institute of Molecular Plant Physiology, Am Muehlenberg 1, 14476 Potsdam, Germany

[2]Institute of Biochemistry and Biology, University of Potsdam, Karl-Liebknecht-Straße 24-25, 14476 Potsdam, Germany

Contact: *breuer@mpimp-golm.mpg.de

4.5.1 Abstract

Thread-like structures are pervasive across scales, from polymeric proteins to root systems to galaxy filaments, and their characteristics can be readily investigated in the network formalism. Yet, network links usually represent only parts of filaments, which, when neglected, may lead to erroneous conclusions from network-based analyses. The existing alternatives to detect filaments in network representations require tuning of parameters over a large range of values and treat all filaments equally, thus, precluding automated analysis of diverse filamentous systems. Here, we propose a fully automated and robust optimization-based approach to detect filaments of consistent intensities and angles in a given network. We test and demonstrate the accuracy of our solution with contrived, biological, and cosmic filamentous structures. In particular, we show that the proposed approach provides powerful automated means to study properties of individual actin filaments in their network context. Our solution is made publicly available as an open-source tool, "DeFiNe", facilitating decomposition of any given network into individual filaments.

Keywords: polymers, cytoskeleton, networks, path cover, computational complexity

4.5.2 Introduction

Many network-like structures in nature are composed of filaments forming intricate interconnected arrays across different scales of organization. For instance, filamentous structures can be observed in networks of cellulose polymers in the primary cell wall of plants and algae [Stamm, 1964; Klemm et al., 2005], cytoskeletal networks of actin filaments or microtubules in cells across all domains of life [Shih and Rothfield, 2006; Liu, 2010; Wickstead and Gull, 2011], networks of neurons [Braitenberg and Schüz, 1998; Lichtman et al., 2008], root systems [Zhu et al., 2011; Galkovskyi et al., 2012; Lobet et al., 2015], as well as solar prominences [Gibson and Fan, 2006; Mackay et al., 2010] and galaxy clusters [Bond et al., 1996; Stoica et al., 2005; Bond et al., 2010;

Input	Method	Features					References
		curved filaments	filament -specific	intensity -based	automated	parsi -monious	
image	texture filter	−	−	+	+	+	[Boudaoud et al., 2014]
	linear programming	−	−	+	+	+	[Wood et al., 2013]
	rotating grid	−	+	+	+	+	[Jacques et al., 2013a]
	filament tracing	+	+	+	○	○	[Cohen et al., 1994; Meijering, 2010; Peng et al., 2015a]
	filament tracking	+	+	+	+	○	[Mayerich and Keyser, 2008]
	open contours	+	+	+	+	○	[Smith et al., 2010; Xu et al., 2015]
network	rule-based decomp.	+	+	−	○	○/−	[Leandro et al., 2009; Qiu and Li, 2014]
	filament cover	+	+	+	+	○/+	current work

Table 4.5.1: **Overview of different approaches for disentangling filamentous networks.** Two main classes of approaches to analyze the filamentous structure of networks can be distinguished, based on whether they operate on image data or on extracted networks. Irrespective of the class, the existing approaches vary in their capacity (+) or inability (−) to detect curved filaments, identify individual filaments, and to include information about the intensity/thickness of filaments. Further, the amount of manual user input as well as the number of parameters required by the algorithms can be feasible (+), laborious (−), or depends on the specific variant of the algorithm (○). For the network-based approaches, the number of required parameters may be different for the extraction of the network from image data and the consequent decomposition of the network into filaments (separated here by /).

Tully et al., 2014]. Network-based studies of these structures have already elucidated important aspects such as the mechanics of cellulose networks [Stamm, 1964; Moon et al., 2011], transport on cytoskeletal actin networks [Akkerman et al., 2011; Bálint et al., 2013], and connectivity patterns in the brain [Kandel et al., 2000; Sporns et al., 2005; Lichtman et al., 2008]. However, the network links usually correspond to segments of the filaments; therefore, the classical network-based analysis neglects the identities of individual filaments. A few powerful exceptions have recently started to emerge [Xu et al., 2014, 2015] which may identify multiple segments that belong to the same filament; yet, since these studies do not capture filament overlaps, filaments are still broken into potentially multiple fragments. Characterization of the mechanical- [Kumar et al., 2006; Bausch and Kroy, 2006; Lu et al., 2008], transport- [Bálint et al., 2013; Osunbayo et al., 2015], and information-transmission related properties [Eccles, 1982; Bennett, 1977] in such network representations may hence lead to erroneous conclusions due to their differences within and between filaments. Thus, analysis of filamentous structures rests upon accurate identification of individual filaments.

Since most of the filamentous structures in natural and man-made systems are studied by using imaging technologies, filaments are identified either directly from the imaging data or from networks extracted from

these data (see Tab. 4.5.1 for succinct review). In the first class of approaches, a texture-based method is employed to infer the overall orientation of objects in an image section [Boudaoud et al., 2014]. However, this method cannot be employed to pinpoint individual filaments. Another method decomposes entire images of filamentous structures into linear segments based on a linear programming formulation [Wood et al., 2013]. While this method utilizes few parameters (e.g., number of filaments), it only models and extracts a representative set of linear filaments. Moreover, filaments have been modeled as linear segments, detected by co-localization with a parallel grid at different orientations and by using manually chosen intensity thresholds along a filament [Jacques et al., 2013a]. While this method is fast and useful for extracting linear filaments (e.g., microtubules), it does not capture bent or tangled filaments and necessitates manual parameter selection. Alternatively, tracing- and tracking-based methods which start from one or multiple image points and predict neighboring points on a putative filament through optimization of an energy function are powerful methods for filament identification. Although these algorithms have led to great insights, especially into the connectome, they typically require user input and do not capture overlapping filaments [Cohen et al., 1994; Mayerich and Keyser, 2008; Meijering, 2010; Peng et al., 2015a]. Using a similar approach, open contour-based methods employ deformable curve models that elongate and align according to an energy functional to match the target filaments. Recent advances in open contour-based approaches enable fully automated filament detection [Xu et al., 2014, 2015], but can account for the overlap of only few filaments at the expense of parameter tuning [Smith et al., 2010].

The second class of approaches for disentangling filamentous structures employs a two-step procedure: First, weighted networks are extracted from image data from different systems and imaging sources. There is a large variety of algorithms for this task [Cohen et al., 1994; Baumgarten and Hauser, 2012; Obara et al., 2012b; Qiu and Li, 2014] which vary, in particular, in the number of parameters. Some of the methods from the first class, presented above, may also be used to obtain such network representations (e.g. [Mayerich and Keyser, 2008; Xu et al., 2015]). Second, the given, weighted networks are decomposed into filaments. The two existing methods for this task [Leandro et al., 2009; Qiu and Li, 2014] define specific junctions for bifurcations and crossings of filaments, depending on the distances between nodes, and assign filament identities according to manually chosen angle thresholds between incoming and outgoing edges. In particular, they strongly restrict the potential overlap of filaments and, due to the angle constraints, allow only crossing but no touching filaments. Most importantly, these methods require manual parameter selection and do not take into account filament intensity/thickness. We note that the step of decomposing a given network may also be beneficially applied to networks obtained, e.g., by open contour-based approaches in which filaments have been fragmented due to omission of filament overlaps [Xu et al., 2014, 2015].

Here, we propose a robust approach to decompose a weighted network into an optimal set of individual filaments. Therefore, our approach addresses the second step in the second class of approaches, presented above. The decomposition is based on a computationally difficult problem, referred to as filament cover problem (FCP), for which we propose suitable approximation algorithms. We test and demonstrate the accuracy of the findings from the approximation algorithms on artificial as well as biological and cosmic filamentous networks by comparison to manually obtained filament covers. In addition, we demonstrate that the proposed, fully automated solution allows facile characterization of well-studied properties of individual filaments, for which alternative approaches require parameter tuning or time-consuming manual tracing. The proposed approach is implemented in a publicly available open-source tool, “DeFiNe” (**De**composing

Filamentous **Ne**tworks), which can be used to decompose any given weighted network into a set of individual filaments for further analyses (http://mathbiol.mpimp-golm.mpg.de/DeFiNe/).

4.5.3 Methods

In this section we introduce the mathematical formulation of our optimization-based approach to decompose filamentous networks, demonstrate its computational intractability, and formulate a suitable approximation scheme. Moreover, we introduce new quality measures which take into account the underlying network structures for the comparison of the obtained filament decompositions with manual assignments used as a gold standard. Finally, we provide a brief overview of the studied data from different biological and physical systems. While we believe that these more technical explanations may promote a deeper understanding of our and related approaches, we encourage readers familiar with the aforementioned topics to proceed directly to the Results.

Mathematical formulation of the filament cover problem

Any filamentous structure may be represented as weighted geometric graph $G = (\mathcal{N}, \mathcal{E})$ with $N = |\mathcal{N}|$ nodes and $E = |\mathcal{E}|$ undirected, weighted edges. Edges represent filament segments and their intensities or thicknesses are reflected by their weights w_e, $e := (n_0, n_1) \in \mathcal{E}$ and $n_0, n_1 \in \mathcal{N}$. Nodes represent endpoints of filament segments and their positions are denoted by v_n, $n \in \mathcal{N}$, whereby, typically, $v_n \in \mathbb{R}^2$ or $v_n \in \mathbb{R}^3$ for networks extracted from image data.

We naturally represent a filament by an edge-path, $p = \left(e_{p,1}, \ldots, e_{p,P}\right)$, $e \in \mathcal{E}$, i.e., by an ordered sequence of $P = \left|p\right|$ adjacent edges, where $e_{p,i}$ denotes the i-th edge of filament p. The quality of a given filament p is assessed by the pairwise filament roughness

$$r_{p,\text{pair}} = \begin{cases} (P-1)^{-1} \sum_{i=1}^{P-1} \left| w_{e_{p,i+1}} - w_{e_{p,i}} \right| & , P > 1 \\ w_{e_{p,1}} & , P = 1 \end{cases}, \tag{4.5.1}$$

where $w_{e_{p,i}}$ denotes the weight of the i-th edge in filament p. The pairwise filament roughness is small if the edge weights along a filament vary smoothly, as expected for natural filaments (but cf. Discussion). For filaments that consist of one edge only, their roughness is given by their edge weight to increase the flexibility of our approach (cf. Appendix 6.4.1). Other roughness measures may be readily introduced that take into account filament thicknesses or alignments. As an additional example, we study the all-to-all filament roughness

$$r_{p,\text{all}} = \begin{cases} (P-1)^{-1} \max\limits_{i,j \in \{1,\ldots,P\}} \left| w_{e_{p,i}} - w_{e_{p,j}} \right| & , P > 1 \\ w_{e_{p,1}} & , P = 1 \end{cases}, \tag{4.5.2}$$

which is the average maximal difference between any two edge weights in a filament p, and again the original weight of the edge is used for a filament of length one. Taking into account that most filaments are only moderately bent, we further consider the maximal filament deflection angle between adjacent edges of a

path p,

$$r_{p,\text{angle}} = \max_{i\in\{1,\ldots,P-1\}} \text{angle}\left(v_{e_{p,i+1,1}} - v_{e_{p,i+1,0}}, v_{e_{p,i,1}} - v_{e_{p,i,0}}\right) \tag{4.5.3}$$

where $v_{e_{p,i,0}}$ and $v_{e_{p,i,1}}$ denote the positions of the start and end nodes of the i-th edge of filament p, respectively. Moreover, $\text{angle}\left(v, v'\right) := \arccos\left(\frac{v\cdot v'}{\sqrt{v\cdot v}\sqrt{v'\cdot v'}}\right)$ is the Euclidean angle of two vectors v and v' and $r_{p,\text{angle}} = 0°$ corresponds to perfectly straight alignment.

The optimal decomposition of a network into individual, smooth filaments then corresponds to solving our filament cover problem (FCP; cf. Appendix 6.4.1 for an overview of related cover problems):

> Given a graph $G = (\mathcal{N}, \mathcal{E})$ and the set $\mathcal{P}$ of all edge-paths in G with roughnesses r_p, $p \in \mathcal{P}$:
>
> Find a subset $\mathcal{P}_{\text{fil}} \subseteq \mathcal{P}$ with minimal total (or average) roughness R such that each element in $\mathcal{E}$ is covered (at least) once.

Here, edges that are covered by more than one path naturally correspond to filament overlaps. Minimizing the average instead of the total roughness yields shorter filaments, as appropriate for some networks (cf. Appendix 6.4.1).

Computational intractability of the filament cover problem and approximation algorithm

The FCP is computationally intractable on general and even planar graphs (cf. Appendix 6.4.2 for motivation and proof). Graphs generated from two-dimensional image data are planar by construction [Baumgarten and Hauser, 2012; Obara et al., 2012b]. The proof is by reduction from the well-studied Hamiltonian path problem which asks, for a given network, whether there is a sequence of adjacent nodes that includes each node exactly once, and which is known to be intractable on planar graphs [Garey et al., 1976]. Moreover, we outline an algorithm for solving the FCP in polynomial time on trees (cf. Appendix 6.4.3).

Since the FCP is computationally intractable on general and even planar graphs, we devise an approximation scheme by formulating the FCP as a fractional integer linear program (cf. Appendix 6.4.4 for motivation and details). For a given set $\mathcal{P}' \subseteq \mathcal{P}$ of input paths with pairwise filament roughnesses r_p, $p \in \mathcal{P}'$, we solve:

$$\begin{aligned} \text{minimize} \quad & \frac{\sum_{p\in\mathcal{P}'} r_p x_p}{\left(\sum_{p\in\mathcal{P}'} x_p\right)^A} \\ \text{subject to} \quad & \sum_{p:e\in p} x_p \geq 1 \,\text{for all}\, e \in \mathcal{E} \\ & x_p \in \{0,1\} \,\text{for all}\, p \in \mathcal{P}', \end{aligned} \tag{4.5.4}$$

where we use $r_p \in \left\{r_{p,\text{pair}}, r_{p,\text{all}}\right\}$ (Eqs. 4.5.1 or 4.5.2; referred to as *pair* and *all*). In the first line, $A \in \{0,1\}$ determines whether the total or the average roughness is minimized (*total*/*avg*). The inequality in the second line allows overlapping filaments and equality holds for an exact cover (*over*/*exact*). For $A = 0$, Eq. 4.5.4

is a binary linear program and for $A = 1$, the fractional problem Eq. 4.5.4 may be rewritten as a binary linear program (cf. Appendix 6.4.4). Binary linear programs may be solved using well-established and efficient algorithms [Schrijver, 1998; Linderoth and Ralphs, 2005].

To solve the FCP for a given network, we further need to collect a set of input paths $\mathscr{P}' \subseteq \mathscr{P}$. Since for a general graph it is not feasible to collect all paths $\mathscr{P}$ (cf. Appendix 6.4.2), we propose two approaches (referred to as *RMST* and *BFS*): (1) We create $T = 100$ random minimal spanning trees (RMST) of G whose $N(N-1)/2$ non-trivial, undirected paths are added to our set $\mathscr{P}'$. To obtain a RMST, each edge is assigned a randomly and uniformly distributed weight and the minimum spanning tree with respect to these weights is computed. (2) We perform a modified breadth-first search (BFS) on the nodes, stop the search for a path p when it violates the straightness criterion $r_{p,\text{angle}} < 60°$ (cf. Eq. 4.5.3), and add all permitted paths to $\mathscr{P}'$. We note that for real-world filamentous graphs, the number of nodes and their degrees are constrained by the filament thickness, while the number of considered loops is further restricted by the straightness criterion, so that our heuristically modified BFS yields a representative set $\mathscr{P}'$ of paths in reasonable time. Moreover, we note that the 60°-criterion is introduced for computational reasons and provides a tolerant estimate for maximal bending of the studied real-world filaments which are typically less bent.

Quality assessment of filament covers via structure-aware partition similarity measures

The accuracy of the filaments covers obtained by solving the FCP is assessed by comparison to manual filament assignments (cf. Fig. 4.5.1B). We quantify the similarity of the two partitions of the set of edges into (potentially overlapping) filaments using the variation of information,VI, the Jaccard index, JI, and the Rand index, RI,

$$\begin{aligned} \mathrm{VI}\left(\mathscr{C},\mathscr{C}'\right) &= 1 + \left(U \log U\right)^{-1} \cdot \\ &\quad \cdot \sum_{i,j} g_{i,j}\left(\log\left(\frac{g_{i,j}}{g_{\cdot,j}}\right) + \log\left(\frac{g_{i,j}}{g_{i,\cdot}}\right)\right), \end{aligned} \tag{4.5.5}$$

$$\mathrm{RI}\left(\mathscr{C},\mathscr{C}'\right) = \frac{h_{=,=} + h_{\neq,\neq}}{h_{=,=} + h_{=,\neq} + h_{\neq,=} + h_{\neq,\neq}}, \tag{4.5.6}$$

$$\mathrm{JI}\left(\mathscr{C},\mathscr{C}'\right) = \frac{h_{=,=}}{h_{=,=} + h_{=,\neq} + h_{\neq,=}}, \tag{4.5.7}$$

where $U = \sum_{i=1}^{C} |\mathscr{C}_i| = \sum_{j=1}^{C'} \left|\mathscr{C}'_j\right|$, $g_{i,j} = \left|\mathscr{C}_i \cap \mathscr{C}'_j\right|$, $g_{\cdot,j} = \sum_{i=1}^{C} g_{i,j}$, and $g_{i,\cdot} = \sum_{j=1}^{C'} g_{i,j}$ [Saporta and Youness, 2002; Meilă, 2005; Denœud and Guénoche, 2006]. The contingency tables $h_{\times,\times'}$, $\times, \times' \in \{=, \neq\}$, provide the numbers of edge pairs which are in the same or different sets in the two partitions, respectively. While these classical measures are widely used [Meilă, 2005; Lancichinetti and Fortunato, 2009b], they may generally yield opposing results and VI is not well-defined for overlapping partitions (cf. Appendix 6.4.6). More severely, these measures do not take into account the structure of the graph underlying the partitions. To remedy this shortcoming, we introduce a suite of measures, the structure-aware Rand and Jaccard indices (cf. Eqs. 4.5.6 and 4.5.7),

$$\mathrm{RI}^d\left(\mathscr{C},\mathscr{C}'\right) = \frac{h^d_{=,=}+h^d_{\neq,\neq}}{h^d_{=,=}+h^d_{=,\neq}+h^d_{\neq,=}+h^d_{\neq,\neq}}, \tag{4.5.8}$$

$$\mathrm{JI}^d\left(\mathscr{C},\mathscr{C}'\right) = \frac{h^d_{=,=}}{h^d_{=,=}+h^d_{=,\neq}+h^d_{\neq,=}}. \tag{4.5.9}$$

Here $h^d_{\times,\times'}$, $\times,\times' \in \{=,\neq\}$, $d \in \mathbb{N}_{>0}$, count the number of edge pairs which are in the same or different sets in the two partitions and which are separated by at most d nodes in G (cf. Appendix 6.4.6 for details). Thus, RI^1 and JI^1 yield structure-aware measures of partition similarity that consider only the partition memberships of adjacent edges (local perspective), while $\mathrm{RI}^\infty \equiv \mathrm{RI}$ and $\mathrm{JI}^\infty \equiv \mathrm{JI}$ do no not take into account the positions of edges in the graph and reproduce the original measures (global perspective; cf. Appendix 6.4.6 for an extensive comparison of similarity measures and intermediates between local and global perspective).

Extraction of weighted networks from image data

We test our method to disentangle filamentous networks on various weighted, geometric networks extracted from image data. The network extraction procedure is similar to those proposed in [Baumgarten and Hauser, 2012; Obara et al., 2012b] (cf. Appendix 6.4.5 for details). We analyze (1) two artificial networks extracted from drawn filamentous patterns, (2) two cytoskeletal networks from confocal microscope images of *Arabidopsis thaliana* hypocotyl actin cytoskeletons [Breuer et al., 2014], (3) 100 additional cytoskeletal networks from a movie over 200 s from the same experimental setup, (4) two neural networks from a fluorescence microscopy image of a branching rat hippocampal neuron *in vitro* [Brandner and Withers, 2014] and a schematic of a cat retinal ganglion cell [Masland, 2001], respectively, and (5) two cosmic networks obtained from images of simulated galaxy clusters [Stoica et al., 2005] (see Tab. 4.5.2 for an overview).

4.5.4 Results

Decomposing filamentous networks is a hard optimization problem

A filamentous network is naturally represented as a weighted graph, whereby the links (i.e., edges) denote segments of filaments and the nodes represent the ends of the segments. The edge weights typically capture the intensity or thickness of the filament segments. In this network representation, a filament corresponds to a path given by an ordered sequence of adjacent edges. To identify individual filaments, we seek a decomposition of the set of edges into paths so that each edge is covered (i.e., belongs to at least one path). Edges belonging to more than one path naturally model filament overlaps. We will refer to such a decomposition as a filament cover. Since a filament cover is non-unique, we introduce a quality measure, called roughness, to assess the quality of each path and the cover itself. Here we mainly consider the pairwise filament roughness given by the average absolute value of weight differences between adjacent edges. This roughness measure quantifies how strongly the thickness varies along a filament and is typically small for biological filaments. Disentangling the filamentous network amounts to solving the filament cover problem (FCP): Find a set of paths of minimum sum of roughness values that covers the network (cf. Methods and Appendix 6.4.1 for

the mathematical formulation). The FCP formulation is quite versatile: For instance, instead of minimizing the total roughness of the filament cover, we may minimize the average roughness. This optimization objective favors shorter filaments and may be more appropriate for specific types of networks. Other roughness measures (e.g., considering the spatial alignment of edges to penalize filaments with strong curvature) are readily introduced and can be considered in a multi-objective optimization approach (cf. Methods and Appendix 6.4.1 for different measures).

While providing a well-defined approach towards disentangling filamentous networks, solving the FCP is computationally prohibitive. Indeed, we show that the FCP is intractable even on planar graphs (cf. Methods and Appendix 6.4.2) which are used to represent filamentous structures extracted from 2D image data [Baumgarten and Hauser, 2012; Obara et al., 2012b]. While the FCP is solvable in polynomial time on trees (cf. Appendix 6.4.3), most biological filamentous structures are not tree-like as they contain loops [Katifori and Magnasco, 2012; Obara et al., 2012b; Breuer et al., 2014]. Therefore, we propose suitable approximation schemes to the FCP for the considered networks (cf. Methods and Appendix 6.4.4 for details and the mathematical formulation). The approximation schemes rely on collecting a large sample of paths in a given graph, followed by the computation of the roughness of each path. The paths are collected by performing a modified breadth-first search (BFS) or by sampling from random minimum spanning trees (RMST). Next, we write the FCP as classical set cover problem [Karp, 1972] which aims at covering the set of edges with a subset of the collected paths of minimum total or average roughness. The set cover approximation of FCP can be formulated and solved as a (fractional) binary linear program for which well-established algorithms exist [Schrijver, 1998]. The output of the program is a set of paths which correspond to the individual filaments of the studied network. Summarizing, the FCP may be solved with different options: The initial set of paths is obtained from a modified BFS (denoted by *BFS*) or sampling of RMSTs (*RMST*), the filaments may overlap (*over*) or not (*exact*), a pairwise (*pair*) or all-to-all filament roughness measure (*all*) is used, and the total (*total*) or average (*avg*) roughness is minimized. Since all these options are categorical, all possible $2^4 = 16$ combinations may be readily checked and no data-specific and computationally demanding gauging of continuous parameters is necessary, as is the case for related approaches [Leandro et al., 2009; Qiu and Li, 2014]. We provide an open-source implementation of our approach, termed "DeFiNe" (**De**composing **Fi**lamentous **Ne**tworks), with a simple and user-friendly graphical user interface available at http://mathbiol.mpimp-golm.mpg.de/DeFiNe/. DeFiNe takes as input a weighted graph in the standard .gml file format [Himsolt, 1997] and outputs a .gml graph with filament identities stored as edge colors as well as a standard, human-readable .csv-table of various individual filament measures for custom analyses.

Disentangling artificial filamentous structures

To test the accuracy of our approach, we investigate an artificial network (Fig. 4.5.1A) of pre-specified filamentous structure (Fig. 4.5.1B; cf. Methods and Appendix 6.4.4 for the extraction of the network; cf. Appendix 6.4.9 for an overview of the different stages of our approach, from an images to a network to filaments). The network contains crossing and overlapping filaments as well as a loop (Fig. 4.5.1B, ⊗, ⊜, and ⊙, respectively). First, we automatically decompose the weighted filamentous network by solving the FCP for a set of input paths from a modified BFS, allowing for overlaps, using the pairwise roughness measure,

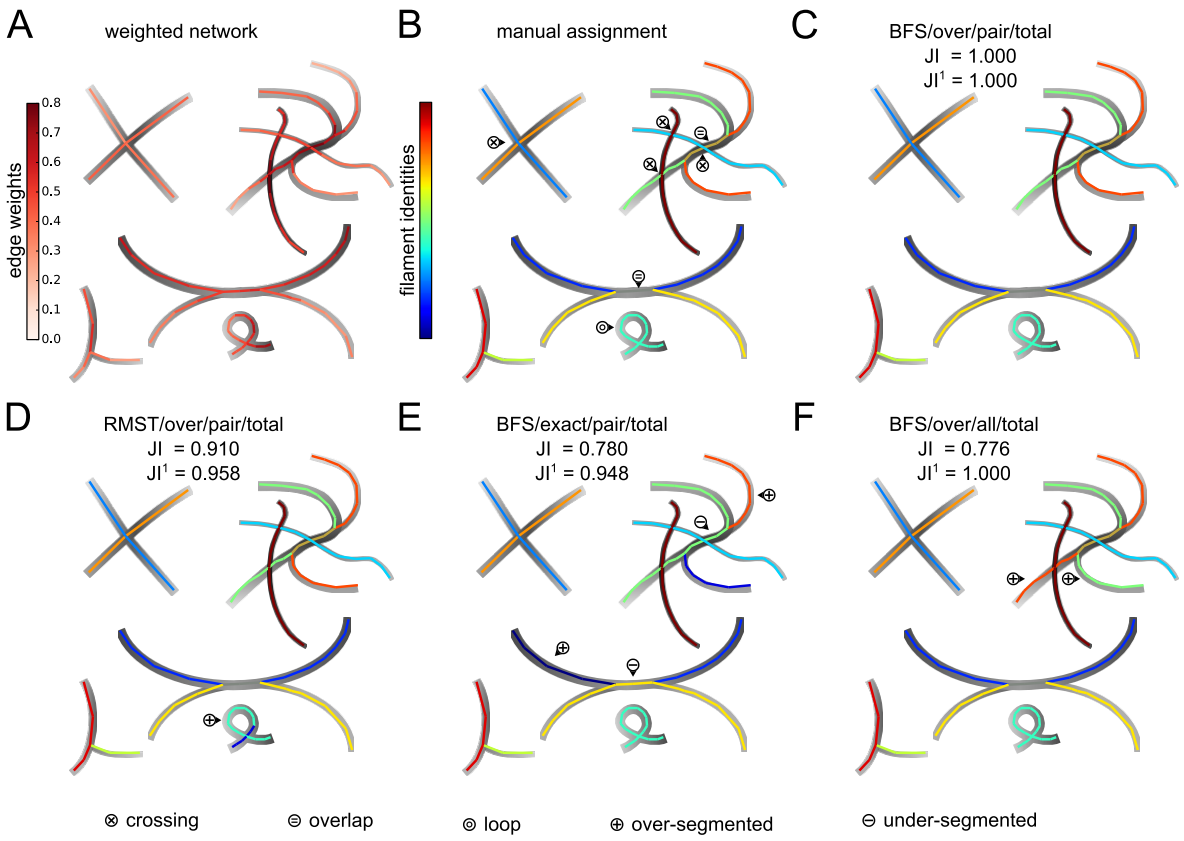


Figure 4.5.1: **Filament covers of artificial network with crossings, overlaps, and a loop.** **(A)** Weighted, artificial network extracted from the underlying drawing, with color-coded edge weights representing the local image intensity. **(B)** Manual decomposition of the network into filaments with color-coded indices. The filaments display crossings (⊗), overlaps (⊜), and a loop (◎). **(C)** Filament cover obtained by solving the FCP using the set of input paths generated by a modified breadth-first-search (*BFS*), allowing overlapping filaments (*over*), employing the pairwise roughness measure (*pair*), and by minimizing the total roughness of the cover (*total*). The automatically obtained filament cover correctly captures crossings, overlaps, and loops, and agrees excellently with the manual assignment (similarity of the two filament covers is measured by the global Jaccard index, JI, and our modified, structure-aware Jaccard index, JI^1, which reflect the fraction of pairs of all or only adjacent edges that are assigned to the same filament, respectively; here $JI = JI^1 = 1$). The filament identities and colors are matched by solving an assignment problem whereby the total number of edges shared by two filaments, from the manual and automated partitioning, is maximized; the same assignment procedure is used for the remaining panels. **(D)** When using paths obtained from sampling random minimum spanning trees (*RMST*) for the FCP, the closed filament loop is not correctly detected and is over-segmented (⊕). **(E)** When solving the exact FCP (*exact*), the loop is correctly detected. However, overlaps are neglected so that no two filaments share an edge, leading to over- and under-segmentation (⊖). **(F)** When minimizing the all-to-all filament roughness (*all*), two half-filaments are interchanged because the maximum weight difference is smaller along the altered filaments.

and minimizing the total roughness of the cover (Fig. 4.5.1C, cf. Eq. 4.5.4). The filament identities and colors are matched by solving an assignment problem (cf. [Kuhn, 1955; Wolsey and Nemhauser, 1999]) such that the total number of edges shared by two filaments, from the manual assignment and the automated cover, is maximized. The agreement between the automated cover and the manual assignment may be measured by classical partition similarity measures such as the Jaccard index JI which counts the fraction of edge pairs which are part of same filament [Meilă, 2005; Denœud and Guénoche, 2006]. However, JI does not take into account the structure of the underlying network. Hence, we introduced a new similarity measure, JI^1, that considers only pairs of adjacent edges in each filament and thus incorporates the network structure (cf. Methods and and Appendix 6.4.6 for details, a generalization to JI^d that considers only pairs of edges which are separated by at most d nodes, and a comparison of various similarity measures). For our artificial network, solving the above FCP yields a decomposition which agrees excellently with the manual assignment ($JI = JI^1 = 1$) as all filaments are correctly detected. Second, we choose a different set of input paths obtained from sampling RMSTs for solving th FCP (Fig. 4.5.1D). While most filaments are correctly detected, the loop (cf. Fig. 4.5.1B) is over-segmented (⊕) because it is not contained in the set of input paths in its entirety (due to looplessness of trees). Third, we solve the exact FCP which does not allow overlapping filaments (Fig. 4.5.1E). Expectedly, the agreement with the manual assignments is lower because filaments are over-segmented into disjoint segments and the supposedly overlapping parts are under-segmented (⊖), i.e., the respective edges are assigned to a single filament instead of multiple filaments. Finally, we employ the all-to-all roughness measure to assess the quality of the filaments (Fig. 4.5.1F, cf. Eq. 4.5.2). Filament crossings, overlaps, and the loop are again correctly detected but parts of two filaments are interchanged (cf. ⊕). This is due to the intensity/thickness of the underlying filaments which is consistently higher for the new detected filaments which are therefore favored by the all-to-all roughness measure. These test cases demonstrate the versatility and the accuracy of the proposed approach to decompose a given network into filaments.

In the analysis of many real-world filamentous structures, the knowledge of the underlying network structure is incomplete or the image data impede filament detection due to low signal-to-noise ratios. To investigate the effect of these obstacles on robust filament detection, we study two scenarios (Appendix 6.4.7): In the first scenario, we remove a single edge from the network, recompute the optimal filament cover, and calculate its agreement with the manual filament assignment as measured by the structure-aware Jaccard index JI^1. We repeat the procedure for all E edges and then proceed with the removal of E randomly chosen doubles of edges, triplets, up to subsets of 50 edges. As expected, the accuracy of the filament cover typically decreases with the number of removed edges, although removal of some specific edges even leads to an increase in accuracy. However, JI^1 decreases very moderately by less than 0.001 per removed edge on average (cf. Appendix 6.4.7). In the second scenario, we assess the robustness of our filament detection approach against noise by adding centered Gaussian noise of increasing standard deviation to the edge weights of the original network. For a given standard deviation, we obtain the optimal filament covers for 100 noisy network instances and compute their similarity, JI^1, to the manual assignment. Again, as expected, the accuracy of the filament cover decreases with increasing noise, but only slightly. On average, increasing the noise by 1% of the original edge weights only decreases JI^1 by less than 0.001. Moreover, we note that with increasing edge noise the accuracy of the filament cover approaches a constant, non-zero JI^1 which reflects that some information about the filament structure maybe obtained from the topology of the network alone,

irrespective of the edge weights (cf. Appendix 6.4.7).

Disentangling biological and cosmic filamentous structures

Since we demonstrated the power of the FCP-based approach on contrived filamentous structures, we next proceed with investigating real biological and cosmic filamentous structures (cf. Methods and Appendix 6.4.5 for the extraction of the networks; cf. Appendix 6.4.9 for an overview of the different stages of our approach). As a first illustrative example of a biological filamentous structure, we extract a weighted network from an image of a hippocampal neuron (Fig. 4.5.2A) and manually obtain a filament assignment with several crossings and loops (Fig. 4.5.2B, ⊗ and ◎, respectively). Solving the FCP (same options as in Fig. 4.5.1E) yields an automated decomposition which captures well the manual assignment, in particular the two loops (Fig. 4.5.2C, $\mathrm{JI}^1 = 0.937$). This is further supported by the distributions of filament lengths (measured by the numbers of edges) as well as the distributions of maximal filament angles (measured between adjacent edges), which are statistically indistinguishable between the manual assignment and the automated decomposition (Fig. 4.5.2D, black and red; Kolmogorov-Smirnov test p-value $p_{\mathrm{KS}} \geq 0.05$). A detailed analysis of the similarity of manual and automated decompositions shows that the classical Rand index RI [Hubert and Arabie, 1985] overestimates the similarity, while the variation of information VI [Meilă, 2003] and the Jaccard index JI severely underestimate the similarity between the manual and automated decomposition when compared to the values of the here-proposed RI^1 and JI^1 (Fig. 4.5.2E, dotted blue, green, and yellow). The latter measures take into consideration the network structure when comparing two network decompositions (Fig. 4.5.2E, solid blue and yellow). We would like to emphasize that the disparities in the estimations of RI and JI result from the consideration of distant, non-adjacent edges which are excluded in RI^1 and JI^1. In addition, we observe that RI^d and JI^d show a non-trivial dependence on the distance, d, between the considered edges, and coincide with the classical similarity measure for large enough distances, i.e., $\mathrm{RI}^\infty \equiv \mathrm{RI}$ and $\mathrm{JI}^\infty \equiv \mathrm{JI}$ (cf. Appendix 6.4.6 for a detailed discussion).

Finally, different flavors of the FCP may be solved , as mentioned above, to obtain decompositions of varying similarity in comparison to the manual assignment (Fig. 4.5.2F). Solving the FCP with paths from the modified BFS, instead of RMSTs, yields consistently higher RI^1- and JI^1-values for the agreement with the manual assignment. This is due to the higher flexibility with respect to the treatment of loops. For the studied networks, a decomposition based on the minimization of the total roughness yields higher RI^1- and JI^1-values in comparison to the minimization of the average roughness. In addition, in terms of RI^1 and JI^1, covers allowing for overlaps yield better agreement with the manual assignment, in comparison to those in which each edge is covered by a single path. However, these expected trends are absent or even reversed for the classical similarity measures VI, RI, and JI (cf. Appendix 6.4.6), which further justifies the usage of the here-proposed RI^1 and JI^1 for comparing decompositions of networks arising in other network-based analyses (cf. e.g. [Newman, 2012]).

As a second biological example, we investigate the filamentous structure of a plant actin cytoskeleton (Fig. 4.5.3A). We create seven manual assignments (one of which is shown in Fig. 4.5.3B) for a quantitative comparison with the automated decomposition (Fig. 4.5.3C, $\mathrm{JI}^1 = 0.655$; same options of the FCP as in Fig. 4.5.1E). The agreement of the automated decomposition with the manual assignment is good, despite several over- or under-segmented filaments (Fig. 4.5.3C, cf. ⊕ and ⊖). For a comprehensive assessment of

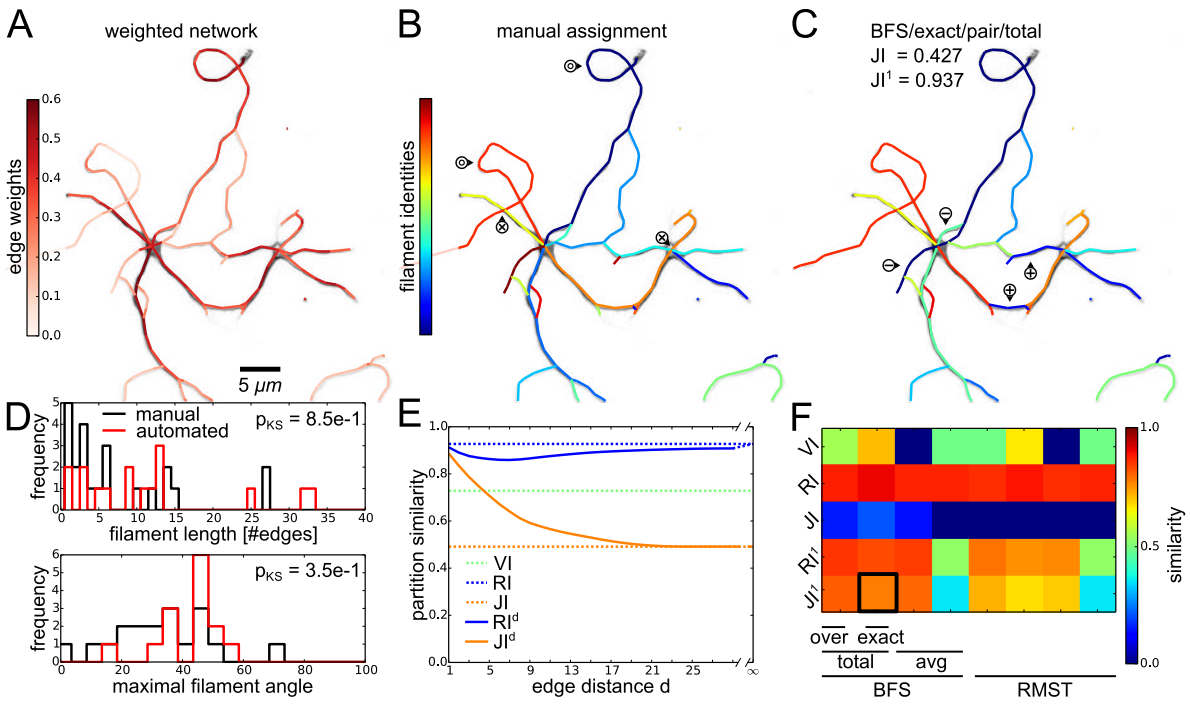


Figure 4.5.2: **Filament covers and analyses of neuronal network.** The weighted hippocampal neuronal network is automatically decomposed into filaments by solving the exact FCP (*exact*) for paths from a modified breadth-first search (*BFS*) and by minimizing the total (*total*) pairwise filament roughness (*pair*). **(A)** Overlay of fluorescence microscopy image of hippocampal neurons and extracted network with color-coded edge weights. **(B)** Manual decomposition of the neuronal network into filaments with color-coded indices and crossings (⊗) and loops (◎). **(C)** Automated partitioning of the network obtained by solving the FCP displays good agreement with the manually obtained partitioning (JI^1 close to 1, see panel (E) for details) with marked illustrative sites of over- (⊕) and under-segmentation (⊖). **(D)** Distributions of numbers of edges per filament (upper panel) as well as distributions of maximum filament angles (lower panel) are similar for manual (black) and automated decomposition (red; Kolmogorov-Smirnov test $p_{KS} \geq 0.05$). **(E)** Different measures of similarity of manual and automated decompositions. The variation of information VI (dashed green) indicates moderate similarity but is not well-defined for general, overlapping decompositions. While the classical Jaccard index JI (dashed yellow) is of small value, the proposed structure-aware extension JI^d increases with decreasing d, i.e., when only edges are considered that are separated by at most d nodes (solid yellow). Moreover, while the classical Rand index RI (dashed blue) is of large value, the proposed structure-aware extension RI^d displays a non-monotonic dependence on d (solid blue). **(F)** Heat map of partition similarities for different similarity measures and options of the FCP, cf. Fig. 4.5.1 for a demonstration of the different options. The FCP options which yield the partition shown in (C) are marked by a black rectangle.

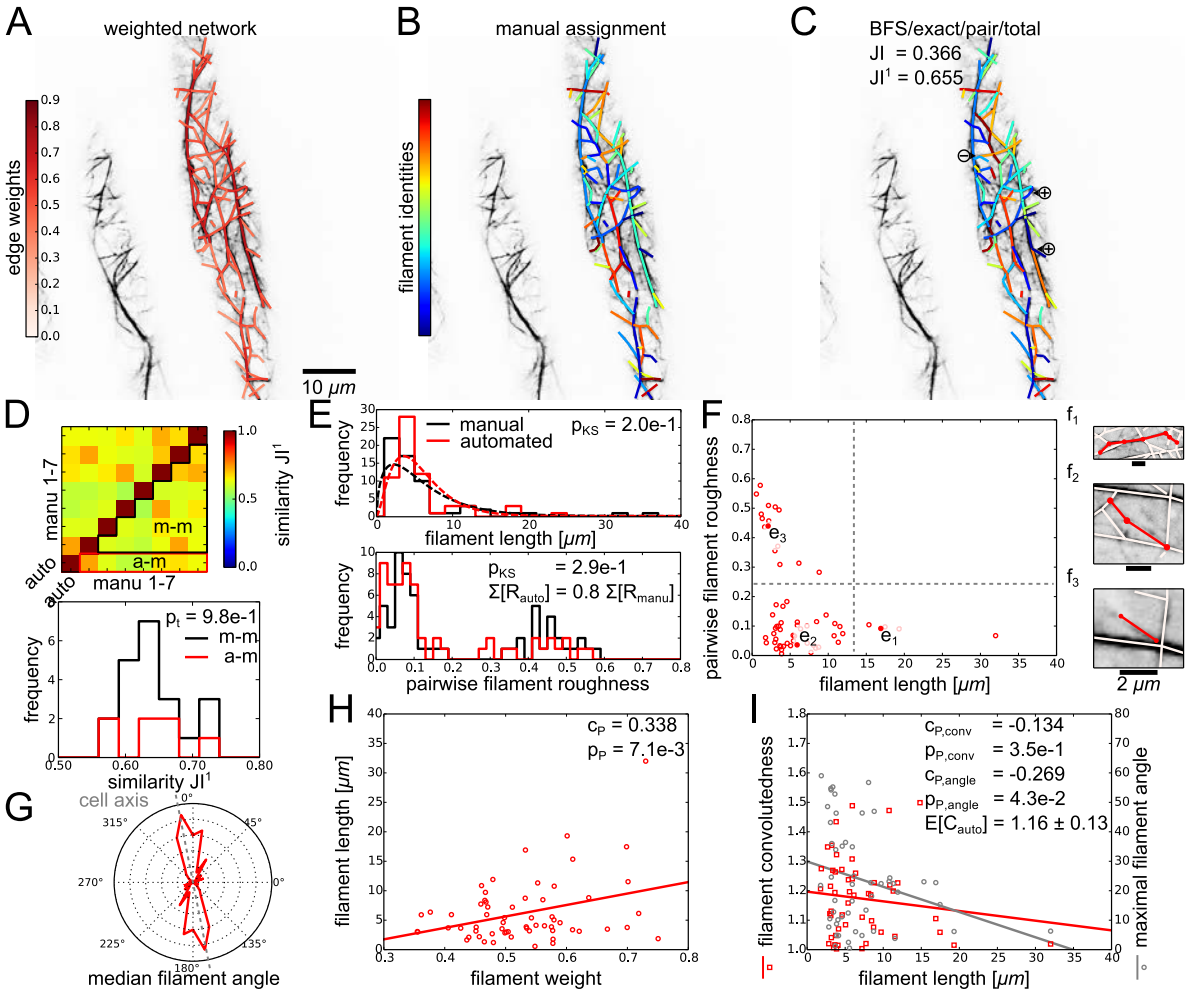


Figure 4.5.3: **Filament covers and analyses of cytoskeletal network.** The weighted cytoskeletal network is decomposed automatically by solving the exact FCP (*exact*) for paths from a modified breadth-first search (*BFS*) and by minimizing the total (*total*) pairwise filament roughness (*pair*). **(A)** Overlay of confocal microscopy image of an actin cytoskeleton and extracted network with color-coded edge weights. **(B)** Manual decomposition of the actin cytoskeleton into filaments with color-coded indices. **(C)** The automated decomposition according to the FCP correctly assigns many of the filaments ($JI^1 = 0.655$). Some occurrences of over- (⊕) and under-segmentation (⊖) are marked. **(D)** Heat map of similarity between automated (cf. (C)) and seven manual decompositions (cf. e.g. (B); upper panel). The similarities between automated and manual decompositions (red, denoted by a-m) do not differ from similarities among the different manual decompositions (black, m-m; lower panel; cf. independent two-sample Student's t-test p-value $p_t \geq 0.05$). **(E)** Distribution of filament lengths for the manual (black) and automated solution (red) are similar (upper panel; cf. Kolmogorov-Smirnov test p-value $p_{KS} \geq 0.05$). Maximum likelihood fits of gamma functions are shown as dashed lines. The distributions of pairwise filament roughnesses are similar (lower panel; cf. $p_{KS} \geq 0.05$), while the total roughness is smaller (cf. summed R-values) for the automated decomposition since it is minimized by the FCP. **(F)** Caption continued on next page.

(F) Scatter plot of pairwise filament roughness versus filament length displays three regions, with representative examples $f_1 - f_3$ (solid dots): (f_1) For long filaments ($\geq 15\,\mu$m), the roughness is moderate (< 0.2), as expected for actin bundles; (f_2) The majority of filaments is short ($< 15\,\mu$m) and of moderate roughness; (f_3) Some typically short filaments show a high roughness (≥ 0.2), namely those which are composed of one network edge only so that their roughness is given by the edge weight itself (cf. Eq. 4.5.1). **(G)** The distribution of median filament angles shows that the majority of filaments is aligned parallel to the cell axis (gray dashed line). **(H)** The filament length correlates with the filament weight (cf. linear regression and Pearson correlation coefficient $c_P > 0$ and p-value $p_P < 0.05$) **(I)** Scatter plot of filament convolutedness versus filament length shows a negative but non-significant correlation (cf. red squares, $c_{P,\text{conv}} < 0$, and $p_{P,\text{conv}} \geq 0.05$) with an average convolutedness of $E[C] = 1.16 \pm 0.13$. The maximum filament angle correlates negatively and significantly with the filament length (cf. gray circles, $c_{P,\text{angle}} < 0$, and $p_{P,\text{angle}} < 0.05$), indicating that longer (and thicker, cf. (G)) filaments are less curved.

this agreement, we compute the pairwise similarities between the automated and all seven manual filament decompositions (Fig. 4.5.3D, upper panel). By comparing the similarities between automated and manual decompositions to the similarities among the different manual decompositions, we find reassuringly that our automated solution is as good as any manual decomposition (Fig. 4.5.3D, lower panel, red and black, respectively; cf. independent two-sample Student's t-test p-value $p_t \geq 0.05$). The agreement between the automated decomposition and the reference manual assignment (cf. Fig. 4.5.3B) is further confirmed by statistical tests which demonstrate that the two distributions of filament lengths from manual assignment and automated decomposition do not statistically differ (Fig. 4.5.3E, upper panel, black and red histograms; cf. $p_{KS} \geq 0.05$). In addition, our results indicate that the filament lengths may be described by a gamma distribution (Fig. 4.5.3E, upper panel, dashed lines; maximal likelihood fits of normal, Weibull, and Rayleigh distributions yield higher values for the Akaike information criterion [Akaike, 1974]), in agreement with theoretical and experimental studies [Burlacu et al., 1992; Ermentrout and Edelstein-Keshet, 1998]. Moreover, the distributions of average pairwise filament roughnesses do not differ between manual assignment and automated decomposition (Fig. 4.5.3E, lower panel; cf. $p_{KS} \geq 0.05$). We note that the sum of filament roughnesses, R, is larger for the manual assignment of filaments than in the automated decomposition, as expected, as R is the objective function of the minimization in the FCP-based formulation.

By investigating the relationship between filament length and pairwise roughness, we can distinguish three regions (Fig. 4.5.3F): Long filaments typically correspond to actin bundles and exhibit small roughnesses (Fig. 4.5.3f_1), the majority of filaments is shorter with comparable roughnesses (Fig. 4.5.3f_2), and some typically short filaments consist of only one edge with roughness given by the edge weight itself (Fig. 4.5.3f_3; cf. Eq. 4.5.1). The angular distribution of filaments indicates that the majority of filaments is aligned parallel to the cell axis (Fig. 4.5.3F, dashed gray line) which has been suggested to support longitudinal cell growth [Waller and Nick, 1997; Sampathkumar et al., 2011]. While these reports of longitudinal alignment of the actin cytoskeleton were based on manual or qualitative measurements, our approach facilitates fully automated quantification of the alignment of individual filaments. Our findings show that the length of a filament correlates with its average weight (Fig. 4.5.3G; Pearson correlation coefficient $c_P > 0$ and p-value $p_P < 0.05$), i.e., thicker actin bundles stretch across the cell while individual thinner actin filaments are more locally confined, as expected [Staiger et al., 2009; Akkerman et al., 2011].

Finally, we study filament convolutedness, given by the ratio of the length of a filament and the largest side of a bounding box enclosing the filament, used as a measure for the curvedness of a filament [Staiger et al.,

2009]. We find that the convolutedness is slightly negatively correlated with the filament length (Fig. 4.5.3I, red; $c_{\mathrm{P,conv}} < 0$ and $p_{\mathrm{P,conv}} \geq 0.05$), in agreement with previous findings in *Arabidopsis thaliana* pollen grain [Staiger et al., 2009] and other plant species [Henty-Ridilla et al., 2013]. In contrast to the automated approach used here, the existing studies of filament convolutedness required manual segmentation which may be biased by the user. Generally, and more severely, using a bounding rectangle to compute the convolutedness of a filament is biased by the orientation of the filament with respect to the x- and y-axis of the image. Therefore, we use the maximal filament angle as a non-biased measure for the maximal, local curvedness of a filament. By investigating the relation between the maximal filament angle and filament length, we find a significant negative correlation (Fig. 4.5.3I, gray; $c_{\mathrm{P,angle}} < 0$ and $p_{\mathrm{P,angle}} < 0.05$). This negative correlation reflects the known increase in stiffness of actin bundles with increasing bundledness and length [Gardel et al., 2004; Claessens et al., 2006]. Thus, our approach provides a fast means to investigate this property for individual filaments in a cellular context without laborious manual filament identification.

To further extend these findings, we extract the cytoskeletal networks from 100 frames of a movie of a plant actin cytoskeleton (cf. Methods). For each frame, we compute the optimal filament covers and analyze the filaments. The additional data support our reported findings (Appendix 6.4.8).

Moreover, we repeat our analyses of the robustness of our approach against incomplete knowledge of the underlying network structure or noisy edge weights for the cytoskeletal network (cf. discussion of Fig. 4.5.1; Appendix 6.4.7). In our first scenario, the removal of increasing numbers of edges typically moderately decreases the accuracy of the obtained filament covers, i.e., their agreement with the manual assignment as measured by JI^1. While the removal of some critical edges leads to a more severe decrease in accuracy, there exist edges whose removal leads to an increase in accuracy. On average, the removal of one additional edge decreases JI^1 by around 0.002. Consequently, a loss of 10% of the cytoskeletal network's $E = 179$ edges still yields $\mathrm{JI}^1 \approx 0.6$ which is comparable to similarity values between different manual assignments (cf. Fig. 4.5.3D; cf. Appendix 6.4.7). In our second scenario, the adding of Gaussian noise of increasing standard deviation to the edge weights similarly, as expected, decreases the accuracy of the obtained filament covers. However, this effect is moderate, i.e, increasing the standard deviation by 1% of the original edge weight decreases JI^1 by less than 0.001. Adding noise with a standard deviation of 20% of the original edges weights still yields $\mathrm{JI}^1 \approx 0.6$. As for the robustness analyses of the contrived network, for strong noise, JI^1 tends to a constant, non-zero value which suggests that some information about the filament structure may be obtained solely from the network topology, irrespective of the edge weights (cf. discussion of Fig. 4.5.1; cf. Appendix 6.4.7).

As a final example, we decompose the network of a simulated galaxy cluster (Fig. 4.5.4A) into individual galaxy filaments (Fig. 4.5.4B). The quantification of galaxy filaments may help to elucidate the acceleration of the universe [Sousbie et al., 2008b] and improve our understanding of large-scale structure formation [Sousbie et al., 2008a]. Moreover, studies have revealed gravitational motion of galaxies along individual filaments [Faltenbacher et al., 2002; Aubert et al., 2004]. Yet, previous studies focused on connected components of the cosmic web, and sought robust methods to identify individual filaments [Stoica et al., 2005; Sousbie et al., 2008a]. Our approach confirms the expected discrepancy between the lengths of the components (i.e., the sum of their edge lengths; Fig. 4.5.4C, upper panel, gray) and the length of individual filaments (Fig. 4.5.4C, upper panel, red; cf. average L-values). Moreover, the decomposition of the cosmic structures enables analyses of individual filament shapes. For example, the convolutedness which measures the curvedness of a

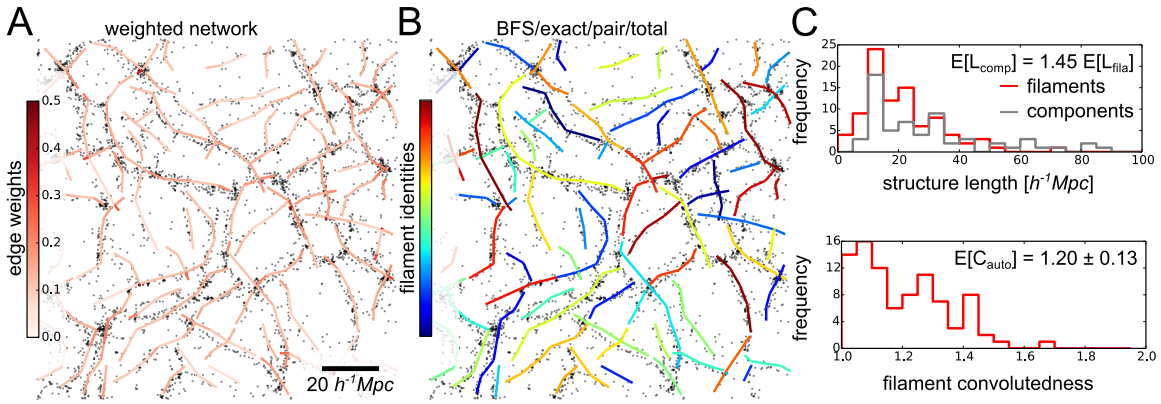


Figure 4.5.4: **Filament covers and analyses of cosmic web.** Image data from: Stoica et al., A&A, 434, 423-432, 2005, reproduced with permission © ESO [Stoica et al., 2005]. The cosmic web is decomposed automatically by solving the exact FCP (*exact*) for paths from a modified breadth-first search (*BFS*) and by minimizing the total (*total*) pairwise filament roughness (*pair*). Distances are given in h^{-1}Mpc, where currently $h \approx 0.7$ is the dimensionless Hubble parameter [Croton, 2013]. **(A)** Overlay of simulated galaxy clusters and extracted network with color-coded edge weights. **(B)** Automated decomposition of the cosmic web into galaxy filaments with color-coded indices. **(C)** The length distribution of galaxy filaments exhibits a peak around $20\,h^{-1}$Mpc and levels off for larger lengths (upper panel, red). As a comparison, the distribution of the total lengths of the connected components levels off more slowly and overestimates the average filament length by a factor of 1.45 (upper panel, gray; cf. average L-values). The distribution of the convolutedness of galaxy filaments suggests a prevalence of straight filaments and its average is comparable to that of the actin network (cf. 4.5.3i; cf. $\mathrm{E}[C] = 1.20 \pm 0.13$).

filament shows small values (Fig. 4.5.4, lower panel), which are interestingly comparable to those found in the actin cytoskeleton (cf. Fig. 4.5.3I; cf. average C-values), indicating the prevalence of straight galaxy filaments.

In Tab. 4.5.2, we summarize the quality of the investigated decompositions of different filamentous networks and the options of the underlying FCP (cf. Appendix 6.4.8 and 6.4.9 for analyses of additional filamentous networks that are not shown in the main text).

4.5.5 Discussion

The decomposition of complex networks into meaningful substructures has facilitated network-based analyses of systems found in nature or designed by humans [Milo et al., 2002; Shen-Orr et al., 2002; Sporns and Kötter, 2004]. These natural and technical networks often embed filaments as basic building units. To enable deeper understanding of network systems with filamentous structure, it is therefore paramount to develop methods for accurate and feasible identification of the underlying filaments. In particular, the distinction between intra- and inter-filament connections enables a more detailed analysis of filamentous structures, including length statistics, spatial alignment, and bending of individual filaments. Such statistics may offer new insights, e.g., into the role of single actin or galaxy filaments in their cellular or cosmic network context, respectively (cf. Figs. 4.5.3E-I and 4.5.4c).

Here, we proposed a robust optimization approach to decompose any given weighted network into a set

		Figure	**Options**	**Similarity**				
				VI	RI $(\equiv \mathrm{RI}^{\infty})$	JI $(\equiv \mathrm{JI}^{\infty})$	RI^{1}	JI^{1}
artificial	overlaps + loop	4.5.1	*BFS/over/pair/tot*	0.792	1.000	1.000	1.000	1.000
	grid-like	6.4.9B	*BFS/exact/pair/tot*	0.889	0.962	0.742	0.941	0.872
neural	hippocampus	4.5.2	*BFS/exact/pair/tot*	0.848	0.906	0.427	0.954	0.937
	retina	6.4.9D	*BFS/exact/pair/tot*	0.792	0.963	0.397	0.905	0.883
cytoskeletal	actin (FABD-labeled)	4.5.3	*BFS/exact/pair/tot*	0.829	0.976	0.366	0.854	0.655
	actin (Lifeact-labeled)	6.4.9F	*BFS/exact/pair/tot*	0.530	0.929	0.193	0.838	0.701
cosmic	galaxy cluster (sparse)	4.5.4	*BFS/exact/pair/tot*	no manual assignment				
	galaxy cluster (dense)	6.4.9H	*BFS/exact/pair/tot*	for comparison				

Table 4.5.2: **Quality of filament covers of artificial, biological, and cosmic networks in comparison to manual decompositions.** A given network is decomposed into filaments by solving the FCP with different options: The initial set of paths is obtained from a modified breadth-first search (*BFS*) or sampling of random minimum spanning trees (*RMST*), the filaments may overlap (*over*) or not (*exact*), a pairwise (*pair*) or all-to-all filament roughness measure (*all*) is used, and the total (*total*) or average (*avg*) roughness is minimized. The table displays the investigated filament covers with high similarity to the manual assignments.

of smooth filaments comprising a filament cover. Since we demonstrated that the filament cover problem is intractable on general networks, we proposed, tested, and validated several alternative approximation schemes. The proposed approximation schemes are gauged at applications from different scientific fields in which filamentous structures naturally arise. We applied our optimization-based approach on contrived test cases as well as biological and cosmic networks, and showed that it reliably identifies crossing, (non-) overlapping, and looped filaments in agreement with expert-based manual assignments.

Our approach offers a number of advantages over the existing alternatives: (1) The proposed optimization approach can be applied to any weighted network. In particular, the approach can be readily applied to any network generated from two- or three-dimensional experimental image data typically gathered in biological studies and analyses of man-made systems (e.g. [Masland, 2001; Paredez et al., 2006; Riedl et al., 2008; Tero et al., 2010]), irrespective of the image source (e.g., light microscopy- or MRI-based). Thus, it may be used to study a variety of natural and technical filamentous structures in search for universal properties which go beyond the characterization of geometric networks [Barthélemy, 2011].

(2) Our approach facilitates the establishment of a link between the dynamics of individual filaments and the dynamics of the whole network. While the dynamics of individual filaments is guided by typically molecular, local processes, the behavior of the entire filamentous structure incorporates and responds to stimuli across different scales. Therefore, the proposed approach provides the starting point towards network-oriented analysis of filaments. More specifically, the filament covers may even be used to track mobile filaments, as has been proposed for images of a few filaments using open contours [Smith et al., 2010], providing a venue for fruitful applications of the method.

(3) The different options of our approach, e.g., different measures of the filament roughness, enable flexible and intuitive customization for different types of networks. For example, the filament roughness measure may include a penalty for filament bending in networks of straight, stiff filaments (such as microtubules [Gittes et al., 1993; van Mameren et al., 2009]), or a penalty for length deviations in networks of filaments of mostly uniform length (such as synthetic polymers that are used, e.g., in drug delivery systems [Ali and Brocchini, 2006; Hartmann and Börner, 2009]).

(4) At the same time, our approach to disentangle a given network is parsimonious, i.e., it has a strictly limited number of categorical options which allow testing of all possible combinations ($4^2 = 16$ in total). In contrast, approaches which rely on multiple continuous parameters require data-specific and computationally expensive gauging of the parameters [Leandro et al., 2009; Qiu and Li, 2014]. When compared to approaches which detect filaments directly from image data, however, the parsimony of our approach is counterbalanced by the parameter requirements of the preceding network extraction procedure.

(5) Nevertheless, approaches that detect filaments directly from image data typically rely on local optimization schemes and thus, e.g., on the order of filament initializations and definitions of local filament properties [Mayerich and Keyser, 2008; Meijering, 2010; Peng et al., 2015a; Xu et al., 2015]. In contrast, our approach offer the advantage that the decomposition into filaments is performed in a single optimization step which holistically considers the global structure of both filaments and network.

(6) Finally, since our approach replies on a general network representation, it may be applied also to networks obtained from other, e.g., open contour-based methods which often do not capture filament overlaps and result in fragmented filaments [Xu et al., 2014, 2015]. In a post-processing step, these fragments may be conveniently merged using our network-based approach (cf. Appendix 6.4.10).

Yet, some caution is warranted: (1) The available options of the FCP yield different decompositions. We showed that paths sampled from a modified BFS enable more flexible and more accurate decompositions in comparison to paths sampled from RMSTs (cf. Fig. 4.5.1); in contrast to minimizing the the average roughness, the minimization of the total roughness favors longer filaments in better accordance with the manual assignments (cf. Fig. 4.5.1); moreover, since filament overlaps in biological systems may lead to an abrupt increase in apparent filament thickness, the proposed all-to-all filament roughness may be more suitable to study such situations than the pairwise filament roughness which favors filaments of slowly varying thickness. Therefore, the suitable choice of feasible and suitable options has to be further investigated. For example, for the actin cytoskeletal networks, it is not obvious if overlapping filaments should be preferred over non-overlapping filaments and if the pairwise roughness is a better measure of filament quality than the all-to-all roughness. Yet, such decision problems are innate not only to all automated decomposition algorithms, but also to the manual assignment based on which the performance is assessed. Thus, exploring different decomposition options by an expert in the field may hint at the right choice.

(2) The quality of the filament cover clearly depends on the quality of the input network. To this end, several algorithms have been proposed for the extraction of various types of networks from image data with low error rates [Cohen et al., 1994; Baumgarten and Hauser, 2012; Meijering, 2010; Obara et al., 2012b; Qiu and Li, 2014; Xu et al., 2015]. Moreover, we investigated different scenarios to test the robustness of our approach against incomplete knowledge of the underlying network structure as well as low signal-to-noise ratios and found that the accuracy of the filament cover is only moderately affected by these obstacles (cf. Appendix 6.4.7).

(3) Another issue are the computational requirements of the FCP. Although our proposed approximation scheme employs a modified BFS and a binary linear program which run fast on the tested networks, it may become infeasible for larger networks comprising more edges or nodes of larger degrees. Therefore, future efforts may focus on devising algorithms which approximate the FCP by employing local searches, i.e., without sampling a large number of paths for the proposed set cover-based approximation scheme.

(4) Finally, we note that many polymers are not simple linear chains but branched tree-like structures [Inoue, 2000; Tomalia and Frechet, 2001]. Also many neurons may be naturally described as tree-like structures [Verwer and van Pelt, 1983; Ascoli et al., 2007]. Our approach can be extended to account for these cases, thus, opening a new field of research. To this end, covering networks with more complex structures, such as stars [Tarsi, 1981; Cohen and Tarsi, 1991; Lin and Shyu, 1996] or, more generally, trees [Even et al., 2004; Horak and McAvaney, 2008] may be employed. Due to intractability of these problems, investigation of approximation schemes like our set cover formulation will be needed. A central question will be the development of measures for the quality of a given star or tree cover.

In conclusion, by decomposing technically and biologically relevant filamentous structures into their constitutive filaments, our approach allows to see both the wood and the trees.

4.5.6 Acknowledgements

Funding: D.B. and Z.N. acknowledge financial support by the Max Planck Society.

Contributions: D.B. implemented the method and analyzed the data; D.B. and Z.N. developed the method, showed the computational complexity of the problem, and wrote the manuscript.

5 Conclusion

In this closing chapter, we briefly summarize our main findings (Section 5.1). Detailed discussions of our specific findings are given in the respective sections in the Results (cf. Chapter 4). Therefore, in the following, we focus on embedding our findings in a broader biological context and propose applications and extensions of our work (Sections 5.2-5.5).

5.1 Summary

Despite the biological and technical relevance of cellular transport by the actin cytoskeleton in plants, previous efforts have focused on one of two extremes (cf. Section 3.5): Either they elucidated molecular details of the underlying biological mechanisms, or they developed abstract system-level models with little or no connection to real-life cells.

Therefore, we set out to a develop and exploit a viable systems perspective that allows incorporation of biological data of cytoskeletal architecture and transport. To this end, we addressed and resolved five key issues stated in the Introduction (cf. Section 3.9): Namely, we proposed two image-based methods for representing experimentally observed cytoskeletal components as networks at a cellular level (cf. Sections 4.1 and 4.4). These representations enabled quantitative and unbiased phenotyping of the cytoskeleton beyond statistics for (a limited number of) manually tracked filaments (cf. Sections 4.1-4.5). Moreover, by comparing the extracted cytoskeletal networks to suitable null models, we showed that the cytoskeletal networks displayed properties that support efficient transport (cf. Sections 4.1 and 4.4). By studying transport of Golgi bodies in plant cells, we offered quantitative evidence that organelle transport is shaped not only by local actin filaments but also by the system-wide organization as well as the cylindrical geometry of the actin cytoskeleton as a whole (cf. Section 4.4). Finally, we devised and employed a method to identify individual actin filaments from a given network representation of the cytoskeleton, which allowed automated investigation of single filament properties in the native cellular context (cf. Section 4.5). Implementations of the developed methods were made available as open-source tools to researcher from different fields.

Altogether, our findings and our proposed frameworks provided quantitative, system-level insights into cellular transport in living cells.

5.2 Application of image-based network representations to super-resolution image data and in large-scale screens

The image-based network representations established in this thesis provide powerful means for quantification or phenotyping of cytoskeletal structures. Our first network framework (cf. Sections 4.1 and 4.2) employs a weighted grid-approximation that captures fewer details of the underlying structure, but is more robust to image noise and potential segmentation error than previous methods. In particular, while other approaches are typically tailored to a specific image source [Baumgarten and Hauser, 2012; Dhondt et al.,

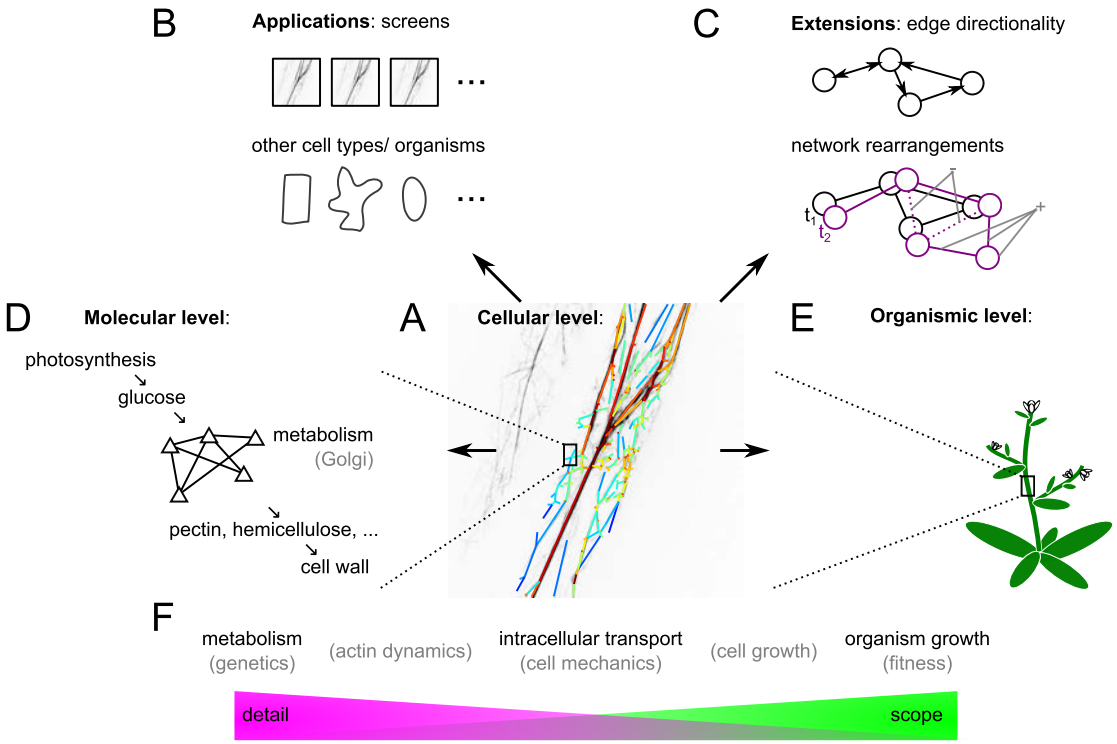


Figure 5.1.5: **Perspectives for future research, from direct applications and extensions of our work to multi-scale models.** **(A)** Starting point: Image-based, network-driven framework of the cytoskeleton as established in this thesis. **(B)** Applications: Chemical and genetic screens may be used to uncover the basis of complex traits of cytoskeletal organization as given by our network phenotypes. **(C)** Extensions: The established weighted but undirected network representation may be further augmented by including directed edges, based on image data of cargo transport. Beyond analyses of (static) cytoskeletal networks at a given point in time, their dynamic rearrangements must be considered (cf. two networks at time points t_1 and t_2) and may be modeled in a geometric network framework which takes into account addition (cf. +) and removal (cf. -) of nodes and edges. **(D)** Molecular level: Plant cell wall material, such as pectin and hemicellulose, are central metabolic products that are synthesized in the Golgi apparatus and deposited in the cell wall via cytoskeleton-based transport and exocytosis, linking the molecular and cytoskeletal level. **(E)** Organismic level: The cytoskeleton plays an important role also at an organismic level, e.g., due to its close functional relation to cell growth and division. **(F)** Trade-off between level of detail and scope of a model (cf. Fig. 3.1) and examples of processes at different levels.

2012; Obara et al., 2012b; Peng et al., 2015a], this framework is applicable to any gray-scale image data from any imaging system and network-like system of interest. Our second network framework (cf. Section 4.4; Fig. 5.1.5A) captures subtle cytoskeletal structures and consequently relies on image data with higher signal-to-noise ratios. Similar approaches for extraction of networks from image data have been proposed, but they required laborious manual user input [Longair et al., 2011; Dhondt et al., 2012] or did not capture intensity-based edge weights that are required as a measure for thickness of cytoskeletal filaments and bundles [Baumgarten and Hauser, 2012; Obara et al., 2012b; Xu et al., 2015]. We demonstrated that our frameworks capture known as well as relevant novel features of cytoskeletal organization in different cell types and under different conditions (cf. Sections 4.1 - 4.3).

While we focused on confocal microscopy data, recent progress in so-called super resolution microscopy has enabled imaging below the diffraction limit [Hell and Wichmann, 1994; Betzig et al., 2006; Li et al., 2015]. In particular, while previous super-resolution techniques suffered from low temporal resolution, current developments achieve increasingly good trade-offs between the spatio-temporal resolution and invasiveness due to high laser intensities, enabling imaging of fast cytoskeletal rearrangements *in vivo* [Li et al., 2015]. While these methods have been applied for imaging of cytoskeletal structures of animal cells [Bálint et al., 2013; Li et al., 2015], the image-based, network-driven framework established for plant cells in this thesis is readily applicable to such super resolution image data.

Moreover, by providing fast, fully automated, and unbiased quantification of cytoskeletal phenotypes, our framework is readily applicable in chemical or genetic screens (Fig. 5.1.5B). Previously, manual examinations of cytoskeletal structures were used in chemical screens to identify potential interaction partners that affect actin cytoskeletal organization and functionality [Baluška et al., 2001; Mathur, 2004; Kim et al., 2005]. Additionally, the increasing wealth of genetic data and bioinformatics approaches available, such as quantitative trait locus (QTL) mapping or genome-wide association studies (GWAS), may be combined with our phenotyping approaches to study the genetic basis of cytoskeletal organization as a complex trait [Steinmetz et al., 2002; McCarthy et al., 2008; Mitchell-Olds, 2010]. Such a combination of approaches may pinpoint specific genes that are associated with a quantitative cytoskeletal phenotype, e.g., network properties reflecting cytoskeletal organization or transport characteristics. Considering the large number of proteins and other gene products that interact with the actin cytoskeleton and modify its dynamics and structure (cf. Section 3.1), these studies may establish a connection between molecular processes underlying the cytoskeleton and its system-level organization that are required for establishing multi-scale models. Generally, the strength of automated imaging approaches such as ours is comparable to other high-throughput technologies, like sequencing approaches or metabolomics [Swedlow et al., 2009; Eliceiri et al., 2012; Schneider et al., 2012].

5.3 Transferability of our findings to system-level coordination of cytoskeletal transport in other cellular settings

Despite differences in the role of the cytoskeleton in long-ranged intracellular transport across cells types and organisms (cf. Sections 3.1 and 3.2; [Cooper, 2000; van den Ent et al., 2001; Liu, 2010]), our developed frameworks may be readily used to study cytoskeletal transport in other scenarios. While we focused on actin-based transport of Golgi in plant cells, we expect similar findings and system-level coordination of

transport for other plant organelles. For example, movement of mitochondria relies on the acto-myosin system [Akkerman et al., 2011; Wang and Hussey, 2015] and there is evidence that mitochondria movement is affected by local actin bundles [Avisar et al., 2008b; Akkerman et al., 2011]. Therefore, we expect that by using a systems-level approach, such as the one described here, the effects of cellular coordination that go beyond local correlations can be revealed and further investigated. In contrast, photodamage avoidance movement of chloroplasts is a more directed type of movement that supposedly requires less coordination [Kasahara et al., 2002]. While the actin cytoskeleton transports many different types of cargo, transport along microtubules has been shown only for few types of cargo in plant cells, such as small cellulose synthase containing vesicles (SmaCCs) [Gutierrez et al., 2009; Wightman and Turner, 2010; Bringmann et al., 2012]. This is in accordance with the hypothesis that microtubules predominantly shape the mechanical properties of the cell and guide cell wall material to its depositioning sites, where it is delivered via actin-based transport [Geisler et al., 2008; Gutierrez et al., 2009; Sampathkumar et al., 2013].

In animal cells, microtubules instead of actin filaments are the main driver of long-range transport [Akkerman et al., 2011; Bálint et al., 2013], likely due to difference in cell sizes and organization (cf. Section 3.2). It has been shown that cytoplasmic streaming in *Drosophila melanogaster* oocytes is related to structural features of microtubules [Ganguly et al., 2012], and that transport of lysosomes in monkey kidney cells depends on microtubules and microtubule cross-overs [Bálint et al., 2013]. Another prominent example are neurodegenerative diseases that may be caused by defects in the axonal cytoskeleton-based transport machinery of neurons [Chevalier-Larsen and Holzbaur, 2006; Kapitein and Hoogenraad, 2011]. Considering the cell's dire need of well-orchestrated delivery and removal of proteins and lipids, effects of cellular coordination that go beyond local correlations of cytoskeletal filaments and transport are likely to be found. We therefore believe that our findings on the coordination of Golgi transport in plant cells are transferable to a range of other scenarios of cellular transport, and can be readily investigated using our established methods.

5.4 Extension of the image-based cytoskeletal network representations by edge directionality and dynamics

Besides direct applications of the established network representations and analyses of the cytoskeleton, there are a number of promising extensions to incorporate additional biological features of the cytoskeleton. In particular, our and any other framework for the extraction of transportation networks from image data captures undirected edges only, assuming bi-directional transport of cargo [Baumgarten and Hauser, 2012; Dhondt et al., 2012; Obara et al., 2012b; Rigort et al., 2012]. However, many networks allow only unidirectional or biased transport (Fig. 5.1.5C, upper), such as venous blood vessels [Marieb and Hoehn, 2007] and fungal networks [Tero et al., 2010; Heaton et al., 2012]. Due to their polarity, also cytoskeletal filaments typically enable only unidirectional movement of motor proteins. This bias may, however, be attenuated by fluctuations in motor-protein movement, by different types of motor proteins, or by bundling of filaments with opposing polarity [Shimmen et al., 2000; Thomas et al., 2009; Akkerman et al., 2011]. Therefore, incorporation of edge directionality into the network framework is central for detailed investigations of cytoskeletal structure and transport. Since imaging techniques that allow inference of single-filament polarity are not in sight, utilization of cargo tracking data, as presented in this thesis, provides a promising route for

the inference of edge directionality (cf. Section 4.4).

Furthermore, current analyses of networks extracted from image data typically focus on static networks [Baumgarten and Hauser, 2012; Dhondt et al., 2012; Obara et al., 2012b; Rigort et al., 2012]. Here, we also showed that the actin network in growing plant cells exhibits properties that support efficient transport, at a given point in time (cf. Sections 4.1-4.4). However, the plant actin cytoskeleton is highly dynamic, and this dynamics has been suggested as a basis of different cytoskeletal functions [Staiger et al., 2009; Henty-Ridilla et al., 2013]. One the one hand, it has been shown that the mutual interaction of cytoskeletal rearrangements and cytoplasmic streaming may lead to the emergence of self-organized streaming patterns [Woodhouse and Goldstein, 2013; Goldstein and van de Meent, 2015]. On the other hand, we hypothesize that rearrangements of the actin cytoskeleton in growing (plant) cells provide an additional level of optimized transport efficiency. Namely, rearrangements of the cortical actin cytoskeleton increase both the effective speed of transported cargo and the area of the plasma membrane that is within reach of cytoskeletal cargo. Our network framework enables accurate incorporation and analysis of such rearrangements by modeling changes in nodes and edges over times (Fig. 5.1.5C, lower; [Dorogovtsev et al., 2000; Pastor-Satorras et al., 2001; Reggiani and Nijkamp, 2009]). Moreover, our proposed method for the identification of filaments from a given network representation may provide a promising link between the filament dynamics, which are guided by local, molecular processes, and the network dynamics on the global, cellular level (cf. Section 4.5).

5.5 Establishing a multi-scale perspective: From molecular processes to cytoskeleton to organismic phenotypes

A major benefit of the employed image-based network approach to plant cytoskeletal structure and transport is its top-down perspective. This perspective captures the system at an organizational level without the need for detailed knowledge of its molecular basis (cf. Section 3.5). At the same time, by focusing on the cytoskeleton at a cellular level, we disregarded the organismic context. While the distinctions between these levels are practical and partly justified by different time scales (typically, molecular processes: nanoseconds – seconds, cellular processes: seconds – hours, organismic processes: hours – years), a complete understanding of living systems requires a multi-scale perspective [Ingber, 1998; Kitano, 2002; Vaziri and Gopinath, 2008].

Regarding the molecular scale (Fig. 5.1.5D), all components of the cellular transport system, cytoskeletal and motor proteins as well as cargo, must eventually be synthesized. Hence, cellular transport is both directly based on and constitutive for translational and metabolic processes [Allan, 1995; Samaj et al., 2004; Kim and Coulombe, 2010]. Future studies may develop in the direction of correlating cytoskeletal transport activity with tissue- or even cell-specific translational or metabolic activity (cf. Section 5.2). In particular, our network-based framework of organelle transport may be used to infer sites of cell material depositioning and uptake (cf. Section 4.4). These predictions may then be verified using confocal imaging data from cells expressing fluorescent exo- and endocytosis markers [Samaj et al., 2004; Chen et al., 2011; Gadeyne et al., 2014; Luo et al., 2015].

At an organismic level (Fig. 5.1.5E), cytoskeletal structures and cytoskeleton-based transport are essential for cell growth, tissue formation, and organismic development. For example, plants with defects in actin

cytoskeletal organization show abnormal cell shapes [Fu et al., 2005; Smith and Oppenheimer, 2005; Ivakov and Persson, 2013] and reduced growth [Gibbon et al., 1999; Baluška et al., 2001]. Another striking example is the twisting of the plant hypocotyl due to impaired guidance of cellulose synthase complexes by the microtubule cytoskeleton [Landrein et al., 2013]. Here again, automated image-based phenotyping of the cytoskeletal networks may speed up analyses, may reduce bias from manual investigations, and enables establishment of quantitative associations between cytoskeletal organization and cellular output.

To conclude, our findings as well as our combination of experimental and theoretical methods offer significant contributions towards a system-level understanding of cytoskeletal transport in living organisms. Moreover, while our network-based top-down approach to the cytoskeleton is inherently independent of knowledge about molecular details and parameters, its future applications can be used to reveal such details. Thus, the work presented here provides quantitative means towards understanding the molecular basis of cytoskeletal transport functionality, as well as its future technological control and utilization.

6 Appendices

6.1 Appendix: Quantitative analyses of the plant cytoskeleton reveal underlying organizational principles

6.1.1 Experimental setup

Here, we describe the experimental setup for recording the cytoskeleton of growing plant cells. Dual-labeled *Arabidopsis thaliana* Columbia-0 seedlings were were previously described in [Sampathkumar et al., 2011]. The seedlings were surface sterilized (ethanol), stratified for 2 days at 4°C and germinated on MS agar plates (1X Murashige and Skoog salts, 8 g/L agar, 1X B5 vitamins, and 10.8 g 8 g/L sugar). All plants were grown in the dark on vertical plates at 21.8°C for 3 days. For treatment with Latrunculin B, seedlings were floated on distilled water containing 150 μM Latrunculin B and a set of control seedlings on pure water in 6-well plates. The seedlings were incubated in the dark with gentle shaking for 4 hours before imaging. For light treatments, a plate of dark-grown seedlings was exposed to light (150 μE m-2 s-1 PAR) for 4 hours while a control plate was kept in the dark, both plates being maintained in a vertical orientation. To study the effect of Latrunculin B, 5 control and 5 treated cells were imaged. For the analysis of the effect of light on the cytoskeleton, 35 control and 26 treated cells were imaged. To fix the seedlings and to avoid mechanical damage, they were mounted between a cover glass and a 1 mm thick 1% agar pad affixed on a circular cover slip. A detailed description of the microscopy setup is given in [Sampathkumar et al., 2011]. Typical exposure times were 400 ms for GFP and 300 ms for mCherry with a time interval of 2 s between subsequent actin and microtubule images, respectively. The cells were recorded for at least 4 min. Only seedlings expressing both fluorescent markers were used for further analyses.

6.1.2 Network reconstruction procedure, different grid topologies and null models

We now explain our method for reconstructing a (edge-weighted, undirected) network from the confocal image series of actin filaments or microtubules and present various extensions and null models. We preprocessed the recorded image series using Fiji [Schindelin et al., 2012]: First, the drift of the seedlings under the microscope was corrected using a stack registration algorithm [Arganda-Carreras et al., 2006]. Here, the microtubule images were registered first because they are less dynamic and easier for the program to align. The more dynamic actin filaments were then subjected imagewise to the same transformations. Second, the image series were rotated so that the shoot apical direction of the cell pointed upwards. The region of interest, i.e., the interior of the cell, was cropped manually and identically for the corresponding actin and microtubule frames. Third, the background noise was reduced by applying a rolling ball background subtraction [Sternberg, 1983], with a ball radius of 50 pixels. The noisy background signal arises largely from fluorescent monomeric actin/tubulin in the cytosol which is not incorporated in

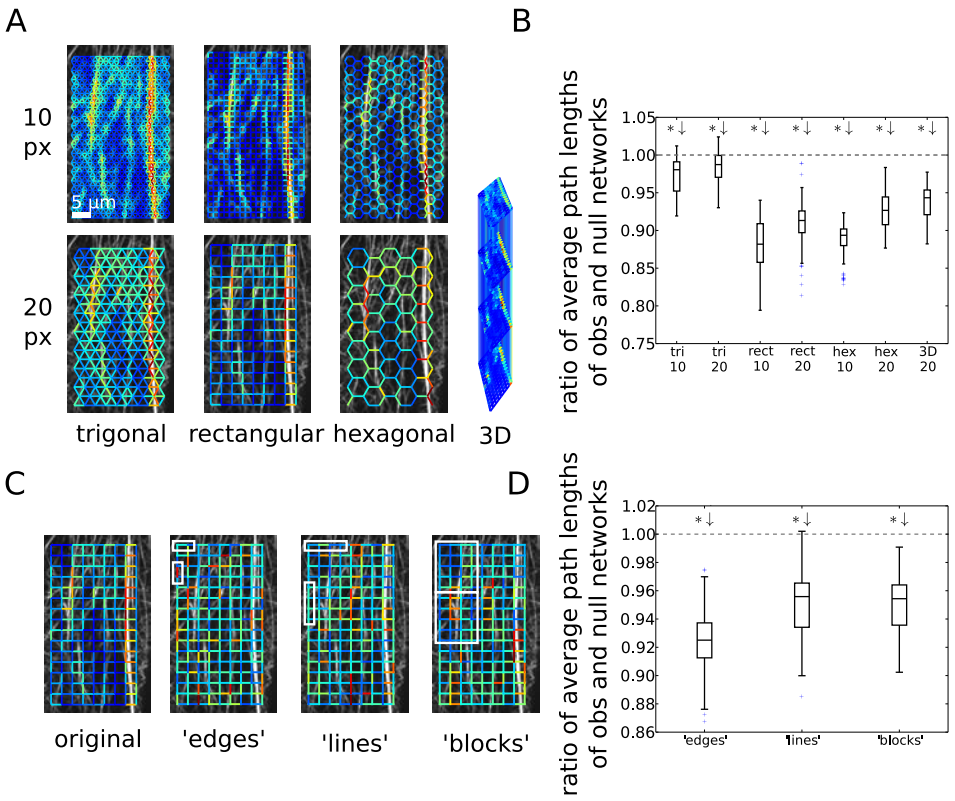


Figure 6.1.1: **Extensions of the network reconstruction framework and different null models.** **(A)** Reconstructions of the actin cytoskeleton of the same cell using different sizes (grid spacings of 10 and 20 pixels) of triangular, rectangular, and hexagonal grids, and a three-dimensional grid with 20 pixel-spacings. **(B)** All studied grid types exhibit smaller average path lengths than their respective null model networks. **(C)** Different null models for the reconstructed network were obtained by shuffling all edge weights, by shuffling connected vertical and horizontal lines, and by rearranging blocks of varying size and shuffling the remaining edge weights, respectively. White rectangles exemplify sections of the cytoskeletal network that were shuffled. **(D)** All proposed null models yield values below one for ratio of the average path lengths of observed and null model networks, suggesting a non-random and efficient organization of the cytoskeletal network.

filaments and, hence, was filtered out in the present study. Finally, photobleaching was compensated by rescaling all images' mean intensities to one.

From these preprocessed images we obtained the cytoskeletal components as complex networks through a two-step procedure as described in the Results in Section 4.1. These steps and all statistical analyses were performed using Python [Van Rossum and Drake, 2011]. We chose an equidistant rectangular grid with a spacing of 10 pixels and the standard deviation of the Gaussian convolutions kernels was 4 pixels in x- and y-direction, unless stated otherwise. An extension of our framework is the three-dimensional network that may be easily reconstructed from three-dimensional confocal microscope image series (Fig. 4.1.1G). First, a network was reconstructed for each z-slice as for the two-dimensional case. Then, for a rectangular grid, the networks of neighboring z-slices were connected by creating a link between nodes with the same x-y-coordinates. The weights of these links were computed by creating Gaussian convolution kernels (pointsymmetric, with the same width as for the edges, i.e., 4 pixels) for its nodes, multiplying them with the two respective z-slices and averaging over the sum of the resultant images. The three-dimensional network reconstruction captures cytoskeletal filaments which leave the cortical plane and may be analyzed using the same network-based methods as for the two-dimensional networks.

To ensure that our findings (e.g., on the transport efficiency of the cytoskeletal network architecture that displays short APLs; see Results in Section 4.1), are valid not only for rectangular grids, we tested other grid types. We reconstructed the cytoskeletal actin network of the same cell based on two-dimensional rectangular, triangular, and hexagonal grids with spacings of 10 and 20 pixels and a three-dimensional grid with a uniform spacing of 20 pixels (Fig. 6.1.1A). For these networks, we compared the APLs to an ensemble of null model networks obtained by edge-shuffling. In all chosen grid types, the APL of the observed network is significantly shorter than expected from the respective null model (Fig. 6.1.1B; one-sample two-sided t-test: all $p-$values < 0.05). Hence, the short APLs of the cytoskeleton are a non-random and biologically relevant feature which does not arise as an artifact of the imposed grid type. As network properties are often dependent on each other, the findings from the comparative analysis suggest that a significant change of other network properties compared to their null model values is largely independent of the underlying grid type, as long as the grid is not too dense, covers the cell too inhomogeneously (e.g., random geometric graphs), or has non-local, long-range links (e.g., scale-free graphs).

Judging the biological relevance of a network property's value requires a meaningful comparison since its value depends on the normalization of the image and is therefore arbitrary. The simplest reference is given by the values of the respective network property which are obtained for null model networks with shuffled edge weights (Fig. 6.1.1C, "edges"). Such networks preserve both the node positions and the distribution of edge weights and thus the total amount of cytoskeletal components in the cell. By comparing the value of a given network property of a reconstructed cytoskeletal network against those of an ensemble of edge-shuffled null model networks, we were able to assess whether a random distribution of cytoskeletal material in the cell results in the same cytoskeletal properties as realized in the observed cell, see Results in Section 4.1. For example, for the APL, the ratio of observed and null model values falls significantly below one (Fig. 6.1.1D, "edges"; one-sample two-sided t-test: $p-$value < 0.05).

We also investigated two alternative null models to strengthen the assessment of the biological relevance of different network properties. Like the first null model, these, too, preserve the positions of the nodes of

the network and the distribution of edge weights in the network. In addition, they leave more of the local cytoskeletal structure intact. While in the first null model, edge weights were shuffled irrespective of the edges to which they were assigned, we now cut all edges forming connected horizontal or vertical lines into several equal sections, respectively, which were in turn shuffled (Fig. 6.1.1C, "lines"; here, horizontal lines are divided into three and vertical lines in four sections). This method better preserves potential strongly weighted paths and hence the filamentous structures of the cytoskeleton. As for the "edge" null model, this "line" null model exhibits longer APLs than the observed network (Fig. 6.1.1D, "lines"; one-sample two-sided t-test: $p-$value < 0.05). Clearly, the focus on horizontal or vertical lines imposes a restriction to the orientation of potential filaments. To circumvent this limitation, we analyzed a third null model in which connected, non-overlapping blocks of nodes were chosen. The subgraphs formed by these blocks were shuffled as well as the remaining edge weights that were not part of any subgraph (Fig. 6.1.1C, "blocks"; here, the network is composed into three times four blocks). This "block" null model also exhibits longer APLs than its biological counterpart (Fig. 6.1.1D, "blocks"; one-sample two-sided t-test: $p-$value < 0.05). More sophisticated null models may be proposed. However, the investigation of three different null models that capture the amount of cytoskeletal components in the cell and their filamentous structure provided consistent results on the non-randomness of various cytoskeletal network properties. Using the simple null model was, therefore, considered reliable for assessing the biological relevance of the studied network properties (see Results in Section 4.1).

6.1.3 Network properties used for quantifying the cytoskeletal organization

We represent the cytoskeleton as a weighted, undirected network and quantify its structure via a number of seminal network properties. Here, we explain the employed properties in more detail and provide careful interpretations of how they relate to the structural (and potentially: functional) features of the cytoskeleton. In the following, we consider a weighted, undirected network $G = (\mathcal{N}, \mathcal{E})$ with a set $\mathcal{N}$ of $N = |\mathcal{N}|$ nodes, a set $\mathcal{E}$ of $E = |\mathcal{E}|$ undirected edges $e = (n, m) \equiv (m, n) \in \mathcal{E}$ and $m, n \in \mathcal{N}$ with weights w_e.

Degree distribution: The degree d_n of a node $n \in \mathcal{N}$ is given by the sum of its edge weights w_e, i.e.,

$$d_n = \sum_{\substack{e \in \mathcal{E} \\ n \in e}} w_e \tag{6.1.1}$$

(cf. Fig. 6.1.2A for the degree of a node of the network in panel E; color-coded node degrees in Figs. 6.1.2E-H). Since the edge weights reflect the intensity of cytoskeletal structures close to the respective edges, the node degrees reflect the cytoskeletal intensities in the vicinity of the respective nodes. Therefore, the standard deviation of the degree distribution (SDD) captures the spatial heterogeneity of the distribution of intensities in the underlying cytoskeleton images,

$$\mathrm{SDD} = \left(\left(N^{-1} \sum_{n=1}^{N} d_n^2 \right) - \left(N^{-1} \sum_{n=1}^{N} d_n \right)^2 \right)^{-1/2}. \tag{6.1.2}$$

In particular, the SDD does not measure the heterogeneity in filament thicknesses or numbers but a com-

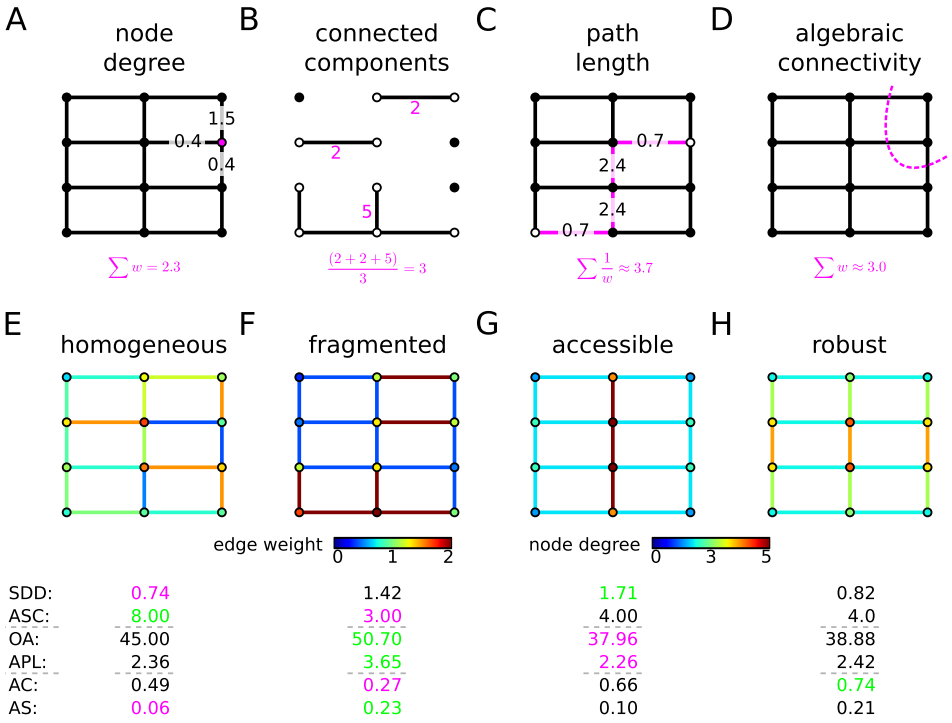


Figure 6.1.2: **Explanations of different network properties and exemplary networks with different structural and functional characteristics.** Panels (A)-(D) display schematic illustrations of different network properties. Panels (E)-(H) show typical paradigmatic networks with color-coded edge weights and node degrees. **(A)** The degree of a node is given by the sum of its edge weights, cf. color-coding in panel (E). **(B)** When removing edges with weights below the 50th percentile from the network in panel (F), the network decomposes into trivial components with just a single node (black circles) and several bigger, non-trivial components (white circles). **(C)** The shortest path length between two nodes is given by a sequence of edges whose sum of inverse weights is minimal, cf. network in panel (G). **(D)** The algebraic connectivity relates to the minimum sum of the weights of edges that need to be removed to disconnect the network, cf. network in panel (H). **(E)** The network exhibits a small standard deviation of the degree distribution (SDD), has a large average size of the connected components (ASC) after thresholding, and a small assortativity (AS). **(F)** The network has a small ASC, its overall angle (OA) indicates a horizontal orientation (OA > 45°), it displays a large average path length (APL), a small algebraic connectivity (AC), and a large assortativity. **(G)** The network has a high SDD, displays a vertical OA (OA < 45°), and a small APL. **(H)** The network was obtained by maximizing the AC for a fixed sum of weights (cf. discussion of Fig. 4.1.4) and, accordingly, displays a high AC.

bination thereof. Comparing Figs. 6.1.2E and G, we find that the SDD of the former network is smaller, indicating a more homogeneous spatial distribution of the cytoskeleton in agreement with the visual impression.

Connected components: By construction, all edges in the reconstructed networks are strictly positive (because the Gaussian convolution kernels are strictly positive and have infinite support and the image intensity is greater than zero somewhere in the image) and, hence, all their nodes are connected. However, when removing edges, e.g., with small weights by arguing that they do not permit transport of cargo, the network may disconnect and split into several connected components (cf. Fig. 6.1.2B for the thresholded version of the network in panel F). The components are called trivial if they consist of a single node only, and non-trivial otherwise. For simplicity, we use the 50th percentile when thresholding the edge weights throughout the manuscript but our findings remain qualitatively unchanged when different, reasonable thresholds are chosen. The average number of nodes, or size, per non-trivial connected component (ASC) is a measure for the fragmentation of the network.

$$\mathrm{ASC} \quad = \quad C^{-1} \sum_{c=1}^{C} N_c, \tag{6.1.3}$$

where $\mathscr{C}$ is the set of $C = |\mathscr{C}|$ non-trivial components and N_c is the number of nodes in component $c \in \mathscr{C}$. Comparing Figs. 6.1.2E and F, we find that the ASC is smaller in the latter which clearly exhibits several small, densely connected fragments separated by weak connections.

Overall angle: As our network representation does not resolve individual filaments, we can not evaluate their orientations individually. However, our approach allows to infer an overall angle (OA) for the orientation of the cytoskeletal structures as a whole. The OA is given by Eq. (6.1.12) and its derivation is explained in ESM4. Two networks with different OA are shown in Figs. 6.1.2F and G with overall horizontal (OA $> 45°$) and vertical (OA $< 45°$) orientations, respectively, as confirmed visually.

Average path length: A path $\mathscr{P}$ between two nodes is a sequence of edges connecting the nodes. A shortest path is a path that minimizes its sum of edge lengths (cf. Fig. 6.1.2C for a shortest path in the network in panel G). Here, for simplicity, we take the length of an edge to be the inverse of its weight. This choice takes into account that parts of the cytoskeleton that yield strong edge weights potentially allow faster/more transport as reflected by small edge lengths.

$$\mathrm{APL} \quad = \quad 2^{-1} N^{-1} (N-1)^{-1} \sum_{n=1}^{N} \sum_{\substack{m=1 \\ m>n}}^{N} \min_{\mathscr{P} \in \mathscr{P}_{n,m}} \sum_{e \in \mathscr{P}} w_e^{-1}, \tag{6.1.4}$$

where $\mathscr{P}_{n,m}$ is the set of all paths from node n to m and w_e^{-1} is the length of the edge $e \in P \in P_{n,m}$. As explained in the discussion of Fig. 4.1.3 in the main text, highway-like structures may yield small average path lengths (APL) as they act as short cuts between distant parts of the network. Such a highway-like structure is given by the network in Fig. 6.1.2G which, accordingly, displays a smaller APL than, e.g., the easily-fragmented network in panel F.

Algebraic connectivity: The algebraic connectivity is the second smallest eigenvalue

$$\mathrm{AC} \quad \equiv \quad \lambda_2 \tag{6.1.5}$$

of the graph Laplacian L,

$$L_{n,m} \quad = \quad \begin{cases} d_n & \text{if } n = m \\ -w_{(n,m)} & \text{if } (m,n) \in \mathscr{E}\,, \\ 0 & \text{otherwise,} \end{cases} \tag{6.1.6}$$

with $n, m \in \mathscr{N}$. By construction of L, its smallest eigenvalue $\lambda_1 = 0$ and the number of zero eigenvalues provides the number of connected components in the graph (cf. e.g. [Newman, 2009]). As our reconstructed networks are always connected (see the discussion of the connected components above) they yield $\lambda_2 \equiv \mathrm{AC} > 0$. The magnitude of the AC is commonly interpreted as a measure for how well-knit the network is, which is related to the minimum sum of the weights of edges that need to be removed in order to disconnect the network (cf. Fig. 6.1.2D). We solved a semi-definite optimization problem described in the discussion of Fig. 4.1.4 to construct a network with a fixed sum of edge weights that maximizes the AC. This network is shown in Fig. 6.1.2H and its AC is larger than, e.g., that of the easily-fragmented network in panel F.
We note that small APL and large AC favor different types of networks (cf. Figs. 6.1.2G and H). This may be explained as follows: In the computation of the APL for each shortest path only one edge may be used at a time. In contrast, the AC is related to cuts (i.e., the removal of sets of edges) that disconnect the network and hence affect multiple edges. Thus, the APL and the AC are independent network properties that provide insights into different potential, transport-related functions of the cytoskeleton.

Assortativity: The assortativity denotes the correlation of the degrees of neighboring nodes

$$\mathrm{AS} \quad = \quad \frac{1}{2E} \sum_{n=1}^{N} \sum_{m=1}^{N} \left(w_{(n,m)} - \frac{d_n d_m}{2E} \right) d_n d_m. \tag{6.1.7}$$

Similar to the SDD (see above) the AS captures the spatial heterogeneity of the cytoskeletal components but contains additional information about its spatial distribution: The AS is high if nodes with high (low) degrees are also connected to nodes with similar degrees, hence detecting regions of spatially clustered cytoskeletal structures. For instance, Fig. 6.1.2E displays a network with low AS because the are no regions of nodes of high or low degree clustered together, while the network in panel F shows high AS values that reflect regions of high and low node degrees, respectively.

6.1.4 Method for determining angles and filament orientations from network structure

We now present a method to evaluate the orientation of AFs and MTs by exploiting their network structures. Starting from a given grid (Fig. 6.1.3A), we placed a stiff rod of length L that is rotated by an angle α and

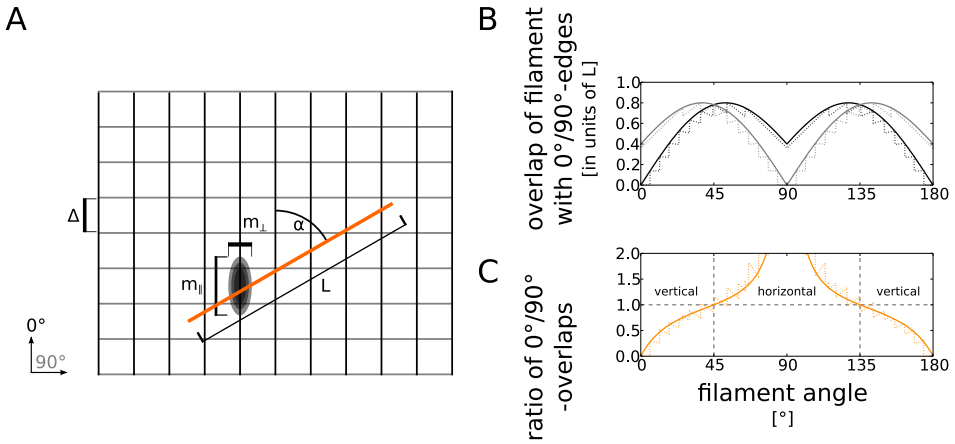


Figure 6.1.3: **Inferring orientation of filament from network structure.** **(A)** Schematic filament of length L at an angle α, placed on a rectangular grid with spacing Δ, vertical (0°, black), and horizontal edges (90°, gray). The shape of the edges' convolution kernels determines the contribution of the filament to each edge, with the two extremes $m_\parallel$ and $m_\perp$ for parallel and perpendicular orientation of the filament with respect to the edge. **(B)** Contribution of filaments ($L = 1000$, solid lines; $L = 10$, dotted lines) to edges of different orientation (0°, black; 90°, gray) for varying angles α. By including that filaments cross only integer numbers of edges, we obtain curves that are strongly discontinuous for short filaments and become less discontinuous for longer filaments. **(C)** The ratio of filament overlaps r with 0°- and 90°-edges determines the filament angle ($L = 1000$, solid line; $L = 10$, dotted line), with $r < 1$ and $r > 1$ corresponding to a vertical and horizontal orientation, respectively.

computed its contribution to the weight of edges with an orientation of angle γ. Assuming a regular grid in which edges with angle γ are distributed with uniform distances Δ, we calculated the number n_γ of crossed γ-edges,

$$n_\gamma \quad = \quad \Delta^{-1} L \left|\sin\left(\alpha-\gamma\right)\right|. \tag{6.1.8}$$

The overlap m of the rod and a γ-edge was computed via the convolution kernel of that edge (see Appendix 6.1.2) and was approximated as

$$m_\gamma \quad = \quad m_\perp + \left(m_\parallel - m_\perp\right)\left|\cos\left(\alpha-\gamma\right)\right|, \tag{6.1.9}$$

where $m_\perp$ and $m_\parallel$ are the contributions of the rod to the edge when they are perpendicular or parallel to each other. The total contribution of the rod to all γ-edges is (Eqs. (6.1.8) and (6.1.9))

$$w_\gamma \quad = \quad n_\gamma m_\gamma = \Delta^{-1} L \left|\sin\left(\alpha-\gamma\right)\right| \left[m_\perp + \left(m_\parallel - m_\perp\right)\left|\cos\left(\alpha-\gamma\right)\right|\right]. \tag{6.1.10}$$

Furthermore, we may include that only integer numbers of edges may be crossed by a filament. Then, $L\sin$ in Eq. (6.1.10) is replaced by $\lfloor L\sin \rfloor$ and the contribution of a filament of finite length becomes discontin-

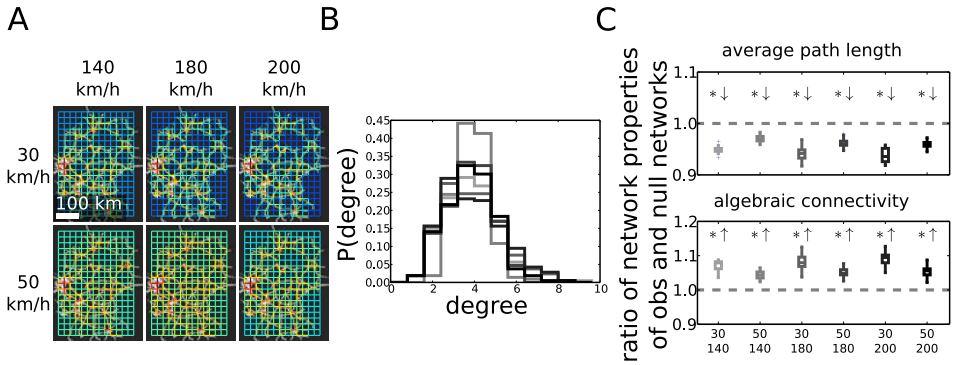


Figure 6.1.4: **The efficiency of the autobahn network is pertained for different choices of speed limits. (A)** Reconstructed autobahn networks with different maximum high speed limits (140, 180, and 200 km/h) and different off-highway speed limits (30 and 50 km/h). **(B)** The degree distributions of the autobahn networks with the different speed limits given in (A) (from light gray to black) are unimodal and centered around their means (excess kurtosis > 0). **(C)** The ratios of average path lengths and algebraic connectivities of observed and null model networks are well below and above one, respectively, for all autobahn networks with speed limits described in (A).

uous and approaches the continuous curve for long filaments (Fig. 6.1.3B). For simplicity, we considered small grid spacings or long filaments and work with Eq. (6.1.10) directly. Each cytoskeleton fluorescence image contains many filaments and their lengths and orientations can not be inferred uniquely from the distribution of edge weights. Because $m_\perp$ and $m_\parallel$ are determined by the convolution kernels, the ratio of the weights of two edge types with different orientations γ and γ' yields an equation for α,

$$\frac{w_\gamma}{w_{\gamma'}} = \frac{|\sin(\alpha-\gamma)|\,[m_\perp+(m_\parallel-m_\perp)\,|\cos(\alpha-\gamma)|]}{|\sin(\alpha-\gamma')|\,[m_\perp+(m_\parallel-m_\perp)\,|\cos(\alpha-\gamma')|]} =: r, \tag{6.1.11}$$

where α may be interpreted as the overall orientation of the cytoskeletal filaments (Fig. 6.1.3C). For $\gamma = 0°$ and $\gamma' = 90°$, Eq. (6.1.11) yields

$$\alpha = \arctan\left(\frac{m_\perp + r\,(m_\parallel - m_\perp)}{m_\perp r + (m_\parallel - m_\perp)}\right). \tag{6.1.12}$$

In our analysis, we refer to $\alpha \in [0°, 45°)$ and $\alpha \in (45°, 90°]$ as an overall vertical and horizontal orientation, respectively. See Results in Section 4.1 for results on the orientation of the cytoskeletal components under different conditions.

6.1.5 Reconstruction and analysis of the German autobahn network

Here, we describe the data and the procedure used for reconstructing the German autobahn as a network for a comparison with the plant cytoskeleton. Further, we present two examples of networks with different

structural and transport-related properties.

An OpenStreetMap (© OpenStreetMap contributors; map data available under the Open Database License (ODbL)) data set of Germany was downloaded (http://download.geofabrik.de/europe/germany.html; © Geofabrik GmbH Karlsruhe), converted to .o5m for faster filtering (http://wiki.openstreetmap.org/wiki/Osmconvert), and filtered for objects of type "highway=motorway" (http://wiki.openstreetmap.org/wiki/Osmfilter). The remaining motorways were parsed in Python. Speed limits were take into account for better analogy with the cytoskeleton that exhibits thinner and thicker bundles that were argued in the main text to allow different net transportation speeds. Because some sections of the autobahn were assigned no speed limits (either because of lacking data or the absence of a speed limit) and to incorporate transportation outside of the autobahn, we chose different settings to ensure the robustness of our findings: Missing autobahn speed limits were set to 140, 180, and 200 km/h and the speed limit in the rest of Germany was set to 30 and 50 km/h, respectively (Fig. 6.1.4A).

The results for speed limits of 200 km/h and 50 km/h (Fig. 4.1.4; section "The cytoskeleton and the German autobahn exhibit similar network properties") demonstrate that the autobahn network displays a unimodal degree distribution that peaks around its mean and that it exhibits significantly shorter path length and a significantly higher AC than the null model networks. The same holds true for all considered speed limits (Fig. 6.1.4B and C; one-sample two-sided t-test: all $p-$values < 0.05). Hence, the findings on the efficiency of the autobahn networks are robust against moderate changes of the speed limits.

6.2 Appendix: Cell-specific organization and optimality of the plant actin cytoskeleton

6.2.1 Multi-objective optimization of weighted networks using semidefinite programming

Optimality of cytoskeletal network organization and trade-offs between different network properties in various cell types was studied using multi-objective optimization via semidefinite programming [Boyd, 2006; Wolkowicz et al., 2012]. Since the assortativity and the average size of the connected components cannot be expressed, or conveniently approximated, via a semidefinite program, we focus here on the Herfindahl index, the average path length, the average maximum flow, and the algebraic connectivity. The studied multi-objective function is a linear combination of our network properties of interest (cf. Eqs. 4.3.1, 4.3.3, 4.3.4, and 4.3.5),

$$\begin{aligned} \max \quad & c_H \cdot v_H^{'} + c_L \cdot v_L^{'} + c_F \cdot v_F^{'} + c_C \cdot v_C^{'} && (6.2.1)\\ \text{subject to} \quad & \sum_{e\in\mathscr{E}} w_e = 1 && (6.2.2)\\ & p_H, p_L, p_F, p_C, && (6.2.3) \end{aligned}$$

where the first two lines are explained in the Materials and Methods and the third line indicates the constraints that are explained below. The first set of constraints for the minimization of the Herfindahl index is given by

$$\begin{aligned} p_H:\ v_H^* &= -\sum_{e\in\mathscr{E}} h_e, && (6.2.4)\\ h_e &= w_e^2, \\ 0 \le \quad h_e, w_e, v_H^* &\le 1\ \forall e \in \mathscr{E}, \end{aligned}$$

where the first two quadratic equations may readily be expressed as a semidefinite constraint,

$$v_H^* \ge \mathbf{w}^T 1^T 1 \mathbf{w} \quad \Leftrightarrow \quad \begin{pmatrix} v_H^* & w \\ w^T & 1 \end{pmatrix} \succeq 0. \tag{6.2.5}$$

While the minimization of the average path length for a given network with fixed edge weights can be written as a linear program, simultaneous optimization of the edge weights leads to a mixed-integer, or mixed-binary, linear program [Jensen and Barnes, 1980; Schrijver, 1998]. Despite efficient solvers [Linderoth and Ralphs, 2005; Achterberg, 2009], such problems are not feasible for networks of the size studied here (for $N = \mathcal{O}\left(10^3\right)$ nodes there are $P = \mathcal{O}\left(10^6\right)$ paths between all pairs of nodes and together with $E = \mathcal{O}\left(10^3\right)$ edges this implies $P \cdot E = \mathcal{O}\left(10^9\right)$ binary variables). Therefore, we used a linear relaxation of the mixed-binary

problem so that the second set of constraints for the minimization of the average path length is given by

$$\begin{aligned}
p_L:\ v_L^* &= -(2EP)^{-1}\sum_{p\in\mathscr{P}}\sum_{e\in\mathscr{E}} b_{p,e}^+ + b_{p,e}^-, \\
l_e a_{p,e}^+ &= b_{p,e}^+ \ \forall p\in\mathscr{P}, e\in\mathscr{E}, \\
l_e a_{p,e}^- &= b_{p,e}^-, \\
l_e &= w_e^{-1} \ \forall e\in\mathscr{E}, \\
\sum_{m\in\mathscr{N}} a_{p,e=(n,m)}^+ - a_{p,e=(m,n)}^- &= \begin{cases} 1 & \text{if } n = p_{\text{source}} \\ -1 & \text{if } n = p_{\text{sink}} \\ 0 & \text{otherwise} \end{cases}, \ \forall p\in\mathscr{P}, \\
0 \le \quad a_{p,e}^+, a_{p,e}^-, v_L^* &\le 1 \ \forall p\in\mathscr{P}, e\in\mathscr{E}, \\
0 \le \quad b_{p,e}^+, b_{p,e}^- &\le \infty, \\
1 \le \quad l_e &\le \infty,
\end{aligned} \tag{6.2.6}$$

where $\mathscr{P} = \{(p_{\text{source}}, p_{\text{sink}}) \mid p_{\text{source}}, p_{\text{sink}} \in \mathscr{N}\}$ is the set of all $P = |\mathscr{P}|$ pairs of nodes. Moreover, we employed the length $l_e = w_e^{-1}$ of an edge, given by its inverse weight. While this constraint is quadratic and non-convex, its relaxation $l_e w_e \ge 1$ may be approximated using a semidefinite programming relaxation,

$$\begin{aligned}
\frac{1}{2}(\mathbf{w},\mathbf{l})\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}(\mathbf{w},\mathbf{l})^T \ge 1 \quad &\Leftrightarrow \quad \frac{1}{2}\text{Trace}\left(V\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\right) \le 1 \\
&\qquad V = (\mathbf{w},\mathbf{l})(\mathbf{w},\mathbf{l})^T, \\
\text{relaxation}: \quad &\Rightarrow \quad \frac{1}{2}\text{Trace}\left(V\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\right) \le 1 \\
&\qquad V \succeq (\mathbf{w},\mathbf{l})(\mathbf{w},\mathbf{l})^T.
\end{aligned} \tag{6.2.7}$$

The third set of constraints for the maximization of the average maximum flow is given by a linear program [Jensen and Barnes, 1980],

$$\begin{aligned}
p_F:\ v_F^* &= P^{-1}\sum_{p\in\mathscr{P}} y_p, \\
0 \le \quad x_{p,e}^+ &\le w_e \ \forall p\in\mathscr{P}, e\in\mathscr{E}, \\
0 \le \quad x_{p,e}^- &\le w_e \\
\sum_{m\in\mathscr{N}} x_{p,e=(n,m)}^+ - x_{p,e=(n,m)}^- &= \begin{cases} y_p & \text{if } n = p_{\text{source}} \\ -y_p & \text{if } n = p_{\text{sink}} \\ 0 & \text{otherwise} \end{cases} \quad \forall p\in\mathscr{P}, \\
0 \le \quad x_{p,e}^+, x_{p,e}^-, y_p, v_F^* &\le 1 \ \forall p\in\mathscr{P}, e\in\mathscr{E}.
\end{aligned} \tag{6.2.8}$$

Finally, the fourth set of constraints for the maximization of the algebraic connectivity is given by a semidefinite program [Boyd, 2006; Sun et al., 2006]

$$\begin{aligned} p_C: \; v_C^* &= \gamma, \\ \gamma 1 - \beta 1 1^T - \mathscr{L} &\preceq 0, \\ 0 \leq \; v_C^*, \gamma, \beta &\leq 1 \;\; \forall p \in \mathscr{P}, e \in \mathscr{E}, \end{aligned} \tag{6.2.9}$$

where $\mathscr{L}$ is the graph Laplacian (cf. Eq. 4.3.6).

6.3 Appendix: System-wide organization of the actin cytoskeleton drives organelle transport in plant cells

6.3.1 Confocal recording of actin cytoskeleton and Golgi bodies in hypocotyl cell of FABD-GFP and tdT-CesA6 dual-labeled *Arabidopsis thaliana* seedlings

Confocal recording of actin cytoskeleton and Golgi bodies in hypocotyl cell of FABD-GFP and tdT-CesA6 dual-labeled *Arabidopsis thaliana* seedlings (cf. digital version or http://mathbiol.mpimp-golm.mpg.de/CytoSeg/Data/CytoSeg_movie.avi).

6.3.2 Generation of artificial cytoskeleton images, manual segmentation of actin images, and comparison of network extraction procedures

The network extraction procedure employed four parameters ν_{width}, ν_{thres}, ν_{size} and ν_{int}. The parameters reflect the width of the detected filaments, the intensity threshold that separates actin foreground from cellular background and the minimal size and minimal average intensity of the segmented filament center lines (cf. Figs. 4.4.1A-D). Their optimal values were determined in a gauging step employing contrived images of known, cytoskeleton-like structures and manual segmentations of biological cytoskeleton images as a gold standard for comparison against the automated segmentation results.

First, for the manual segmentations, we randomly chose 20 of the investigated actin cytoskeleton images. For these images, we performed the preprocessing steps described in Materials and Methods in Section 4.4 and manually segmented the center lines of the filaments (cf. 4.4.1H).

Second, we created 20 empty images of 100×500 pixels in size, comparable to the typical cellular region of interest. For each image, we created ten auxiliary images by drawing two random points uniformly distributed over the image area, adding a line of unit intensity between these points, blurring the line with a Gaussian filter of random standard deviation uniformly drawn from the interval $[0.5, 1.5]$ and rescaling the average image intensity to one. We superimposed all ten auxiliary image to obtain a cytoskeleton-like structure of filaments of different thicknesses. Finally, we added gamma distributed noise such that both mean and standard deviation of the contrived image intensity were similar to those of the biological images (cf. 4.4.1G). While noise in digital camera images is typically a combination of Gaussian-distributed sensor noise, gamma-distributed speckle noise and Poisson-distributed shot noise, we adhered to simple gamma distributed noise for simplicity here, which was similar to that in the biological cytoskeleton images.

Finally, we briefly review other, existing methods for the extraction of networks from different image sources and systems of interest. The first class of approaches typically relies on two-dimensional image data and employs classical image segmentation for network extraction: From high contrast dark-field microscopy images of leaves, venation networks were extracted in a supervised procedure [Dhondt et al., 2012]. However, due to the high signal-to-noise ratio (SNR), no filters to enhance the signal of the curvilinear veins were applied, as necessary for the cytoskeletal images. Edge weights capturing the thickness of the underlying structures were not computed, but are essential for understanding the cytoskeleton. Moreover, only

the largest connected component of the venation network was kept, which removes smaller, disconnected parts of the network that commonly occur in the cytoskeleton. Finally, while the supervised adjustment of extraction parameters increases flexibility, it requires manual input, thereby also introducing user bias.

Using a fully automated extraction procedure, photographic image series were used to extract networks formed by slime moulds [Baumgarten and Hauser, 2012]. Again, the high SNR did not necessitate enhancement of curvilinear structures. A background subtraction was performed by subtracting the first frame of the image series before the growth of the slime mould which is not applicable to the cytoskeletal image data. The thickness of the veins was measured by counting the average number of pixels per vein length which is reasonable for slime mould images. Yet, this measure ignores the underlying image intensity which is a better estimator of thickness for cytoskeletal filaments and bundles whose diameter is typically well below the resolution limit of confocal microscopes [Cooper, 2000]. The parameters of the procedure were determined manually without gauging or comparison to a known gold standard.

Our network extraction procedure was inspired by another, more sophisticated method for the extraction of fungal networks from photographs [Obara et al., 2012b]. While there the curvilinear vein structures were enhanced using a contrast-independent phase-congruency filter [Obara et al., 2012a], we employed a faster and widely used tubeness filter for simplicity [Sato et al., 1998]. As for the slime mould network above, the vein thickness was only determined based on the average pixel numbers per vein length and we extended the thickness computation by taking into account pixel intensities. Moreover, we adopted the gauging of the free image analysis parameters by computing the average smallest distances between the center lines of manually segmented gold standard images and automated segmentations. However, [Obara et al., 2012a] considered only one direction, $d_{\text{manu}\to\text{auto}}$, which clearly favors undersegmentation, i.e. an excess of pixels in the automated segmentation (cf. Fig. 4.4.1I). Instead, we included the opposite direction and minimized the Haussdorf distance $d_{\text{HD}} = \frac{1}{2}\left(d_{\text{manu}\to\text{auto}} + d_{\text{auto}\to\text{manu}}\right)$ in the gauging step to avoid both under- and oversegmentation (cf. Fig. 4.4.1J; [Mayerich et al., 2012; Xu et al., 2015]). Thus, although the average smallest distance $d_{\text{manu}\to\text{auto}} \lesssim 1$ pixel in [Obara et al., 2012a,b] was smaller than our $d_{\text{HD}} \lesssim 3$ pixels, this is due to the poorer SNR of the original cytoskeletal images as well as the trade-off between under- and oversegmentation incorporated in d_{HD}.

Another class of approaches for the detection of networks from image data does not rely on segmentation but encompasses tracing-, tracking- or open contour-based approaches [Mayerich and Keyser, 2008; Meijering, 2010; Xu et al., 2014; Peng et al., 2015a; Xu et al., 2015]. Instead of (global) image segmentation, these approaches typically identify two or more points on the network and find a connecting path through the network by optimization of an (usually local) energy function. Since many of these approaches have been originally developed for the reconstruction of neural networks, they are directly suitable for the extraction of three-dimensional network structures. However, since most existing approaches in this class require manual user input, we focus our discussion on a recent, fully automated open contour-based approach, called SOAX [Xu et al., 2014, 2015], whose strengths and limits are large representative for the related approaches: Especially for large sets of three-dimensional image data, SOAX is faster than segmentation-based approaches. Furthermore, we note that SOAX incorporates resampling of the three-dimensional image data to compensate the decrease in SNR towards to distant side of the cell (cf. discussion of periodic boundary conditions in the main text). However, while our highly time-resolved recordings of the cytoskeleton displayed low SNR ≈ 1, the image data studied in [Xu et al., 2015] displayed high SNR ≈ 3 which rendered filters for the

(A) actin network edge e			**(B) actin network**		
property	**symbol**	**Eq.**	**property**	**symbol**	**Eq.**
Euclidean edge length	$a_{e,E}$	-	avg. frac. nodes / conn. comp.	S	6.3.5
filament edge length	$a_{e,F}$	-	avg. edge capacity	$\mathrm{E}[a_c]$	6.3.6
filament edge length	$a_{e,B}$	6.3.1	assortativity	A	6.3.7
edge weight	$a_{e,w}$	-	rel. crossing number	X	6.3.8
edge capacity	$a_{e,c}$	6.3.2	avg. shortest path length	$\mathrm{E}[L]$	6.3.9
edge length	$a_{e,l}$	6.3.3	CV of shortest path lengths	$\mathrm{CV}[L]$	6.3.10
edge angle	$a_{e,a}$	6.3.4	algebraic connectivity	C	6.3.11
edge degree	$a_{e,\mathrm{deg}}$	6.3.13	edge angles	$\mathrm{CV}[a_a]$	6.3.12
edge rank	$a_{e,\mathrm{rank}}$	6.3.14	**(C) Golgi flow network edge e**		
edge path betweenness	$a_{e,\mathrm{path}}$	6.3.15	**property**	**symbol**	**Eq.**
edge flow betweenness	$a_{e,\mathrm{flow}}$	6.3.16	number	$g_{e,n}$	6.3.18
			wiggling	$g_{e,w}$	6.3.19
			intensity	$g_{e,i}$	6.3.20
			direction	$g_{e,d}$	6.3.21
			velocity	$g_{e,v}$	6.3.22
			combinations	e.g. $g_{e,d+v}$	-

Table 6.3.3: **List of studied (edge) properties of actin and Golgi flow networks. (A)** Edge properties of actin networks that were used to compare networks across conditions and time and edge properties used to predict organelle flow (cf. (C)). Some edge properties are local ($a_{e,E}$ to $a_{e,a}$) while some consider the role of the edge in the network context ($a_{e,\mathrm{deg}}$ to $a_{e,\mathrm{flow}}$). **(B)** Properties of actin network that were used for quantification of cytoskeletal phenotypes and assessment of transport efficiency. **(C)** Edge properties of Golgi flow network derived from Golgi tracking data, taking into account numbers, intensities, velocities and directions of Golgi as well as combinations of these properties.

enhancement of curvilinear structures unnecessary. Nevertheless, such filters [Sato et al., 1998; Obara et al., 2012a] may be readily applied in open contour-based approaches as well. Moreover, [Xu et al., 2015] employed manual segmentations and careful parameter gauging to optimize the network extraction quality and found $d_{\mathrm{HD}} \lesssim 1$ pixels which outperforms our approach, but benefited from higher SNR. Finally, while our presented approach, including the image preprocessing, involves only 6 parameters, SOAX employs more than 20 parameters in total.

Despite the differences in the image segmentation procedures, which have been designed for different image sources and systems of interest, systematic gauging (cf. Fig. 4.4.1) provides a means to adjust the free parameters of any of the above segmentation procedures. While automated segmentation methods are steadily improving, a quantitative comparison with manual, expert-driven segmentations remains crucial.

6.3.3 List of studied (edge) properties of actin and Golgi flow networks

After extracting the actin cytoskeletal networks from image data, we computed various seminal properties to quantify cytoskeletal phenotypes and evaluate transport efficiency (cf. Fig. 4.4.2). The actin cytoskeleton at a given time is represented by a weighted, undirected network $G = (\mathcal{N}, \mathcal{E})$ of $N = |\mathcal{N}|$ nodes with positions $x_n \in \mathbb{R}^2$, $n \in \mathcal{N}$, and $E = |\mathcal{E}|$ edges with weightings a_e, $e = (n,m) \equiv (m,n) \in \mathcal{E}$. We considered different weightings of the actin edges (Tab. 6.3.3A): The Euclidean length $a_{e,E}$ of an edge e is the Euclidean distance between its two nodes and its filament length $a_{e,F}$ is given by the arc length of the curvilinear filament

segment (cf. Materials and Methods in Section 4.4). In addition, we derived the filament bending as the ratio of filament and Euclidean length,

$$a_{e,B} \quad \equiv B = \quad a_{e,E}^{-1} a_{e,F}, \tag{6.3.1}$$

which is close to one for straight filaments (cf. definition of convolutedness in [Staiger et al., 2009]). The edge weight $a_{e,w}$ is the intensity of the image summed along the filament segment. The edge capacity,

$$a_{e,c} \quad = \quad a_{e,F}^{-1} a_{e,w}, \tag{6.3.2}$$

captures the average weight per unit length of the filament segment and reflects the average thickness of a filament segment. The edge length,

$$a_{e,l} \quad = \quad a_{e,c}^{-1}, \tag{6.3.3}$$

is given by the inverse capacity. Moreover, we quantified the alignment of filament segments by the angle between the respective edge and the cell axis unit vector x_{cell},

$$a_{e\equiv(n,m),a} \quad = \quad \arccos\left(\frac{(x_n - x_m)\, x_{\text{cell}}}{\|x_n - x_m\|\, \|x_{\text{cell}}\|}\right). \tag{6.3.4}$$

The organization of the extracted networks was quantified by various seminal network properties (Tab. 6.3.3B): To study the connectedness of a given network, we removed all edges with capacities below the 50th percentile and computed the average fraction of nodes per connected components,

$$S \quad = \quad (NI)^{-1} \sum_{i=1}^{I} N_i, \tag{6.3.5}$$

where I is the number of connected components and N_i, $i \in \{1, \ldots, I\}$, is the number of nodes in component i. As a measure of actin bundling across the cell, we computed the average edge capacity,

$$\mathrm{E}[a_c] \quad = \quad E^{-1} \sum_{e=1}^{E} a_{e,c}. \tag{6.3.6}$$

As a measure of heterogeneity of the actin distribution, we calculated the assortativity,

$$A \quad = \quad \frac{1}{2E} \sum_{n=1}^{N} \sum_{m=1}^{N} \left(w_{(n,m)} - \frac{d_n d_m}{2E}\right) d_n d_m, \tag{6.3.7}$$

which is positive if nodes are preferentially connected to nodes of similar degrees, hence detecting regions of spatially clustered bundling of the actin cytoskeletal.

The assessment of the actin networks' transport efficiency employed ensembles of randomized null model networks. The null model as well as the extracted networks may exhibit crossing edges (cf. Materials and

Methods in Section 4.4) and the frequency of such crossings was measured by the relative crossing number,

$$X = E^{-1} \sum_{\substack{e,f \in \mathscr{E} \\ e \neq f}} e \times f, \tag{6.3.8}$$

where $e \times f = 1$ if edges e and f cross and 0 otherwise. Moreover, the average shortest path length was computed as a standard measure of transport efficiency,

$$\mathrm{E}[L] = \frac{1}{2N(N-1)} \sum_{n=1}^{N} \sum_{\substack{m=1 \\ m>n}}^{N} L_{n,m}, \tag{6.3.9}$$

where $L_{n,m} = \min_{\mathscr{P} \in \mathscr{P}_{n,m}} \sum_{e \in \mathscr{P}} a_{e,c}^{-1}$ and $\mathscr{P}_{n,m}$ is the set of all paths from node n to m [West, 2001]. The dispersal of the transport efficiency was measured by the CV of the shortest path lengths,

$$\mathrm{CV}[L] = \frac{\mathrm{SD}[L]}{\mathrm{E}[L]}, \tag{6.3.10}$$

where SD denotes the standard deviation. The robustness of the transportation network against disruptions was evaluated by the algebraic connectivity,

$$C = \lambda_2, \tag{6.3.11}$$

which is the second smallest eigenvalue of the graph Laplacian $\mathscr{L}_G = D_G - A_G$ and is zero for a disconnected network [West, 2001]. Here, D_G is the diagonal matrix of the capacity-weighted node degrees and A_G is the capacity-weighted connectivity matrix of G. Finally, the alignment of filament segments in the cell was measured using the CV of edge angles,

$$\mathrm{CV}[a_a] = \frac{\mathrm{SD}[a_a]}{\mathrm{E}[a_a]}. \tag{6.3.12}$$

To correlate the flow of organelles with the structure of the actin network, we derived two networks with identical structure and different edge weightings that we referred to as actin and Golgi flow networks, respectively (cf. Fig. 4.4.4; Tabs. 6.3.3A and C). The studied (local) edge properties of the actin network were given by the edge weight $a_{e,w}$ and capacity $a_{e,c}$. Moreover, we investigated the (global) importance of a given edge in the actin network context using several betweenness measures. First, we introduced the edge degree as a semi-local measure of actin organization,

$$a_{e,\mathrm{deg}} = \sum_{\substack{f \in \mathscr{E} \\ e \sim f}} a_{f,c}, \tag{6.3.13}$$

where $e \sim f$ indicates adjacent edges. This measure may be conveniently computed in the line graph LG of

G, whereby each node in LG represents an edge in G and two nodes in LG are adjacent if the corresponding edge share a node in G [West, 2001]. In the line graph LG, the edge degree $a_{e,\mathrm{deg}}$ simply corresponds to the capacity-weighted degree of the node e. Analogously, we computed another famous property of an edge e in the line graph LG with capacity-weighted connectivity matrix A_{LG}, i.e. its page rank

$$a_{e,\mathrm{rank}} = \alpha \sum_{f \in \mathscr{E}} A_{LG,e,f} D^{-1}_{LG,e,f} a_{f,\mathrm{rank}}, \tag{6.3.14}$$

where D_{LG} is the diagonal matrix of capacity-weighted node degrees of the line graph and classically $\alpha = 0.85$ [Brin and Page, 1998; Langville and Meyer, 2005]. Moreover, we computed the shortest path edge betweenness which counts the number of shortest path between any two nodes in the network that pass through a given edge,

$$a_{e,\mathrm{path}} = \sum_{\substack{n,m \in \mathscr{N} \\ n \neq m}} \begin{cases} 1 & \text{, if } e \in \mathscr{P}^{\mathrm{shortest}}_{n,m} \\ 0 & \text{, otherwise,} \end{cases} \tag{6.3.15}$$

where $\mathscr{P}^{\mathrm{shortest}}_{n,m} = \mathrm{argmin}_{\mathscr{P} \in \mathscr{P}_{n,m}} \sum_{e \in \mathscr{P}} a^{-1}_{e,c}$. Similarly, we computed the maximum current flow edge betweenness which is the maximum flow across an edge summed over all pairs of nodes that are treated as unit sinks and sources of the flow, respectively,

$$a_{e,\mathrm{flow}} = \sum_{\substack{n,m \in \mathscr{N} \\ n \neq m}} \mathscr{F}_{n,m,e}, \tag{6.3.16}$$

where $\mathscr{F}_{n,m,e}$ is the maximum flow between nodes n and m across edge e.

The studied edge properties of the introduced Golgi flow network considered the movement of Golgi close to an edge (Tab. 6.3.3C). At a given time, the center points Golgi track segments are given by $y_{t \equiv (u,v)} = 2^{-1}(x_u + x_v)$, $t \in \{1, \ldots, T\}$, where y_u and y_v denote the endpoints of the track segment $t \equiv (u, v)$. The minimum distance between the track segments center point y_t and an edge e was computed as

$$d_{t,e \equiv (n,m)} = \min_{\lambda \in [0,1]} \left\| y_t - (x_n + \lambda (x_m - x_n)) \right\|. \tag{6.3.17}$$

Then, as a first simple measure of Golgi flow, we computed the number of close-by Golgi,

$$g_{e,n} = \sum_{t=1}^{T} \Theta\left(d_{\mathrm{thres}} - d_{t,e}\right), \tag{6.3.18}$$

where $d_{\mathrm{thres}} = 20\,\mathrm{pixels}$ throughout the manuscript. To measure the wiggling behavior of Golgi, we com-

puted the angle between consecutive track segments,

$$g_{e,w} = \sum_{t=1}^{T} \Theta\left(d_{\text{thres}} - d_{t,e}\right) \omega_t, \tag{6.3.19}$$

where $\omega_{t\equiv(u,v)} = \arccos\left(\frac{(y_u - y_v)\left(y_{u'} - y_{v'}\right)}{\|y_u - y_v\|\left\|y_{u'} - y_{v'}\right\|}\right)$ and $y_{u'}$ and $y_{v'}$ are the endpoint of the track segment in the next time step. If there was no next track segment in the track, we set $\omega_{t\equiv(u,v)} = 0$. Moreover, the average intensity ι_t of the Golgi was used a measure for their size (Appendix 6.3.5),

$$g_{e,i} = g_{e,n}^{-1} \sum_{t=1}^{T} \Theta\left(d_{\text{thres}} - d_{t,e}\right) \iota_t. \tag{6.3.20}$$

The direction of the Golgi track segments with respect to the respective actin edges was computed as

$$g_{e,d} = g_{e,n}^{-1} \sum_{t=1}^{T} \Theta\left(d_{\text{thres}} - d_{t,e}\right) \delta_{t,e}, \tag{6.3.21}$$

where $\delta_{t\equiv(u,v),e\equiv(n,m)} = \frac{(x_n - x_m)(y_u - y_v)}{\|x_n - x_m\|\|y_u - y_v\|}$. Finally, the average Golgi velocity in the vicinity of an edge was calculated as

$$g_{e,v} = g_{e,n}^{-1} \sum_{t=1}^{T} \Theta\left(d_{\text{thres}} - d_{t,e}\right) \nu_t, \tag{6.3.22}$$

where $\nu_t = \|y_t\|$.

6.3.4 Inference of network design principles using two different null models and negative controls

We demonstrated that the actin cytoskeleton displays network properties supportive of efficient transport processes (cf. Fig. 4.4.2). In particular, by proposing suitable null models, we showed that this transport efficiency arises from the specific organization of the cytoskeleton in biological cells, hence indicating an evolutionary basis. Here, we discuss and justify our proposed null models in more detail.

The first null model randomly and uniformly distributes the node positions across the cell and assigns the edges to new, randomly chosen pairs of nodes whose distance matches the Euclidean length of the respective edge, while keeping the number of edges crossings low (cf. Figs. 4.4.2H-J and Materials and Methods in Section 4.4). This procedure is a modified version of the Erdős–Rényi model with hidden variables that has been used to identify structural features of various real world networks [Caldarelli et al., 2002; Hayashi, 2006]. In our version, the hidden node variables are given by their positions and the probability of adding an edge between two nodes depends on their Euclidean distance and the fraction of edges from the original network of identical (binned) length that have not already been added to the null model network. Moreover, as an extension to the hidden variable model, for each added edge, we tried 50 possible pairs of nodes and assigned the edge such that the number of intersections with already added edges was smallest.

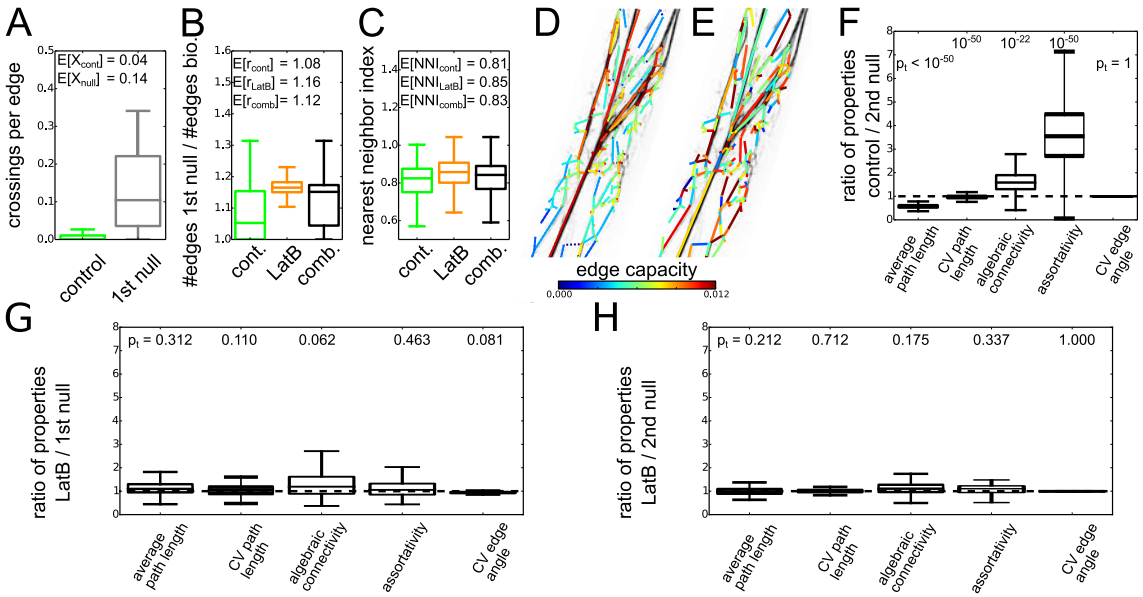


Figure 6.3.3: **Assessment of biological significance of actin network properties using two null models.** Differences between cytoskeletal networks from control cells (green), cells treated with LatB (orange) and randomized null model networks (gray). **(A)** Boxplot of the number of crossings per edge in the original actin networks with $\mathrm{E}[X] = 0.04 \pm 0.02$ and network derived from the first null model with $\mathrm{E}[X] = 0.14 \pm 0.11$. **(B)** Boxplot of the ratio r of the number of edges in the extracted actin networks and the corresponding first null model networks. By construction, the ratio is larger than 1 and the number of edges in the first null model networks was increased by $\mathrm{E}[r] = 1.12 \pm 0.07$ on average. **(C)** Boxplot of the nearest neighbor index NNI which captures the spatial distribution of node positions in the extracted actin networks. Values below 1 indicate a stronger spatial clustering than expected from a random and uniform distribution of node positions and the observed nearest neighbor index was $\mathrm{E}[\mathrm{NNI}] = 0.83 \pm 0.09$ on average for the extracted networks. **(D)** Example of an original extracted actin network with edge colors reflecting the edge capacities. **(E)** Example of network obtained from our second null model, i.e. by shuffling of the edge properties (cf. (C)). The structure of the randomized network is identical to that of the original network. **(F)** Boxplot of the ratios of various network properties of the extracted networks and an ensemble of networks obtained from the second null model. Again, the significance of the differences when considering the second null model remained unchanged when compared to the differences found from employing the first null model (independent two-sample t-tests p-values $p_t < 0.05$ were considered significant; cf. 4.4.2I). **(G)** Analysis of biological significance of various network properties of the reconstructed actin network for LatB-treated cells using the first null model. In contrast to the cytoskeletal networks of the untreated control cells, no significant differences were observed (all $p_t \geq 0.05$; cf. 4.4.2I). **(H)** Analysis of various network properties using the second null model showed no biological significance of cytoskeletal network properties in LatB-treated cells (all $p_t \geq 0.05$; cf. also 4.4.2I).

This null model preserves the distribution of edge capacities, whose sum reflects the amount of filamentous actin in the cell. The null model further maintains the length distribution of filament segments. Although the generated null model networks are in general not planar, their average relative crossing number is small, with $\mathrm{E}[X] = 0.14 \pm 0.11$ crossings per edge (Fig. 6.3.3A). We note that while the extracted actin skeleton is planar by construction (all filament crossing are nodes; cf. Fig. 4.4.1D), the reconstructed actin network may exhibit crossing edges since the edges are modeled as straight lines between the nodes, so that $\mathrm{E}[X] = 0.04 \pm 0.02$ already for the original networks. Moreover, since the null model networks are generally not connected we added edges of minimum total Euclidean length to connect the network. Yet, across all generated null model networks, the number of edges E_{null} of the null model networks exceeds the number of edges $E_{\mathrm{bio.}}$ of the corresponding extracted networks only by a factor $\mathrm{E}[r] = 1.12 \pm 0.07$ (Fig. 6.3.3B; cf. Materials and Methods in Section 4.4), where

$$r = \frac{E_{\mathrm{null}}}{E_{\mathrm{bio.}}}. \tag{6.3.23}$$

Furthermore, in contrast to the nodes in the null model networks, the nodes of the extracted actin networks were typically not randomly and uniformly distributed across the cell. We quantified the spatial distribution of node positions using the nearest neighbor index (NNI; [Cover and Hart, 1967]),

$$\mathrm{NNI} = \frac{\frac{1}{N^2(N-1)} \sum_{n,m=1}^{N} d_{n,m}}{\frac{1}{2}\left(\frac{A_{\mathrm{cell}}}{N}\right)^{1/2}}, \tag{6.3.24}$$

where $d_{n,m} = \|x_n - m_m\|$ is the Euclidean distance between nodes n and m, $n, m \in \mathcal{N}$, $N = |\mathcal{N}|$, and A_{cell} is the area of the cellular region of interest. Averaged across all extracted networks, we found $\mathrm{E}[\mathrm{NNI}] = 0.83 \pm 0.09$ (Fig. 6.3.3C). This NNI below 1 indicates a stronger spatial clustering of nodes than expected by chance. This clustering is consistent with the visible actin-free regions in the cells (cf. Fig. 4.4.2A). In addition, the extraction procedure does not allow nodes at neighboring pixels, violating the assumption of a uniform distribution of nodes (cf. Fig. 4.4.1 and Materials and Methods in Section 4.4).

To overcome these shortcomings, we introduced a second, more restricted null model in which all node positions of the original extracted network (Fig. 6.3.3D) were kept and only edge properties were shuffled (Fig. 6.3.3D). This procedure has been previously used in [Breuer et al., 2014; Breuer and Nikoloski, 2014] with a similar purpose, i.e. to study the biological relevance of properties of a grid-approximation of the cytoskeleton. The null model preserves the total amount of filamentous actin and further leaves the connectedness and planarity of the original network unchanged. All investigated properties of the extracted networks showed the same significant differences to the second null model as to the first null model (Fig. 6.3.3F; cf. Fig. 4.4.2I; independent two-sample t-tests p-values $p_t < 0.05$ were considered significant). Namely, the average path length and the CV of the path lengths were smaller than expected by chance, while the algebraic connectivity and the assortativity were larger then expected by chance. By construction, the distribution of edge angles in the second null model networks was identical to that of the original networks, yielding a unit ratio of the CV of edge angles. In conclusion, our two different null models yielded consistent results and were able to capture biologically relevant signals.

Moreover, as a negative control, we assessed the biological relevance of transport-related properties of cy-

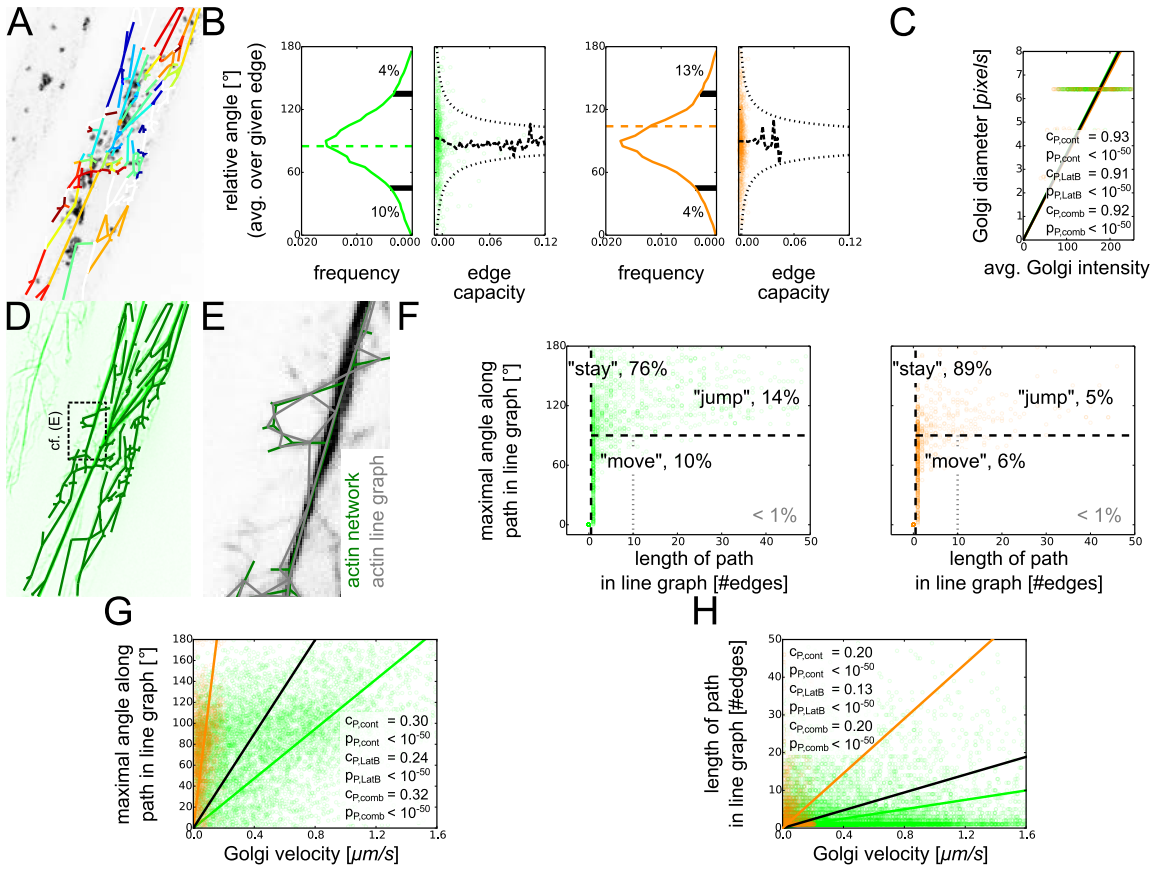


Figure 6.3.4: **Additional analyses of Golgi movement and wiggling in dependence of actin cytoskeleton.** Results for cells of untreated control plants (green), cells of plants treated with the actin-disrupting drug LatB (orange) and the combined data (black). **(A)** Overlay of cellular recording of Golgi and Golgi flow network with edge colors representing the relative angle of close-by Golgi tracks, capturing the wiggling behavior of Golgi. **(B)** Caption continued on next page.

toskeletal networks extracted from recordings of LatB-treated cells (Figs. 6.3.3G and H; cf. Fig. 4.4.2). Indeed, none of the studied properties of the cytoskeletal networks of the LatB-treated cells showed significant differences for either of the two null models (all $p_t \geq 0.05$). This absence of deviations from transport-related properties of randomized networks in chemically perturbed cells further supports our claim that the employed null models are suitable to uncover biological design principles.

6.3.5 Golgi wiggling in dependence of actin cytoskeleton and frequencies of Golgi switching between filaments

In our quantitative analysis of cellular transport dynamics, we combined automated tracking data of Golgi with automated extraction of actin cytoskeletal networks (cf. Fig. 4.4.3). Here, we discuss two of these analyses in more detail, the investigation of Golgi wiggling in dependence of actin structures and the movement of Golgi along and between filaments.

(B) Distributions of average relative angles of close-by Golgi tracks per edge in dependence of capacity of the respective edge for control and LatB-treated cells. The average relative angles approached 90° with increasing edge capacities (right panel; dashed lines indicate averages and dotted lines are discussed below). The distributions of average relative angles exhibited peaks close to 85° and 104° for control and LatB-treated cells, respectively (left panel; dashed lines, cf. Fig. 4.4.3G). For the control cells, 10% of the edges showed average relative angles below 45°, reflecting predominant directed movement of close-by Golgi and 4% showed average relative angles above 135°, reflecting predominant wiggling behavior of close-by Golgi. For LatB-treated cells, these fractions were roughly reversed with 4% and 13%, respectively. **(C)** Scatter plot of Golgi diameters and average intensities estimated from the Golgi detection step (cf. Materials and Methods in Section 4.4) showed a strong positive correlation (Pearson correlation coefficients $c_P \approx 0.9$ and all p-values $p_P < 0.05$). **(D)** Overlay of cellular recording of actin cytoskeleton from control cell (green) and extracted actin network (dark green). **(E)** Section of overlay of an extracted actin network (dark green; cf. (D)) and its line graph (gray) which has a node for each edge in the original network and a link between nodes that represent adjacent edges. **(F)** Scatter plots of the length of the shortest path connecting the two edges which are closest to a given track segment's start- and end-points and the associated maximal angle along the shortest path. Golgi bodies that did not move to a different edge were assigned to the "stay" class. Track segments were assigned to the "move" or "jump" class if they moved to a different edge while the corresponding shortest path between the two edges was below or above 90°, respectively. **(G)** Scatter plot of the length of the shortest path connecting the two edges which are closest to a given track segment's start- and end-points and the associated Golgi velocity shows positive correlations for control and LatB-treated cells as well as the combined data ($c_P \approx 0.3$ and all $p_P < 0.05$). **(H)** Scatter plot of the maximal angle along the shortest path connecting the two edges which are closest to a given track segment's start- and end-points and the associated Golgi velocity shows moderate correlations for control and LatB-treated cells as well as the combined data ($c_P \lesssim 0.2$ and all $p_P < 0.05$).

We confirmed that the Golgi wiggling behavior is not Golgi-specific [Akkerman et al., 2011] and showed further that the prevalence of Golgi wiggling behavior is stationary over the course of the recording period and does not depend on the distance from the actin cytoskeleton (cf. Figs. 4.4.3H-J). However, the actin cytoskeleton is composed of filaments and bundles of varying thickness and it has been suggested that arrays of fine actin filaments promote wiggling [Akkerman et al., 2011]. To test this hypothesis in detail, we employed our extracted, weighted network representation of the actin cytoskeleton. We constructed a Golgi flow network in which the edges of the actin network were assigned a measure of Golgi wiggling (Fig. 6.3.4A, cf. Materials and Methods in Section 4.4). To this end, for each edge, we considered all Golgi in a vicinity of 20 pixels of the edge and computed their average relative angle of movement (cf. e.g. Fig. 4.4.3H), where angles above 90° were considered as wiggling behavior. For both control and LatB-treated cells, this average relative angle showed no significant correlation with the capacities of the actin edges, i.e. their average thickness (Fig. 6.3.4B). In particular, the average relative angles approached 90° for larger edge capacities which may be explained by the larger number of Golgi close to thick edges (cf. Fig. 4.4.4L; linear regression yielded a slope of $s = 391$). Therefore, assuming that the relative angle were distributed randomly across Golgi tracks, the standard deviation of average relative angle decreased and the expected average relative angle approached 90° (cf. dotted black lines). These findings suggest that the thickness of close-by actin bundles does not influence the Golgi wiggling behavior. However, this is still compatible the observation that arrays of fine actin filament increase Golgi wiggling [Akkerman et al., 2011] since the thickness of individual filaments studied here does not capture the surrounding actin environment. We note that Golgi size is strongly correlated with the average Golgi intensity (Fig. 6.3.4C; Pearson correlation coefficients $c_P \approx 0.9$ and p-values $p_P < 0.05$) so that Golgi intensity may be used as an estimator of Golgi size (cf. Fig. 4.4.4).

We further quantified the movement of Golgi along and between filaments and distinguished three classes (cf. Fig. 4.4.3M). To this end, we assigned the start- and end-points of each track segment to their nearest edge in the actin network that we refer to as start- and end-edge, respectively. From the original actin network (cf. Fig. 6.3.4D, green), we constructed the line graph (cf. Fig. 6.3.4E, gray), i.e. a graph which has a node for each edge in the original network and a link between nodes that represent adjacent edges (cf. [Hemminger and Beineke, 1978; West, 2001] and Appendix 6.3.3). We computed the shortest paths from the start- to the end-edges through the line graph of the actin network. For a given shortest path, we calculated the path length and the maximal angle between any two adjacent edges along the path. We then classified different types of Golgi movement (Fig. 6.3.4F; the classes are referred to as "stay", "move" and "jump"), depending on the minimum number of edges traversed by the Golgi and the associated maximal angle between traversed edges of the cytoskeletal network. The frequencies of Golgi in these different classes showed stationary Golgi, Golgi likely moving along a given filament, as well as Golgi switching between different, non-adjacent filaments (cf. Fig. 4.4.3M). A closer inspection further showed that less than 1% of Golgi track segments in the "move"-class traversed more than 10 edges of the actin network. This supports the assumption that the most Golgi in this class did not switch to a different filament but moved along a single filament.

Next, we investigated the relationship between Golgi velocity and redistribution across AFs (cf. Fig. 6.3.4F). Both, the maximal angle along the shortest path from the start- to the end-edge and its path length were moderately correlated with the velocity of the respective Golgi (Figs. 6.3.4G and H; $c_P \approx 0.3$ and $c_P \lesssim 0.2$, respectively, and all $p_P < 0.05$). This indicates that faster Golgi are more likely to move to a different AF.

6.3.6 Directionality of actin edges and correlations of actin edge properties

Our analyses of the transport capacity of the actin cytoskeleton rely on the assumption of undirected edges, i.e. edges that allow bi-directional transport (cf. Fig. 4.4.2). To elucidate the biological plausibility of this assumption, were constructed an additional type of Golgi flow networks (cf. Fig. 4.4.4) by weighting the edges according to the average angle between the respective edge and the close-by Golgi track segments (Fig. 6.3.5A). For an edge that allows predominantly uni-directional transport, this average angle is expected to be below 45° or above 135°. However, only around 15% of the edges showed such predominantly uni-directional transport (Fig. 6.3.5B, left) and we found no correlation between the uni-directionality of transport of an edge and its capacity, i.e. its thickness (Fig. 6.3.5B, right).

When investigating the flow of Golgi along the actin cytoskeleton, we considered several local and global edge properties of the actin network as regressors or predictors of the Golgi flow (cf. Fig. 4.4.4 and Appendix 6.3.3). Across the studied networks, there were on average $E = 218 \pm 52$ edges (dependent variables) whose flow properties were predicted using sets of 5 different actin edge properties (predictors). Since the number of predictors was much smaller than the number of dependent variables, overfitting was not an issue. Moreover, the reliability of the multiple linear regression results might be affected by colinearity of the predictors which may be quantified by the condition number n. Across the studied network, the median of the condition number $\mathrm{Md}[n] = 16.93 \pm 4.36$ was below 20, suggesting that our findings were not severely affected by colinearities [Belsley, 2006]. Here, the median was used to exclude the effect of outliers and the confidence interval was accordingly determined by the median absolute deviation $\mathrm{Mad}[n] = \mathrm{Md}[|n - \mathrm{Md}[n]|]$.

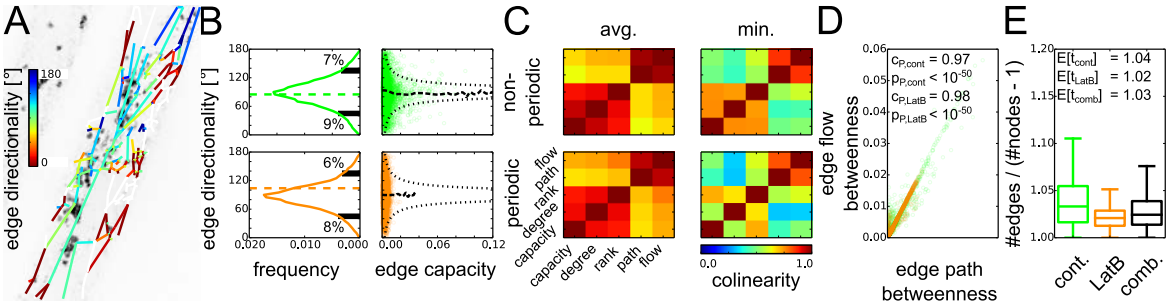


Figure 6.3.5: **Directionality of Golgi movement along edges in actin networks and correlations of actin edge properties.** Results for cells of untreated control plants (green), cells of plants treated with the actin-disrupting drug LatB (orange) and the combined data (black). **(A)** Overlay of cellular recording of Golgi and Golgi flow network with edge colors representing edge directionality, capturing the average angle between network edge and close-by Golgi track segments. **(B)** Distribution of edge directionality in dependence of capacity of the respective edge for control (upper panels) and LatB-treated cells (lower panels). The edge directionality was not correlated with the edge capacity, i.e. thicker edges did not imply stronger uni-directionality of transport. For the control cells, 16% of the edges showed predominantly uni-directional transport with angles below 45° or above 135°. For the LatB-treated cells, 14% of the edges showed predominantly uni-directional transport. Refer to Fig. 6.3.4B for detailed explanation of the panels. **(C)** Heat maps of the colinearity between the studied actin edge properties for the original, non-periodic networks (upper panels) and the periodic networks (lower panels). The colinearities were averaged across all time points of all control cells (left panels) or the minimum value was taken (right panels). In all cases, the edge path and flow betweenness were strongly correlated. The edge capacity, degree and rank were mutually correlated on average but the minimum colinearity revealed lower colinearity, especially for the periodic networks. **(D)** Scatter plot of edge flow betweenness versus edge path betweenness showed strong correlations with Pearson correlation coefficients $c_P = 0.97$ and $c_P = 0.98$ for control and LatB-treated cells, respectively. **(E)** Boxplot of the ratio t of the number of edges in the extracted actin networks and the corresponding spanning trees (the spanning tree has, by definition, $N-1$ edges). Since the original actin networks are connected by construction, the ratio is larger than 1 and there were on average $\mathrm{E}[t] = 1.03 \pm 0.02$ times more edges in the actin networks than in the corresponding spanning trees.

Analyzing the colinearity of the predictors in more detail, we computed the colinearity of any two actin edge properties x and $x^{'}$ (Fig. 6.3.5C), via

$$C_{x,x'} = \frac{\sum_{e=1}^{E} a_{e,x} a_{e,x'}}{\left(\sum_{e=1}^{E} a_{e,x} a_{e,x}\right)^{\frac{1}{2}} \left(\sum_{e=1}^{E} a_{e,x'} a_{e,x'}\right)^{\frac{1}{2}}}. \tag{6.3.25}$$

Some of the studied actin edge properties were strongly correlated, such as the edge path betweenness and the edge flow betweenness (Fig. 6.3.5D). Although these two properties generally measure different aspects of the importance of an edge in the network context, they are identical for tree-like networks of unit edge capacities and lengths [Brandes and Fleischer, 2005]. We therefore quantified how tree-like our studied actin networks were by computing the ratio of their number of edges E and the number of edges $N-1$ in the corresponding spanning trees (Fig. 6.3.5E),

$$t = \frac{E}{N-1}. \tag{6.3.26}$$

Indeed, there were on average only $\mathrm{E}[t] = 1.03 \pm 0.02$ times more edges in the actin networks than in the corresponding spanning trees, thus explaining the strong correlation between the edge path and flow betweennesses.

6.4 Appendix: DeFiNe: an optimization-based method for robust disentangling of filamentous networks

6.4.1 Mathematical formulation of the filament cover problem

The structure of a filamentous network is described by a weighted geometric graph $G = (\mathcal{N}, \mathcal{E})$ with $N = |\mathcal{N}|$ nodes and $E = |\mathcal{E}|$ undirected, weighted edges. Edges represent filament segments and nodes represent their endpoints. The positions of the nodes are v_n, $n \in \mathcal{N}$, whereby, typically, $v_n \in \mathbb{R}^2$ or $v_n \in \mathbb{R}^3$ for networks extracted from image data. We focus on geometric networks because filaments are embedded in space, but our approach is readily applicable to non-geometric graphs. The edge weights are w_e, $e := (n_0, n_1) \in \mathcal{E}$ and $n_0, n_1 \in \mathcal{N}$.

To decompose the graph G into individual filaments it is natural to decompose it into paths, i.e., to solve a path cover problem (PCP). The PCP has been intensively studied on different types of graphs and with various restrictions (e.g. [Rao and C., 1990; Andreatta and Mason, 1995; Lin et al., 1995; Pak-Ken, 1999; Lin et al., 2006; Brešar et al., 2011]). There are several potential routes (cf. [Andreatta and Mason, 1995] for an overview of the PCP for testing printed circuits): (1) We may either use node- or edge-paths, where a path $p = \left(a_{p,1}, \ldots, a_{p,P}\right)$ is an ordered sequence of $P = \left|p\right|$ pairwise adjacent nodes ($a \in \mathcal{N}$) or edges ($a \in \mathcal{E}$), respectively, and $a_{p,i}$ denotes the i-th node or edge of filament p. (2) The paths may be either node-disjoint, edge-disjoint, or unrestricted. (3) The objective of the PCP may be either to obtain a cover of minimum cardinality or minimum weight.

For our purpose, the decomposition of a filamentous network into individual smooth filaments, it seems reasonable to look for an edge-path cover where each edge is covered by (at least) one path and the total (or average) roughness is minimized. Edges that are covered by more than one path naturally correspond to filament overlaps. The minimization of the average instead of the total roughness favors shorter paths which may be appropriate for some networks.

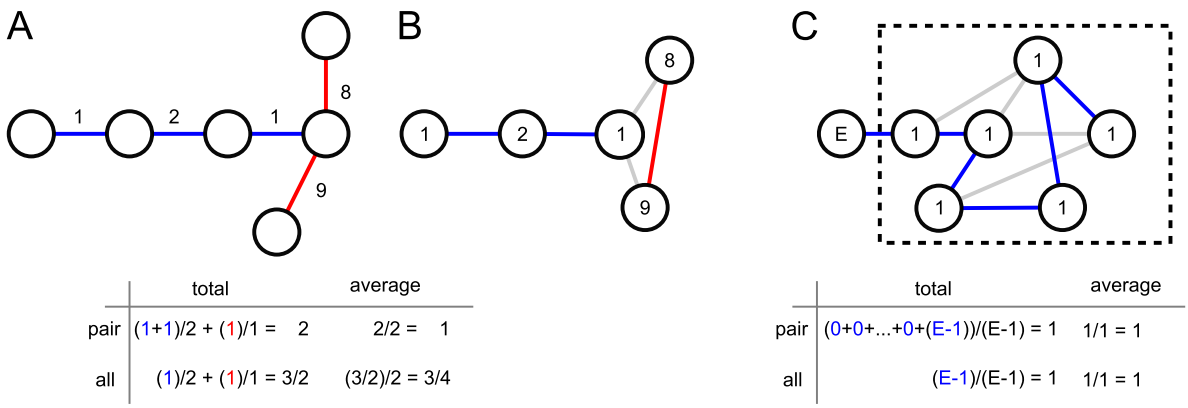
To define our filament cover problem (FCP) more rigorously, we introduce the roughness r_p of path p and the set $\mathcal{P}$ of all paths in G:

> Given a set $\mathcal{E}$ of edges and a set $\mathcal{P}$ of paths with roughnesses r_p, $p \in \mathcal{P}$:
>
> Find a subset $\mathcal{P}_{\text{fil}} \subseteq \mathcal{P}$ with minimal total (or average) roughness R such that each element in $\mathcal{E}$ is covered (at least) once.

The roughness measure r_p of a path p can be chosen arbitrarily and may involve, e.g., the edge weights or the edge alignments. An intuitive choice is the pairwise filament roughness of p (cf. Eq. 4.5.1),

$$r_{p,\text{pair}} = \begin{cases} (P-1)^{-1} \sum_{i=1}^{P-1} \left| w_{e_{p,i+1}} - w_{e_{p,i}} \right| & , P > 1 \\ w_{e_{p,1}} & , P = 1 \end{cases}, \tag{6.4.1}$$

where $w_{e_{p,i}}$ denotes the weight of the i-th edge in filament p. The pairwise filament roughness is the average absolute value of the difference between weights of adjacent edges. It reflects the consistency of the edge



	total	average
pair	(1+1)/2 + (1)/1 = 2	2/2 = 1
all	(1)/2 + (1)/1 = 3/2	(3/2)/2 = 3/4

	total	average
pair	(0+0+...+0+(E-1))/(E-1) = 1	1/1 = 1
all	(E-1)/(E-1) = 1	1/1 = 1

Figure 6.4.1: **Proof of NP-hardness of the filament cover problem. (A)** Optimal filament cover of an exemplary (edge-weighted) graph. Table with cover roughnesses R for minimization of total or average roughness and pairwise or all-to-all filament roughness measure, respectively. **(B)** Corresponding (node-weighted) line graph with equivalent path cover and the same roughness results as for the (edge-weighted) graph in (A). **(C)** Extension of an arbitrary graph with node weights 1 by a node of weight E. Here, finding a (node-weighted) path cover of roughness $R = (0+0+\cdots+(E-1))\,/\,(E-1) = 1$ or less is equivalent to finding a Hamiltonian path. This equivalence holds for covers minimizing the total or average roughness of the cover and using the pairwise or all-to-all filament roughness measure (Eqs. 6.4.1 and 6.4.2), see table.

weights along a filament which is typically smaller within than across filaments (but cf. Discussion). Moreover, if the path consists of a single edge we take its weight as a roughness measure. This choice increases the flexibility of the obtainable filament covers and is necessary to avoid a cover by only individual edges which contribute zero weight when weighted only according the first line in Eq. 6.4.1. Another measures for the quality of a filament is the all-to-all filament roughness (cf. Eq. 4.5.2)

$$r_{p,\mathrm{all}} = \begin{cases} (P-1)^{-1} \max\limits_{i,j\in\{1,\ldots,P\}} \left| w_{e_{p,i}} - w_{e_{p,j}} \right| & , P > 1 \\ w_{e_{p,1}} & , P = 1 \end{cases}, \tag{6.4.2}$$

which is the average maximal difference between any edge weights in a path p, and again the original weight of the edge is used for a path of length one. Taking into account that most filaments are only moderately bent, we may further wish to minimize the maximal filament deflection angle between adjacent edges of a path p (cf. Eq. 4.5.3),

$$r_{p,\mathrm{angle}} = \max_{i\in\{1,\ldots,P-1\}} \mathrm{angle}\left(v_{e_{p,i+1,1}} - v_{e_{p,i+1,0}}, v_{e_{p,i,1}} - v_{e_{p,i,0}} \right) \tag{6.4.3}$$

where $v_{e_{p,i,0}}$ and $v_{e_{p,i,1}}$ denote the positions of the start and end nodes of the i-th edge of filament p, respectively. Moreover, $\mathrm{angle}\left(v, v'\right) := \arccos\left(\frac{v\cdot v'}{\sqrt{v\cdot v}\sqrt{v'\cdot v'}} \right)$ is the Euclidean angle of two vectors v and v' and $r_{p,\mathrm{angle}} = 0°$ corresponds to perfectly straight alignment.

6.4.2 Computational intractability of the filament cover problem

The FCP is difficult to solve. This is intuitively clear as the number of paths (let alone the number of path covers) increases rapidly with the number of nodes N. Even in planar graphs, the number of closed paths visiting each node once was shown to increase at least exponentially with N [Buchin et al., 2007; Biswas et al., 2012]. We show now that the FCP is NP-hard, even for planar, cubic graphs. Planar graphs can be drawn on a plane without crossing edges. They are of particular relevance since graphs that are generated from two-dimensional image data are planar by construction [Baumgarten and Hauser, 2012; Obara et al., 2012b]. Cubic graphs have only nodes of degree three. A proof of NP-hardness of a problem for planar, cubic graphs directly implies its NP-hardness on general graphs. The basic idea of a typical proof of computational complexity is as follows [Garey and Johnson, 1979]: A problem of known complexity is selected. By providing a constructive transformation or reduction, a bijection between the known problem and the problem in question is established, i.e., any yes-instance of the decision-version of the known problem is mapped to a yes-instance of the decision-version of the problem of interest and analogously for the no-instances. This reduction proves that the two problems fall into the same class of computational complexity. Our proof is by reduction from the Hamiltonian path problem (HPP) on planar, cubic graphs which is known to be NP-complete [Garey et al., 1976]. The HPP asks, for a given graph, whether there is a node-path which visits each node exactly once.

First, we note that finding a filament cover on an edge-weighted graph G is equivalent to finding a node-path cover on its node-weighted line graph $L(G)$ (Fig. 6.4.1A and B). The line graph $L(G)$ of a graph G has a node of weight w_e for each edge e in G and edges connecting two nodes if the corresponding edges share a node in G.

Second, for a given line graph $L(G)$, we construct a graph such that finding a node-path cover of weight 1 or less is equivalent to solving the HPP. To that end, we add one edge with a terminal node to the line graph and set all original node-weights to 1 and the new node-weight to E (Fig. 6.4.1C). Then, only a Hamiltonian path ensures a minimal weight of $R = C^{-A}\sum_{i=1}^{C} r_{p_i} = \frac{1}{1}\frac{(0+\cdots+(E-1))}{(E-1)} = 1$, for both pairwise and all-to-all filament roughness $r_p = \left\{r_{p,\text{pair}}, r_{p,\text{all}}\right\}$ (cf. Eqs. 6.4.1 and 6.4.2) and both minimization of total and average filament roughness, i.e., $A \in \{0,1\}$.

Finally, we show that finding a Hamiltonian path on a line graph of a planar, cubic graph is NP-complete. It was shown that the HPP is NP-complete on general line graphs via a reduction from the HPP in cubic graphs [Bertossi, 1981]. This reduction remains valid when planar, cubic graphs are used instead of cubic graphs, for which NP-completeness of HPP is known [Garey et al., 1976]. Therefore, the decision version of the FCP is NP-complete and the FCP is NP-hard, as claimed. Since the FCP is NP-hard on planar, cubic graphs, it is (at least) NP-hard on general graphs.

6.4.3 The filament cover problem on trees is solvable in polynomial time

While we showed that the FCP is NP-hard on general and even planar, cubic graphs, it is solvable in polynomial time on trees. The polynomial algorithm outlined here is similar to those proposed to find an unrestricted node-path cover where each vertex may be included in multiple paths of minimum cardinality or minimum weight [Lin et al., 2006].

The basic idea is to assume that a certain path covering a certain edge is in the cover (in a tree, there are at most $N(N-1)/2 = \mathcal{O}\left(N^2\right)$ paths to choose from). Upon removal, the tree is split into potentially multiple forests (at most $\mathcal{O}(N)$), each tree of which is decomposed in the same way. The procedure is repeated for each edge (clearly $\mathcal{O}(N)$ in a tree). Thus, this results in a dynamic programming algorithm which has an overall polynomial time complexity of $\mathcal{O}\left(N^4\right)$.

The above procedure assumes non-overlapping paths and may be extended to limitedly overlapping paths. For the completely unrestricted case, there would be $\mathcal{O}\left(2^{\#\text{paths}}\right) = \mathcal{O}\left(2^{N^2}\right)$ combinations for covering a given edge to chose from in the first step, and the time complexity of the algorithm would be exponential. However, the problem remains polynomial if we allow only k-fold overlaps, $k = \mathcal{O}(1)$, i.e., each edge may be covered by at most k paths. In the first step of the above algorithm, a given edge may then be covered by at most $\mathcal{O}\left(\begin{pmatrix} N(N-1)/2 \\ k \end{pmatrix}\right) = \mathcal{O}\left(N^{2k}\right)$ edges and consequently the time complexity of the full algorithm is $\mathcal{O}\left(N^{2k+2}\right)$.

6.4.4 Approximation algorithm for the filament cover problem

Since the FCP is NP-hard even on planar, cubic graphs, we need suitable approximation algorithms. In particular, the approximation algorithms should allow overlapping filaments as well as looped filaments. A natural choice seems to be the formulation of the FCP as a set cover problem (SCP) [Karp, 1972]:

> Given an object set $\mathcal{U}$, called universe, and a set $\mathcal{S}$ of sets with costs c_s, $s \in \mathcal{S}$:
>
> Find a subset $\mathcal{S}_{\text{set}} \subseteq \mathcal{S}$ with minimal total (or average) cost such that each element in $\mathcal{U}$ is covered (at least) once.

In our case, the universe corresponds to the set of edges of the given graph ($\mathcal{U} \hat{=} \mathcal{E}$), a set corresponds to a path ($s \hat{=} p$), the cost of a set corresponds to the roughness of a path ($c_s \hat{=} r_p$), and the set cover corresponds to the desired filament cover ($\mathcal{S}_{\text{set}} \hat{=} \mathcal{P}_{\text{fil}}$). We note, that this formulation of the SCP allows overlapping sets, $s \cap s' \neq \emptyset$, $s, s' \in \mathcal{S}$, which directly translates into overlapping filaments in our FCP. By requiring that each element in $\mathcal{U}$ is contained in $\mathcal{S}_{\text{set}}$ exactly once, we may exclude filament overlaps.

An open task is then the generation of a suitable set of paths ($\mathcal{S} \hat{=} \mathcal{P}$). Since for a general graph it is not feasible to find all paths $\mathcal{P}$ (cf. the motivation of the NP-hardness proof of the FCP above), we need to find a representative subset ,$\mathcal{P}'$, of paths. We propose two approaches: (1) We sample paths from $T = 100$ random minimal spanning trees (RMST) of G. To obtain a RMST, each edge is assigned a uniformly distributed random weight and the minimum spanning tree with respect to these weights is computed. Each tree has $N(N-1)/2$ non-trivial, undirected paths that we add to our set $\mathcal{P}'$. However, the paths in a tree cannot contain loops. (2) We perform a modified breadth-first search (BFS) on the nodes, store the generated paths, and stop the search for a path p when it violates a straightness criterion, e.g., $r_{p,\text{angle}} < 60°$ (cf. Eq. 6.4.3) which is used throughout the paper. We add all permitted paths to $\mathcal{P}'$. We note that for all real-world filamentous graphs, due to filament thickness, there are spatial constraints on the number of nodes of a graph as well as on the node degrees. Moreover, for the filamentous networks considered here, the radius of curvature of a filament is typically not much smaller than the region of interest. The number of loops

is further reduced by the straightness criterion which eliminates paths with a small radius of curvature. Hence, the number of loops in the network is restricted and our heuristically modified BFS allows for loops and yields a representative set $\mathscr{P}'$ in reasonable time.

The SCP may be expressed as a binary fractional linear program [Vazirani, 2001], and we analogously write the FCP as

$$\begin{aligned} \text{minimize} \quad & \frac{\sum_{p\in\mathscr{P}} r_{p,\text{pair}}\, x_p}{\left(\sum_{p\in\mathscr{P}} x_p\right)^A} \\ \text{subject to} \quad & \sum_{p:e\in p} x_p \geq 1 \,\text{for all}\, e\in\mathscr{E} \\ & x_p \in \{0,1\} \,\text{for all}\, p\in\mathscr{P}', \end{aligned} \tag{6.4.4}$$

where in the first line $A\in\{0,1\}$ determines whether the total or the average roughness is minimized. In the second line, equality holds for an exact cover. For $A=0$, Eq. 6.4.4 is a binary linear program that may be solved using well-established and efficient algorithms [Schrijver, 1998; Linderoth and Ralphs, 2005].

For $A=1$, the fractional problem may be rewritten as a binary linear program as well [Wu, 1997; Yue et al., 2013]. To that end, we introduce new variables $y=\left(\sum_{p\in\mathscr{P}} x_p\right)^{-1}$ and $z_p = x_p y$, $p\in\mathscr{P}'$. The latter expression is non-linear but may be replaced by a set of binary linear equations, yielding

$$\begin{aligned} \text{minimize} \quad & \sum_{p\in\mathscr{P}} r_p z_p \\ \text{subject to} \quad & \sum_{p:e\in p} z_p \geq y \,\text{for all}\, e\in\mathscr{E} \\ & \sum_{p\in\mathscr{P}} z_p = 1 \\ & y \geq 0 \\ & y - z_p \leq M - M x_p \\ & z_p \leq y \\ & z_p \leq M x_p \\ & z_p \geq 0 \\ & x_p \in \{0,1\} \,\text{for all}\, p\in\mathscr{P}. \end{aligned} \tag{6.4.5}$$

Here, M is a sufficiently large constant that needs to exceed any y (cf. the Big M method [Griva et al., 2009]). Since $y=\left(\sum_{p\in\mathscr{P}'} x_p\right)^{-1} \leq 1$ for the cover of any non-empty graph, we choose $M=2$.

Thus, there are a number of options in our FCP: The input set of paths may be obtained by using a modified BFS or from sampling RMSTs or (denoted by either *BFS* or *RMST*). The filaments may overlap or not (*over*/*exact*). The objective of the FCP may be the minimization of the total or the average roughness (*total*/*avg*). The roughness of a filament may be measured by the pairwise or the all-to-all filament roughness (*pair*/*all*). Solutions of the FCP with different options are compared in the Results.

An implementation of the presented approximation schemes to the FCP with the described options is supplied as an open-source tool, "DeFiNe" (**De**composing **Fi**lamentous **Ne**tworks), under GLP3 at http://mathbiol.mpimp-golm.mpg.de/DeFiNe/. DeFiNe is programmed in Python [Van Rossum and Drake,

2011] and employs the packages SciPy [Olivier et al., 2002], NetworkX [Hagberg et al., 2008], and Cvxopt [Dahl and Vandenberghe, 2006] and PyGTK [Finlay, 2005] for a simple and user-friendly graphical user interface. DeFiNe takes as input a weighted graph in the standard .gml file format [Himsolt, 1997] and outputs a standard .gml graph with filament identities stored as edge colors. Node coordinates may be included in the input file to enable the modified BFS that takes into account edge alignments. Furthermore, manual filament assignments may be included in the input file and the similarity with the automatically obtained filament cover is assessed as described below. In addition, DeFiNe generates a standard, human-readable .csv-table of various individual filament measures for custom analyses. The filamentous structure as well as the manual filament assignments shown in Fig. 4.5.1 are available as a .gml file under the above internet address for demonstration purposes.

6.4.5 Extraction of weighted networks from images

The procedure used to extract weighted networks from image data is similar to those proposed in [Baumgarten and Hauser, 2012; Obara et al., 2012b]: (1) The original gray-scale image are pre-processed to enhance the filamentous structures. Here, a vesselness filter with kernel width of 2 pixels was used for simplicity [Frangi et al., 1998]. (2) In the filtered image, the filamentous structures are separated from the background by applying an adaptive median threshold with a block size of 49 pixels, whereby moderate variations of this size leave our findings largely unchanged. (3) The resultant binary image is skeletonized to obtain the filament center lines [Haralick et al., 1987]. (4) Then, the nodes of the network under construction are extracted as terminal points, branching points, or crossings of skeleton branches. (5) An edge is inserted between two nodes if they are directly connected via the skeleton. (6) Finally, the edges are weighted by integrating the intensity of the underlying original gray-scale image smoothed with a Gaussian filter with a standard deviation of 5 pixels along the filament and taking its average per unit length of the filament. For the images of the simulated galaxy clusters, the structures obtained by a model-based filter from [Stoica et al., 2005], Figure 6 left middle and bottom rows, are directly employed as binary images and the networks representations are obtained as described above.

6.4.6 Quality assessment of filament covers via structure-aware partition similarity measures

The extraction of the filamentous networks from image data enables comparison of the automated filament cover with manual filament assignments. Both, automated cover and manual assignment may be regarded as partitions (where we allow overlapping subsets as well). As measures for the similarity of the automated and manual partitions we use the variation of information,VI, the Jaccard index, JI, and the Rand index, RI, which are commonly used and were shown to estimate similarity reliably for distant and close partitions alike [Saporta and Youness, 2002; Meilă, 2005; Denœud and Guénoche, 2006]. For given partitions $\mathscr{C}$ =

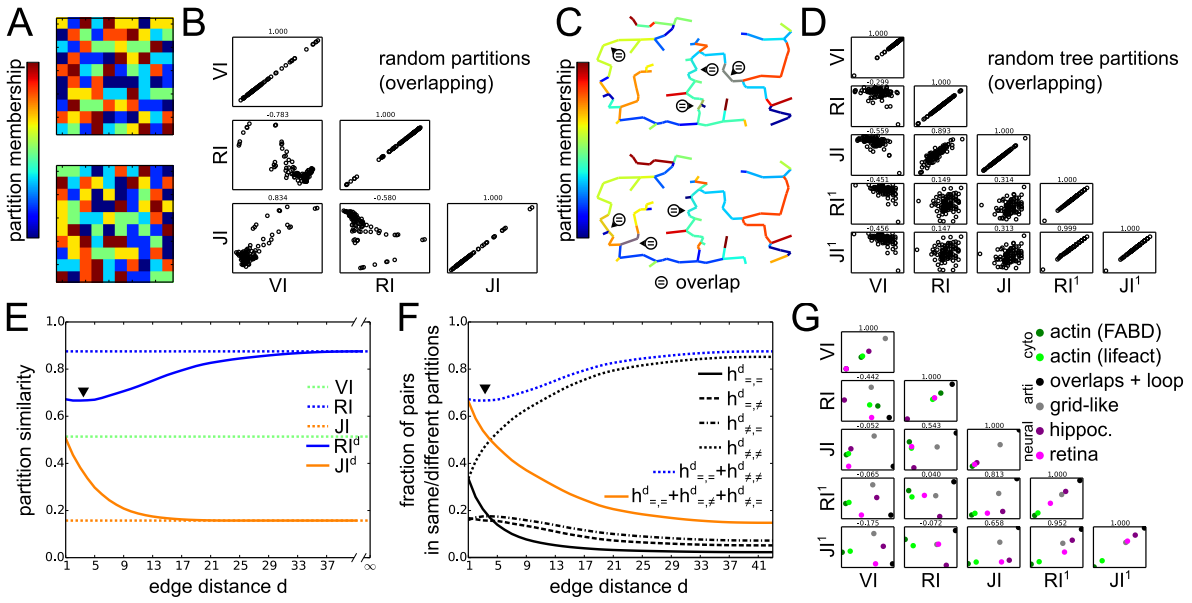


Figure 6.4.5: **Comparison of classical and extended partition similarity measures.** Analysis of 100×2 random partitions of sets of 100 numbers plus 10 duplicate ones into $5-10$ partitions (A-B). Analysis of a 100×2 path covers of Euclidean minimum spanning trees with 100 nodes distributed uniformly in the unit square, where the paths are drawn randomly and added if the overall overlap of paths is below 10 edges (C-F). Analysis of the similarities between the manual and automated decompositions of the networks studied in the paper (G). **(A)** Color-representation of two exemplary random partitions as explained above. **(B)** The classical partition similarity measures VI, RI, and JI are not correlated (cf. Kendall rank correlation coefficients $\tau < 0.9$) and may lead to opposing conclusions for the similarity of different partitions. **(C)** Two exemplary random tree path covers with overlaps ($\ominus$). **(D)** The classical partition similarity measures VI, RI, and JI show no correlation among themselves (except for the pairing of RI and JI), nor with the structure-aware RI^1 and JI^1. In contrast, RI^1 and JI^1 are very strongly correlated (cf. $\tau > 0.9$) and yield consistent results for the similarity of different partitions. **(E)** Similarity of the partitions shown in (C) in dependence on maximal distance d between considered pairs of edges (cf. Fig. 4.5.3 for a detailed discussion). The RI^d shows a non-monotonous dependency on d (triangle). **(F)** This non-monotonicity of RI^d may be explained by the entries of the contingency table $h_{\times,\times'}$, $\times, \times' \in \{=, \neq\}$. For small distances d, the fraction of true positives (solid black, $h_{=,=}$) drops slower than the fraction of true negatives (dotted black, $h_{\neq,\neq}$) and for larger d, this trend is reversed. Hence, their sum (dashed blue) shows a minimum at intermediate distances d (triangle). In contrast, when summed up (solid yellow), the fast drop in the fraction of true positives dominates over the slightly non-monotony of the false positives and negatives. **(G)** For the investigated artificial and biological networks, the classical measures VI, RI, and JI yield partially opposing results on the similarity of the manual assignment and the automated decomposition (cf. $\tau < 0$). The structure-aware similarity measures RI^1 and JI^1 are strongly correlated and yield consistent results (cf. $\tau = 0.952$).

$\{C_1,\ldots,\mathscr{C}_C\}$ and $\mathscr{C}' = \left\{C'_1,\ldots,\mathscr{C}'_{C'}\right\}$, they are computed via

$$\begin{aligned} \mathrm{VI}\left(\mathscr{C},\mathscr{C}'\right) &= 1+\left(U\log U\right)^{-1}\cdot \\ &\quad \cdot\sum_{i,j} g_{i,j}\left(\log\left(\frac{g_{i,j}}{g_{\cdot,j}}\right)+\log\left(\frac{g_{i,j}}{g_{i,\cdot}}\right)\right), \end{aligned} \tag{6.4.6}$$

$$\mathrm{RI}\left(\mathscr{C},\mathscr{C}'\right) = \frac{h_{=,=}+h_{\neq,\neq}}{h_{=,=}+h_{=,\neq}+h_{\neq,=}+h_{\neq,\neq}}, \tag{6.4.7}$$

$$\mathrm{JI}\left(\mathscr{C},\mathscr{C}'\right) = \frac{h_{=,=}}{h_{=,=}+h_{=,\neq}+h_{\neq,=}}, \tag{6.4.8}$$

where $U=\sum_{i=1}^{C}|\mathscr{C}_i|=\sum_{j=1}^{C'}\left|\mathscr{C}'_j\right|$, $g_{i,j}=\left|\mathscr{C}_i\cap\mathscr{C}'_j\right|$, $g_{\cdot,j}=\sum_{i=1}^{C}g_{i,j}$, and $g_{i,\cdot}=\sum_{j=1}^{C'}g_{i,j}$. The contingency tables $h_{\times,\times'}$, $\times,\times'\in\{=,\neq\}$, provide the numbers of edge pairs which are in the same or different sets in the two partitions, respectively, and is related to $g_{i,j}$ as shown in [Hubert and Arabie, 1985]. All measures are restricted to the unit interval with larger values reflecting higher similarity [Meilă, 2003].

While these measures of partition similarity are widely used [Meilă, 2005; Lancichinetti and Fortunato, 2009b], they pose some difficulties. The variation of information, VI, is only well-defined for disjoint partitions, which occur for non-overlapping filaments. While the Jaccard index, JI, and the Rand index, RI, cover intersecting partitions they may generally yield opposing results. We demonstrate this inconsistency by investigating two types of random partitionings: First, for 100 repetitions, we randomly partitioned 2 sets of 100 numbers and up to 10 duplicates (to simulate overlapping filaments) into $5-10$ random partitions (Fig. 6.4.5A). While VI and JI were correlated (Fig. 6.4.5B; cf. Kendall rank correlation coefficient $\tau>0$), the other two combinations showed a strong negative correlation (cf. $\tau<0$). Second, to study filament covers that resemble the decomposition of real filamentous networks more closely, we constructed a relative neighborhood graph [Toussaint, 1980; Supowit, 1983] with 100 nodes uniformly distributed in the unit square and computed a random minimum spanning tree (Fig. 6.4.5C). For 100 repetitions of this procedure, we partitioned the resultant tree into filaments by choosing a path at random and adding it to the decomposition if the total overlap of any two paths already in the decomposition is below 10 edges (cf. $\ominus$ for overlaps). Again, the correlation among the classical similarity measures was poor or negative (Fig. 6.4.5D; except for the correlation between RI and JI; $|\tau|<0.6$). Although other measure for the similarity of intersecting partitions have been proposed [Goldberg et al., 2010; Lancichinetti and Fortunato, 2009a; Lancichinetti et al., 2009], we adhere to RI and JI for simplicity.

More severely, however, the above similarity measures do not take into account the structure of the graph G underlying the (edge-)partitions induced by the obtained filament covers. To date, we are only aware of structure-aware similarity measures for the comparison of partitions whose items are distributed in Euclidean space [Zhou et al., 2005; Bae et al., 2010; Coen et al., 2010]. Yet, these approaches do not take into account the explicit graph structure of the partitions. To remedy this shortcoming, we introduce a suite of measures, the structure-aware Rand and Jaccard index, RI^d and JI^d, respectively. To that end, the contin-

gency tables $h_{\times,\times'}$ in Eqs. 6.4.7 and 6.4.8 are replace by distance dependent $h^d_{\times,\times'}$,

$$\mathrm{RI}^d\left(\mathscr{C},\mathscr{C}'\right) = \frac{h^d_{=,=}+h^d_{\neq,\neq}}{h^d_{=,=}+h^d_{=,\neq}+h^d_{\neq,=}+h^d_{\neq,\neq}}, \tag{6.4.9}$$

$$\mathrm{JI}^d\left(\mathscr{C},\mathscr{C}'\right) = \frac{h^d_{=,=}}{h^d_{=,=}+h^d_{=,\neq}+h^d_{\neq,=}} \tag{6.4.10}$$

where $h^d_{\times,\times'}$, $\times,\times' \in \{=,\neq\}$, $d \in \mathbb{N}_{>0}$, count the number of edge pairs which are in the same or different sets in the two partitions, respectively, and which are separated by at most d nodes in G. More precisely, we define

$$h^d_{\times,\times'} = \Big\{\#(e_0,e_1) \mid e_0 \in \mathscr{C}_i \cap \mathscr{C}'_{i'},\, e_1 \in \mathscr{C}_j \cap \mathscr{C}'_{j'}, \text{ with } i \times j \text{ and } i' \times' j' \text{ and } D_{L(G)}(e_0,e_1) \leq d\Big\}, \tag{6.4.11}$$

where $\#(e_0,e_1)$ is the number of edges (e_0,e_1) and $D_{L(G)}(e_0,e_1)$ is the length of the shortest path between nodes in the line graph $L(G)$ of G corresponding to the edges e_0 and e_1. For example, $h^0_{=,=}$ counts the the number of adjacent edges which are in the same set in both partitions (local perspective). In contrast, $h^\infty_{\times,\times'} \equiv h_{\times,\times'}$ reproduce the original measures which do not take into account the positions of edges in the graph (global perspective).

To investigate the performance of our extended, structure-aware partition similarity measures, RI^d and JI^d, we apply them to the artificial graph-based random partitions described above (cf. Fig. 6.4.5C). Indeed, when considering the partition membership of neighboring edges only, i.e., RI^1 and JI^1, the similarity measures yield very consistent results (Fig. 6.4.5D; cf. $\tau = 0.999$) in contrast to the lower correlation with the classical similarity measures (cf. $\tau < 0.9$). Investigating the dependency of RI^d and JI^d on the distance d for the tree filament covers shown in Fig. 6.4.5D, we find that RI and JI (Fig. 6.4.5E; dotted blue and yellow) over- and underestimate the partition similarity with respect to RI^1 and JI^1 (Fig. 6.4.5E; solid blue and yellow). Furthermore, we find that RI^d is non-monotonic in d (Fig. 6.4.5E; cf. the black triangle). These errors in estimation are explained by the large fraction of false negatives ($h_{\neq,\neq}$) and the small fraction of true positives ($h_{=,=}$), respectively, which dominate for large distances d, i.e., the limit in which the graph structure is ignored (Fig. 6.4.5F; dotted black and solid black). Due to the differential increase/decrease of $h_{\neq,\neq}/h_{=,=}$, their combination and therefore RI^1 is non-monotonic (Fig. 6.4.5F; dashed blue). Finally, we observe opposing results of the classical partition similarity measures also for the filament covers of artificial and biological filamentous networks investigated in the main text, while our extended, structure-aware measures RI^1 and JI^1 provide consistent similarity results (Fig. 6.4.5G).

6.4.7 Robustness of filament covers against incomplete knowledge of underlying network structure and image noise

Our approach enables accurate decomposition of a given filamentous network into its constitutive filaments (cf. Results). However, the preceding extraction of the network from image data is often non-trivial (cf. Methods). Therefore, to assess the robustness of our approach, we test how the accuracy of our filament de-

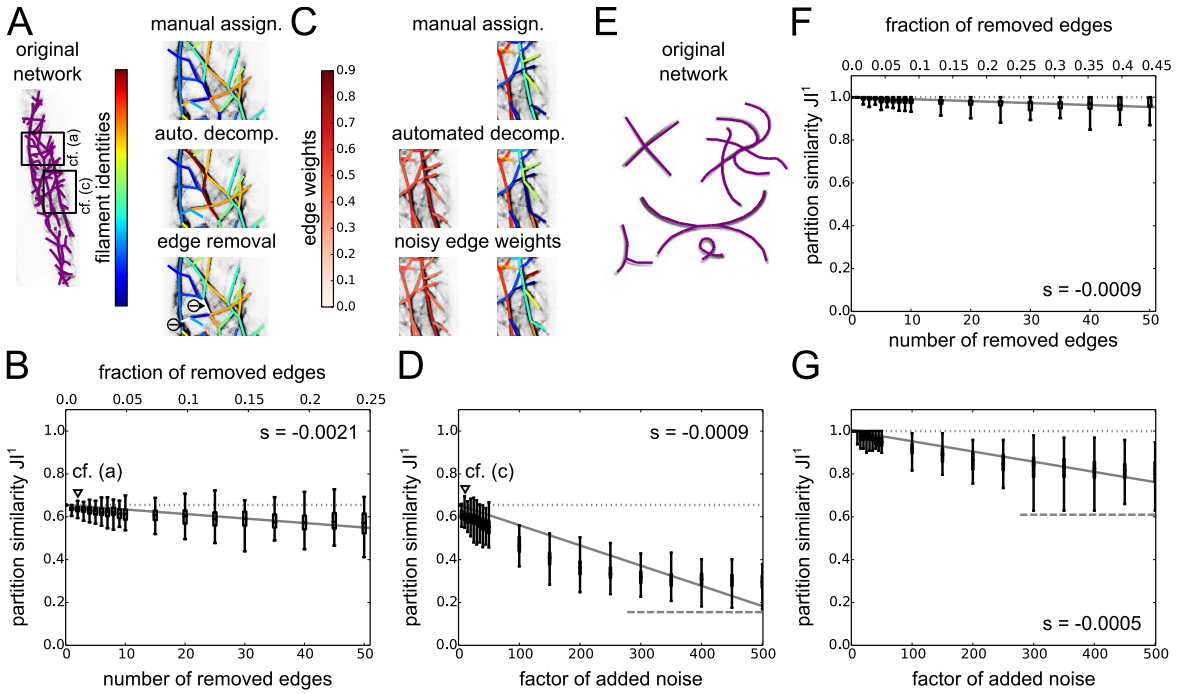


Figure 6.4.6: **Analyses of robustness of filament covers against incomplete knowledge of network and image noise.** A cytoskeletal and a contrived network are decomposed automatically by solving the FCP with options given in Fig. 4.5.3C and Fig. 4.5.1C, respectively. **(A)** Overlay of extracted actin network structure and original image data (left panel). Sections of cytoskeletal network with edge colors representing the manual assignment, the optimal filament cover obtained for the full, non-disrupted network, and the optimal filament cover after removal of two edges which are shown in white (right panels). **(B)** Similarity of manual filament assignment and automated filament covers after removal of increasing numbers of edges, measured by structure-aware Jaccard index JI^1. On average, JI^1 decreases with the number of removed edges as shown by a linear fit with slope $s = -0.0021$ (solid gray line). Occasionally, the removal of edges increases the accuracy of the filament cover above the accuracy of the original solution (dotted gray line and triangle; cf. panel (A)). **(C)** Sections of cytoskeletal network with edge colors representing the original edge weights and the edge weights after adding Gaussian noise (left panels). Sections of cytoskeletal network with edge colors representing the manual assignment, the optimal filament cover obtained for the full, non-disrupted network, and the optimal filament cover after adding Gaussian noise (right panels). **(D)** Similarity JI^1 of manual filament assignment and automated filament covers after adding Gaussian noise. On average, JI^1 decreases with increasing noise factor as shown by a linear fit with slope $s = -0.0009$ (solid gray line). Occasionally, the noisy edge weights lead to an increase in accuracy of the filament cover above the accuracy of the original solution (dotted gray line and triangle; cf. panel (C)). The decrease levels off for large noise factors and JI^1 approaches a constant value (dashed gray line). **(E)** Overlay of extracted contrived network structure and original image data. **(F)** Results for the contrived network analogue to those presented for the cytoskeletal network in panel (B). The average change in JI^1 per removed edge is captured by a linear fit with slope $s = -0.0009$. **(G)** Results for the contrived network analogue to those presented for the cytoskeletal network in panel (f). The average change in JI^1 per unit increase in the noise factor is captured by a linear fit with slope $s = -0.0005$.

composition is affected (1) by incomplete knowledge of the true underlying network structure and (2) by image noise which affects the edge weights of the extracted network. We perform these analyses for the actin cytoskeleton shown in Fig. 4.5.3 (Fig. 6.4.6A, left panel) and the contrived network shown in Fig. 4.5.1 (Fig. 6.4.6E).

(1) First, we start from the original, weighted network and randomly remove one of the E edges to model erroneous segmentation. For the disrupted network, we recompute the optimal filament cover (with the same options as in Figs. 4.5.1C and 4.5.3C, respectively) and calculate its agreement with the original manual segmentation (measured by the structure-aware Jaccard index JI^1; the removed edge is assigned a dummy label). We repeat the procedure for E networks from which a single, randomly chosen edge has been removed. Next, we repeat the procedure for E networks from which a randomly chosen double of edges has been removed. We then proceed with triplets, quartets, and so on up to subsets of 50 randomly chosen edges.

As expected, the removal of increasing numbers of edges typically decreases the agreement of the automated filament cover with the manual assignment for the cytoskeletal as well as the contrived network (Fig. 6.4.6B and F). For both networks, however, the decrease is slow and JI^1 increases only by around 0.002 per removed edge (cf. Fig. 6.4.6B and F, solid gray line indicates linear fit). Interestingly, for the actin cytoskeleton, the removal of certain edges may even increase the accuracy of the filament cover (Fig. 6.4.6A, right panels show manual filament assignment and automated filament cover the original network, and an exemplary filament cover obtained after the removal of two edges, colored white here, which improves the agreement with the manual assignment; cf. Fig. 6.4.6B, dotted gray line and triangle).

(2) Second, we simulate image noise by adding centered Gaussian noise Δw to the edge weights of the original network with

$$\mathrm{E}[\Delta w] = 0, \tag{6.4.12}$$

$$\mathrm{Sd}[\Delta \mathrm{w}] = \left(1 + \frac{f}{100}\right) w. \tag{6.4.13}$$

We normalize the standard deviation of the added noise by the original edge weights to avoid extreme fluctuations, and f is referred to as noise factor. For each noise factor, we construct 100 networks, recompute the optimal filament covers, and measure their agreement with the manual filament assignment, as in the first scenario above.

For both the contrived and the cytoskeletal network, the accuracy of the filament cover decreases with increasing noise, as expected (Fig. 6.4.6D and G). However, this decrease in accuracy is slow and JI^1 decreases by less than 0.001 when increasing the standard deviation of the noise by 1% of the original edge weights, i.e., when increasing the noise factor by one (cf. Fig. 6.4.6D and G, solid gray lines indicate linear fits). We note that with increasing edge noise the accuracy of the filament cover approaches a constant, non-zero JI^1 which reflects that some information about the filament structure maybe obtained from the topology of the network alone, irrespective of the edge weights (cf. Fig. 6.4.6D and G, dashed gray lines).

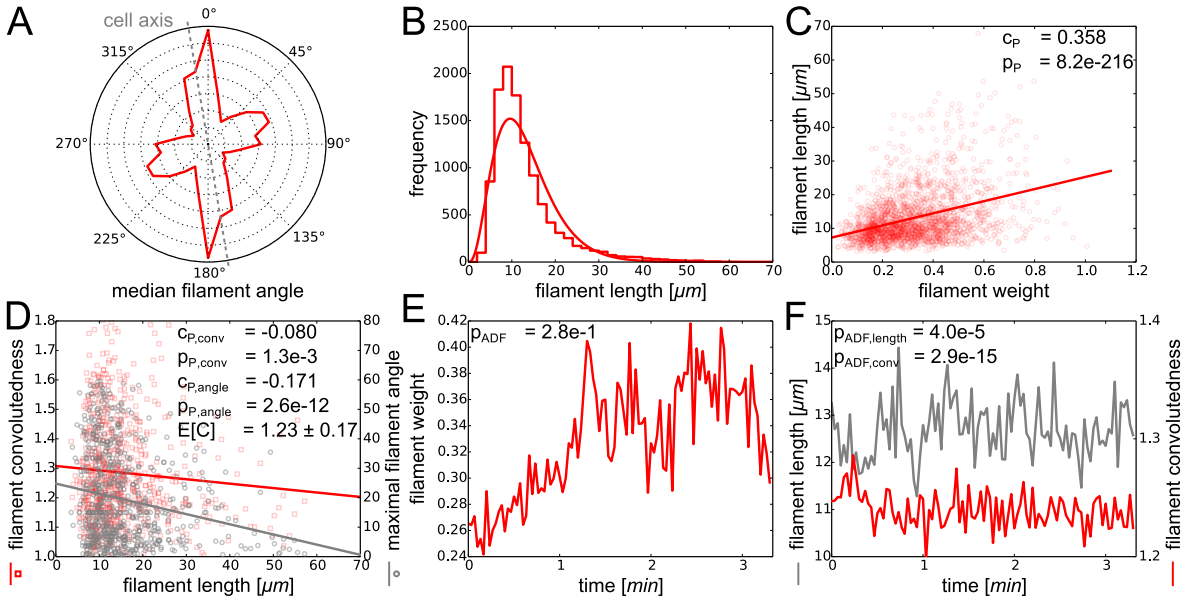


Figure 6.4.7: **Filament analyses of** 100 **cytoskeletal networks.** Results from filament decompositions of 100 cytoskeletal networks extracted from a movie of a plant cytoskeleton over 200 s. The cytoskeletal networks are decomposed automatically by solving the exact FCP (*exact*) for paths from a modified breadth-first search (*BFS*) and by minimizing the total (*total*) pairwise filament roughness (*pair*; cf. Fig. 4.5.3). **(A)** The distribution of median filament angles shows that the majority of filaments is aligned parallel to the cell axis (gray dashed line). **(B)** Filament lengths (bars) follows a gamma distribution (line shows maximum likelihood fit). **(C)** Filament length correlates with filament weight (cf. linear regression and Pearson correlation coefficient $c_P > 0$ and p-value $p_P < 0.05$) **(D)** Scatter plot of filament convolutedness versus filament length shows a negative correlation (cf. red squares, $c_{P,\text{conv}} < 0$, and $p_{P,\text{conv}} < 0.05$) with an average convolutedness of $\mathrm{E}[C] = 1.23 \pm 0.17$. The maximum filament angle correlates negatively with the filament length (cf. gray circles, $c_{P,\text{angle}} < 0$, and $p_{P,\text{angle}} < 0.05$), indicating that longer (and thicker, cf. (C)) filaments are less curved. **(E)** Time series of average filament weight over 200 s shows large fluctuations and is non-stationary (cf. augmented Dickey-Fuller test p-value $p_{\text{ADF}} \geq 0.05$). **(F)** Time series of filament length and convolutedness are stationary over the recording period (cf. $p_{\text{ADF,length}} < 0.05$ and $p_{\text{ADF,conv}} < 0.05$).

6.4.8 Filament analysis for networks extracted from movie of plant actin cytoskeleton

To further strengthen our statistical analyses of cytoskeletal actin filaments (cf. Fig. 4.5.3), we investigate a complete movie of a plant cytoskeleton of 100 frames over 200 s (cf. Methods for details). For each frame, we extract a weighted network representation of the cytoskeleton as described above (cf. Methods for details) and solve the FCP with options described in Fig. 4.5.3, i.e., we solve the exact FCP (*exact*) for paths from a modified breadth-first search (*BFS*) and by minimizing the total (*total*) pairwise filament roughness (*pair*). Analysis of various properties of the automatically obtained filaments confirms our findings in Fig. 4.5.3: The filaments show a preferential alignment parallel to the cell axis throughout the movie (Fig. 6.4.7A). The distribution of filament lengths, pooled across the duration of the movie, confirms the reported gamma distribution (Fig. 6.4.7B; maximal likelihood fits of normal, Weibull, and Rayleigh distributions yield higher values for the Akaike information criterion [Akaike, 1974]). Filament length is correlated with filament weight, i.e., longer filaments are typically thicker (Fig. 6.4.7C; Pearson correlation p-value $p_{\mathrm{P}} < 0.05$). Moreover, the correlation between different measures of filament curvedness, i.e., the filament bending and the maximal filament angle, are consistently negatively correlated with the filament length (Fig. 6.4.7D; $p_{\mathrm{P}} < 0.05$).

In addition to these previously analyzed features of filamental organization, we study the course of different filament properties over time: The average filament weight shows large fluctuations and is non-stationary over the recording period (Fig. 6.4.7E; cf. augmented Dickey-Fuller test p-value $p_{\mathrm{ADF}} \geq 0.05$). This non-stationarity suggests substantial changes in the prevalence of fine actin filament and thick bundles, respectively, and prompts further investigations. However, we found that the average filament length as well as the average filament bending remain stationary over the course of 200 s (Fig. 6.4.7E; cf. $p_{\mathrm{ADF,length}} < 0.05$ and $p_{\mathrm{ADF,conv}} < 0.05$). Since the length distribution of filaments tunes the mechanical properties of filamentous networks [Kasza et al., 2010; Bai et al., 2011], this stationarity of the average filament length may be of immediate biological relevance. The stationarity of the average filament bending may be a direct consequence of the roughly constant filament length distribution (cf. Fig. 6.4.7D) in combination with the resultant physical constraints of actin filament length on filament bending.

6.4.9 Overview of different stages of filament decomposition of artificial, biological, and cosmic networks

We test our method of decomposing a given weighted network into filaments by solving the FCP for different filamentous networks. In addition to the four networks presented in the main text and the 100 frames analyzed in Appendix 6.4.8, we investigate four more networks of different types and show the different stages of our analysis. Starting from gray-scale image data of contrived, neural, cytoskeletal and cosmic network structures (Fig. 6.4.9, 1st column), we pre-process the images to obtain a binary representation of the filament center lines (Fig. 6.4.9, 2nd column), and extract a weighted network representation as described in the Methods (Fig. 6.4.9, 3rd column). For the contrived and biological and the cosmic networks, we manually assign filament identities and compute the connected components, respectively (Fig. 6.4.9, 4th column). Finally, we decompose the networks into filaments by solving the FCP with different options (Fig. 6.4.9, 5th column). For the first contrived network, we allow overlapping filaments (Fig. 6.4.9A) while for the second, grid-like contrived network (Fig. 6.4.9B), the neural networks (Fig. 6.4.9C and D), the cytoskeletal networks

(Fig. 6.4.9E and F), and the cosmic webs (Fig. 6.4.9G and H), we obtain exact filament covers with options described in Fig. 4.5.1E. The agreement of manual assignments and automated filament decompositions of the studied networks is measured by the classical and the structure-aware Jaccard indices JI and JI^1 and shows good agreement (JI^1 close to 1, and cf. discussion of Fig. 4.5.3D) despite occasional over- (cf. ⊕) or under-segmentation (cf. ⊖) of filaments.

6.4.10 Open contour-based filament decomposition and filament cover-based post-processing

Finally, we demonstrate how our filament cover-based approach may be used to post-process and improve filament decompositions obtained from other, e.g., open contour-based approaches. For the demonstration, we select SOAX [Xu et al., 2015], a fully automated, stretching open active contour-based approach which is available as an open-source software tool to extract a network-like representation (i.e., coordinates of filament center lines as well as junctions are provided) from image data. As a test case, we study the contrived filamentous structure investigated in Fig. 4.5.1. For a fair comparison of our and the open contour-based approach, we apply SOAX to the pre-processed and segmented image data (cf. Methods and Fig. 6.4.10A, second panel) to which we further apply a Gaussian filter of unit standard deviation to obtain smooth intensity gradients required by the algorithm. SOAX is run using the default parameters and the resulting filament identities are manually assigned to match those of the manual solution (Fig. 6.4.10B). To quantify the quality of the decomposition, we manually assign filament identities in our original network representation (cf. Fig. 6.4.10A, third panel) according to the open contour-based result (cf. Fig. 6.4.10B) and compare the result to the manual assignment (cf. Fig. 6.4.10A, fourth panel). The structure-aware Jaccard index $JI^1 = 0.938$ is close to 1 and indicates good agreement between open-contour based decomposition and manual filament assignment. We note that some junctions/nodes obtained from SOAX are split in two in comparison to our extracted networks (cf. intersecting ○).

Moreover severely, some filaments are over-segmented and thus fragmented (cf. ⊕), especially overlapping filaments which are not captured in the open contour-based approach (cf. ⊖). To remedy this shortcoming, we apply our filament cover-based approach to post-process the open contour-based decomposition and merge over-segmented filament fragments. To this end, we convert the open contour-based filament representation into a weighted network, where edge weights represent average filament segment intensities as before (cf. Methods and Fig. 6.4.10C). As before, a collection of paths $\mathscr{P}'$ is sampled using a breadth-first search (*BFS*) and their pairwise roughness values r_p, $p \in \mathscr{P}'$, are computed according to Eq. 6.4.1 (*pair*). Then, to take into account the initial open contour-based filament decomposition $\mathscr{F}$ as a starting point, in which certain edges have already been assigned to the a given filament, we modify the roughness values of the sampled paths: For each initial filament or fragment that is fully contained within a sampled path, the roughness of that path is decreased by a large value, $R_{\text{filament}} = 10^4$, which is larger than any r_p to favor the inclusion of these filaments or fragments in the optimal filament cover. Since the subtraction of R_{filament} yields negative roughness values which would lead to the inclusion of all these paths, we add another, even

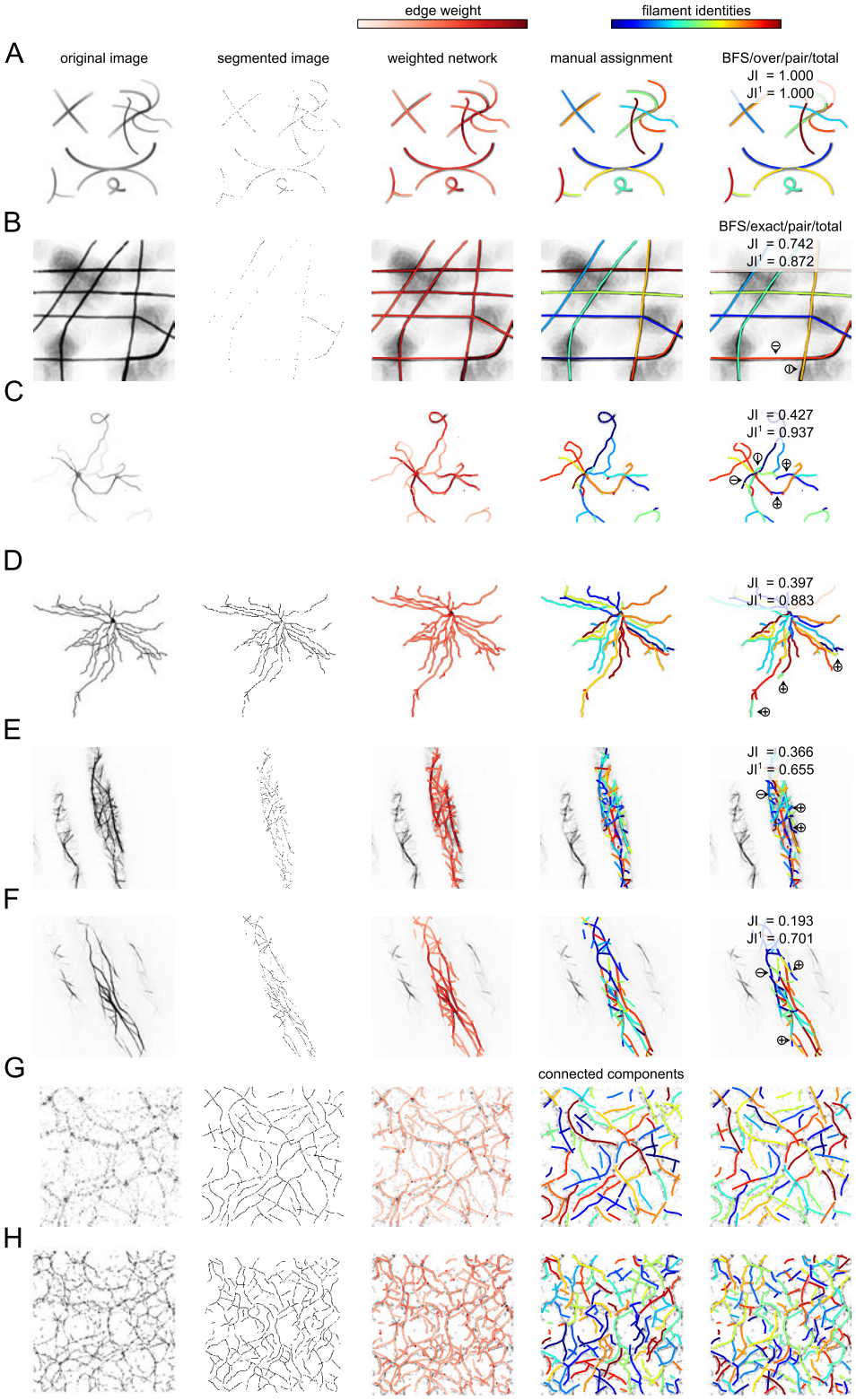

edge weight
filament identities
A
original image
segmented image
weighted network
manual assignment
BFS/over/pair/total
JI = 1.000
JI¹ = 1.000
B
BFS/exact/pair/total
JI = 0.742
JI¹ = 0.872
C
JI = 0.427
JI¹ = 0.937
D
JI = 0.397
JI¹ = 0.883
E
JI = 0.366
JI¹ = 0.655
F
JI = 0.193
JI¹ = 0.701
G
connected components
H

Figure 6.4.9: **Overview of studied networks, manual assignments, and filament covers obtained from solving the FCP.** Original gray-scale image data (1st column), binary images of filament center lines (2nd column), extracted networks with color-coded edge weights (3rd column), manual filament assignments of contrived and biological networks and connected connected components of cosmic networks, respectively (4th column), and automatically obtained filament covers (5th column). Agreement between manual decompositions and automated filament cover is quantified by a number of measures (cf., e.g., Methods and Fig. 4.5.2), here the classical and the structure-aware Jaccard indices JI and JI^1 are shown. **(A)** Contrived network with crossing and overlapping filaments and a loop (cf. Fig. 4.5.1). **(B)** For a contrived, grid-like network, the automated decomposition correctly detects most of the filaments (JI^1 close to 1). Only the filament in the bottom right corner with a kink is over-segmented (⊕) because the curvature restriction of the initial paths does not allow such large angles (about 90° here, cf. Eq. 6.4.3). **(E)** Neural network of hippocamal neuron (cf. Fig. 4.5.2). **(D)** The decomposition of the network of a retinal ganglion cell shows good agreement with the manual results (JI^1 close to 1). A few filaments are over-segmented (⊕), e.g., due to kinks in the filaments that are not captured by the initial set of paths (cf. the center ⊕). **(E)** Cytoskeletal network of actin filaments (cf. Fig. 4.5.3). **(F)** For the actin network extracted from the confocal recording of a Lifeact-labeled cytoskeleton, the automated partitions agrees well with the manual results (JI^1 close to 1, and cf. discussion of Fig. 4.5.3D). A few examples of over- and under-segmentation (⊖) are marked. **(G)** Cosmic web of galaxies (cf. Fig. 4.5.4). **(H)** The dense web of simulated galaxies consists of many connected components that are further decomposed into filaments (cf. Fig. 4.5.4 for a discussion). Image data for panels (G) and (H) from: Stoica et al., A&A, 434, 423-432, 2005, reproduced with permission © ESO.

larger constant $R_{\text{offset}} = 10^8 > R_{\text{filament}}$ to all roughness values, i.e.,

$$r_p' = r_p - \sum_{\substack{f \in \mathscr{F} \\ f \subset p}} R_{\text{filament}} + R_{\text{offset}}. \tag{6.4.14}$$

For these modified roughness values r_p', we solve the FCP by minimizing the total roughness (*total*) and allowing for overlaps (*over*; Fig. 6.4.10D). The resulting post-processed filament decomposition merges several filament fragments which were over-segmented by the open-contour based approach and shows very good agreement of JI = 0.776 and JI^1 = 1.000 with the manual filament assignment. Interestingly, in this decomposition, parts of two filaments are interchanged (cf. ⊕) as in Fig. 4.5.1F for different FCP options. In conclusion, for any approach that detects filaments from image data and yields a weighted network representation, our filament cover-based approach may provide a helpful means to further post-process and enhance the accuracy of the obtained filament decomposition.

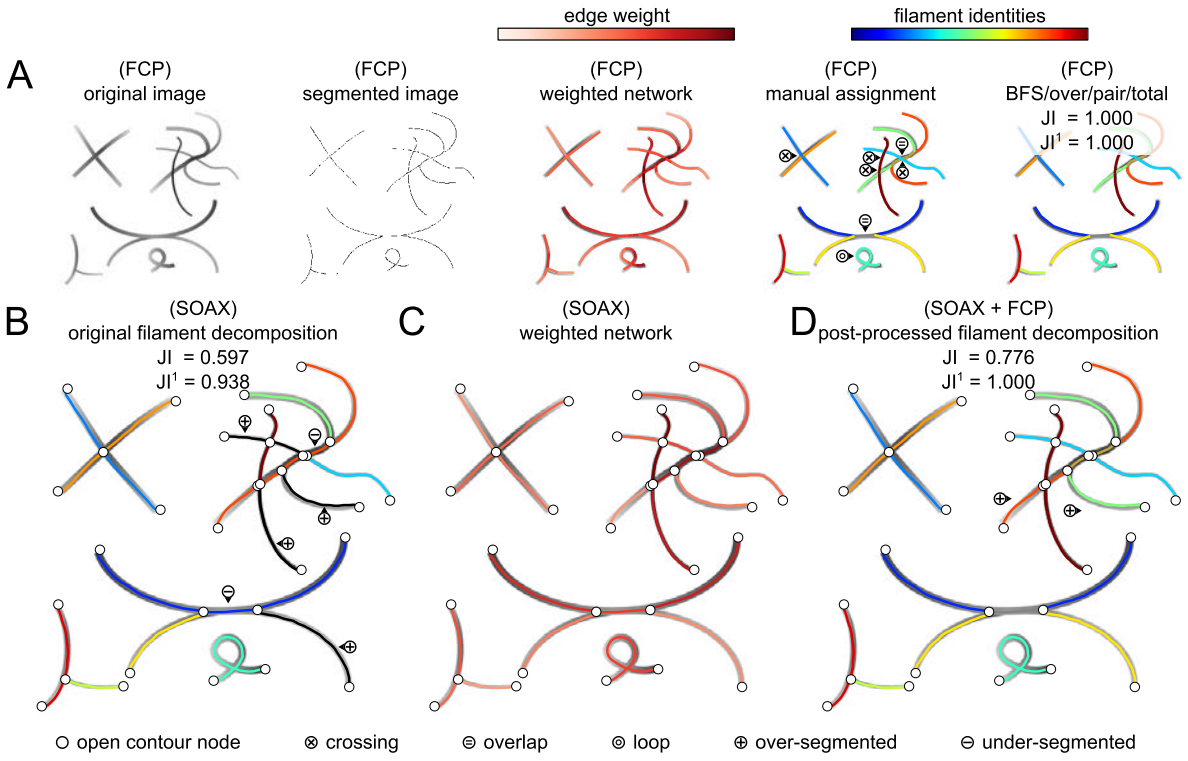


Figure 6.4.10: **Open contour-based filament detection and filament cover-based post-processing.** **(A)** Different stages of our filament cover problem (FCP)-based analysis for a contrived filament structure, from original image to segmented filament center lines and weighted network representation, manual filament assignment and automated solution (cf. Fig. 6.4.9A for further explanations). **(B)** Filaments and junctions (cf. circles) identified from the segmented filament center line image using SOAX, an stretching open active contour-based approach [Xu et al., 2015]. Color-coded filament identities were manually assigned to match those of the manual solution in (A) and excess filament fragments were colored black. While the agreement with the manual solution is good (JI^1 close to 1), some filaments are over-segmented (cf. ⊕) and thus fragmented, especially at locations of filament overlaps (cf. ⊖). **(C)** Weighted network representation of the contrived filamentous structure obtained from SOAX. **(D)** Using the filament assignments from SOAX in (B) as a starting point, our filament cover-based approach is used to post-process the filament decomposition, which merges broken filament fragments and improves the agreement with the manual solution ($JI^1 = 1$).

7 Bibliography

Achterberg, T. (2009). SCIP: solving constraint integer programs. *Math Prog Comp*, 1:1–41.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Trans Autom Control*, 19(6):716–723.

Akkerman, M., Overdijk, E. J., Schel, J. H., Emons, A. M. C., and Ketelaar, T. (2011). Golgi body motility in the plant cell cortex correlates with actin cytoskeleton organization. *Plant Cell Physiol*, 52(10):1844–1855.

Ali, M. and Brocchini, S. (2006). Synthetic approaches to uniform polymers. *Adv Drug Delivery Rev*, 58(15):1671–1687.

Allan, V. (1995). Membrane traffic motors. *FEBS Lett*, 369(1):101–106.

Allard, J. F., Ambrose, J. C., Wasteneys, G. O., and Cytrynbaum, E. N. (2010). A mechanochemical model explains interactions between cortical microtubules in plants. *Biophys J*, 99(4):1082–1090.

Allen, N. S. and Allen, R. D. (1978). Cytoplasmic streaming in green plants. *Annu Rev Biophys Bioeng*, 7:497–526.

Ando, D., Korabel, N., Huang, K. C., and Gopinathan, A. (2015). Cytoskeletal network morphology regulates intracellular transport dynamics. *Biophys J*, 109(8):1574–1582.

Andreatta, G. and Mason, F. (1995). Path covering problems and testing of printed circuits. *Discrete Appl Math*, 62(1):5–13.

Apelt, F., Breuer, D., Nikoloski, Z., Stitt, M., and Kragler, F. (2015). Phytotyping4D: a light-field imaging system for non-invasive and accurate monitoring of spatio-temporal plant growth. *Plant J*, 82(4):693–706.

Arganda-Carreras, I., Sorzano, C. O., Marabini, R., Carazo, J. M., Ortiz-de Solorzano, C., and Kybic, J. (2006). Consistent and elastic registration of histological sections using vector-spline regularization. In *Computer vision approaches to medical image analysis*, pages 85–95. Springer, Berlin.

Ascoli, G. A., Donohue, D. E., and Halavi, M. (2007). Neuromorpho.org: a central resource for neuronal morphologies. *J Neurosci*, 27(35):9247–9251.

Aubert, D., Pichon, C., and Colombi, S. (2004). The origin and implications of dark matter anisotropic cosmic infall on L haloes. *Mon Not R Astron Soc*, 352(2):376–398.

Ausmees, N., Kuhn, J. R., and Jacobs-Wagner, C. (2003). The bacterial cytoskeleton: an intermediate filament-like function in cell shape. *Cell*, 115(6):705–713.

Avisar, D., Prokhnevsky, A. I., and Dolja, V. V. (2008a). Class VIII myosins are required for plasmodesmatal localization of a closterovirus Hsp70 homolog. *J Virol*, 82(6):2836–2843.

Avisar, D., Prokhnevsky, A. I., Makarova, K. S., Koonin, E. V., and Dolja, V. V. (2008b). Myosin XI-K is required for rapid trafficking of Golgi stacks, peroxisomes, and mitochondria in leaf cells of Nicotiana benthamiana. *Plant Physiol*, 146(3):1098–1108.

Badoual, M., Jülicher, F., and Prost, J. (2002). Bidirectional cooperative motion of molecular motors. *Proc Natl Acad Sci*, 99(10):6696–6701.

Bae, E., Bailey, J., and Dong, G. (2010). A clustering comparison measure using density profiles and its application to the discovery of alternate clusterings. *Data Min Knowl Disc*, 21(3):427–471.

Bagchi, S., Tomenius, H., Belova, L. M., and Ausmees, N. (2008). Intermediate filament-like proteins in bacteria and a cytoskeletal function in Streptomyces. *Mol Microbiol*, 70(4):1037–1050.

Bai, M., Missel, A. R., Levine, A. J., and Klug, W. S. (2011). On the role of the filament length distribution in the mechanics of semiflexible networks. *Acta Biomater*, 7(5):2109–2118.

Bálint, S., Verdeny Vilanova, I., Sandoval Álvarez, A., and Lakadamyali, M. (2013). Correlative live-cell and superresolution microscopy reveals cargo transport dynamics at microtubule intersections. *Proc Natl Acad Sci*, 110(9):3375–3380.

Baluška, F., Jasik, J., Edelmann, H., Salajova, T., and Volkmann, D. (2001). Latrunculin B-induced plant dwarfism: plant cell elongation is F-actin-dependent. *Dev Biol*, 231(1):113–124.

Banerjee, N. and Park, J. (2015). Modeling and simulation of biopolymer networks: classification of the cytoskeleton models according to multiple scales. *Korean J Chem Engineer*, 32(7):1–11.

Barabási, A.-L. and Oltvai, Z. N. (2004). Network biology: understanding the cell's functional organization. *Nat Rev Genet*, 5:101–113.

Barthélemy, M. (2011). Spatial networks. *Phys Rep*, 499(1):1–101.

Barthélemy, M. and Flammini, A. (2008). Modeling urban street patterns. *Phys Rev Lett*, 100(13):138702.

Bartles, J. R. (2000). Parallel actin bundles and their multiple actin-bundling proteins. *Curr Opin Cell Biol*, 12(1):72–78.

Bassell, G. and Singer, R. H. (1997). mRNA and cytoskeletal filaments. *Curr Opin Cell Biol*, 9(1):109–115.

Baumgarten, W. and Hauser, M. J. (2012). Computational algorithms for extraction and analysis of two-dimensional transportation networks. *J Comput Interdiscip Sci*, 3:107–16.

Bausch, A. and Kroy, K. (2006). A bottom-up approach to cell mechanics. *Nat Phys*, 2(4):231–238.

Belsley, D. A. (2006). *Encyclopedia of Statistical Sciences*, chapter Conditioning diagnostics. Wiley Online Library, Bethesda, 2nd edition.

Benetatos, P. and Zippelius, A. (2007). Anisotropic random networks of semiflexible polymers. *Phys Rev Lett*, 99(19):198301.

Benhamou, S. (2007). How many animals really do the Lévy walk? *Ecology*, 88(8):1962–1969.

Bennett, M. (1977). *Handbook of Physiology*, chapter Electrical transmission: a functional analysis and comparison to chemical transmission. Wiley Online Library, Bethesda, 1st edition.

Berger, F., Keller, C., Klumpp, S., and Lipowsky, R. (2015). External forces influence the elastic coupling effects during cargo transport by molecular motors. *Phys Rev E*, 91(2):022701.

Bertossi, A. A. (1981). The edge Hamiltonian path problem is NP-complete. *Inform Process Lett*, 13(4):157–159.

Betzig, E., Patterson, G. H., Sougrat, R., Lindwasser, O. W., Olenych, S., Bonifacino, J. S., Davidson, M. W., Lippincott-Schwartz, J., and Hess, H. F. (2006). Imaging intracellular fluorescent proteins at nanometer resolution. *Science*, 313(5793):1642–1645.

Bianconi, G., Pin, P., and Marsili, M. (2009). Assessing the relevance of node features for network structure. *Proc Natl Acad Sci*, 106(28):11433–11438.

Biswas, S., Durocher, S., Mondal, D., and Nishat, R. I. (2012). Hamiltonian paths and cycles in planar graphs. In *Combinatorial Optimization and Applications*, pages 83–94. Springer, Berlin.

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., and Hwang, D.-U. (2006). Complex networks: structure and dynamics. *Phys Rep*, 424(4):175–308.

Boevink, P., Oparka, K., Santa Cruz, S., Martin, B., Betteridge, A., and Hawes, C. (1998). Stacks on tracks: the plant Golgi apparatus traffics on an actin/ER network. *Plant J*, 15(3):441–447.

Bohn, S., Andreotti, B., Douady, S., Munzinger, J., and Couder, Y. (2002). Constitutive property of the local organization of leaf venation networks. *Phys Rev E*, 65(6):061914.

Bond, J., Kofman, L., and Pogosyan, D. (1996). How filaments of galaxies are woven into the cosmic web. *Nature*, 380(6575):603–606.

Bond, N. A., Strauss, M. A., and Cen, R. (2010). Crawling the cosmic network: identifying and quantifying filamentary structure. *Mon Not R Astron Soc*, 409(1):156–168.

Borgatti, S. P. (2005). Centrality and network flow. *Soc Networks*, 27(1):55–71.

Bornholdt, S. and Röhl, T. (2003). Self-organized critical neural networks. *Phys Rev E*, 67:066118.

Bornholdt, S. and Schuster, H. G. (2002). *Handbook of Graphs and Networks*. Wiley, Berlin, 1st edition.

Boudaoud, A., Burian, A., Borowska-Wykret, D., Uyttewaal, M., Wrzalik, R., Kwiatkowska, D., and Hamant, O. (2014). FibrilTool, an ImageJ plug-in to quantify fibrillar structures in raw microscopy images. *Nat Protoc*, 9(2):457–463.

Boyd, S. (2006). Convex optimization of graph Laplacian eigenvalues. In *Proc Intl Cong Math*, volume 3, pages 1311–1319. Europ Math Soc Pub.

Braitenberg, V. and Schüz, A. (1998). *Cortex: Statistics and Geometry of Neuronal Connectivity*. Springer, Berlin, 1st edition.

Brandes, U. (2008). On variants of shortest-path betweenness centrality and their generic computation. *Soc Networks*, 30(2):136–145.

Brandes, U. and Fleischer, D. (2005). *Centrality measures based on current flow*. Springer, Berlin, 1st edition.

Brandner, D. and Withers, G. (2014). Development of the axon and dendritic arbors in cultured hippocampal neurons. CC-BY 3.0, http://www.cellimagelibrary.org/contributors/742096, Date of access: 07/09/2015.

Brešar, B., Kardoš, F., Katrenič, J., and Semanišin, G. (2011). Minimum k-path vertex cover. *Discrete Appl Math*, 159(12):1189–1195.

Breuer, D., Ivakov, A., Nowak, J., Persson, S., and Nikoloski, Z. (2015). System-wide organisation of the actin cytoskeleton drives organelle transport in plant cells. *submitted*.

Breuer, D., Ivakov, A., Sampathkumar, A., Hollandt, F., Persson, S., and Nikoloski, Z. (2014). Quantitative analyses of the plant cytoskeleton reveal underlying organizational principles. *J R Soc Interface*, 11(97):20140362.

Breuer, D. and Nikoloski, Z. (2014). img2net: automated network-based analysis of imaged phenotypes. *Bioinformatics*, 30(22):3291–3292.

Breuer, D. and Nikoloski, Z. (2015a). Cell type-specific organization and optimality of the plant actin cytoskeleton. *submitted*.

Breuer, D. and Nikoloski, Z. (2015b). Define: an optimisation-based method for robust disentangling of filamentous networks. *Sci Rep*, 5:18267.

Brin, S. and Page, L. (1998). The anatomy of a large-scale hypertextual web search engine. In *Proc 7th WWW Conf*, Geneva. IW3C2.

Bringmann, M., Landrein, B., Schudoma, C., Hamant, O., Hauser, M.-T., and Persson, S. (2012). Cracking the elusive alignment hypothesis: the microtubule-cellulose synthase nexus unraveled. *Trends Plant Sci*, 17(11):666–674.

Buchin, K., Knauer, C., Kriegel, K., Schulz, A., and Seidel, R. (2007). On the number of cycles in planar graphs. In *Computing and Combinatorics*, pages 97–107. Springer, Berlin.

Buchnik, L., Abu-Abied, M., and Sadot, E. (2014). Role of plant myosins in motile organelles: is a direct interaction required? *J Integr Plant Biol*, 57(1):23–30.

Burlacu, S., Janmey, P., and Borejdo, J. (1992). Distribution of actin filament lengths measured by fluorescence microscopy. *Am J Physiol-Cell Ph*, 262(3):C569–C577.

Buschmann, H., Green, P., Sambade, A., Doonan, J., and Lloyd, C. (2011). Cytoskeletal dynamics in interphase, mitosis and cytokinesis analysed through agrobacterium-mediated transient transformation of tobacco BY-2 cells. *New Phytol*, 190(1):258–267.

Byrd, R. H., Lu, P., Nocedal, J., and Zhu, C. (1995). A limited memory algorithm for bound constrained optimization. *SIAM J Sci Comput*, 16(5):1190–1208.

Cai, C., Henty-Ridilla, J. L., Szymanski, D. B., and Staiger, C. J. (2014). Arabidopsis myosin XI: a motor rules the tracks. *Plant Physiol*, 166(3):1359–1370.

Cai, G. and Cresti, M. (2012). Are kinesins required for organelle trafficking in plant cells? *Front Plant Sci*, 3:170.

Cai, G., Parrotta, L., and Cresti, M. (2015). Organelle trafficking, the cytoskeleton, and pollen tube growth. *J Integr Plant Biol*, 57(1):63–78.

Caldarelli, G., Capocci, A., De Los Rios, P., and Munoz, M. A. (2002). Scale-free networks from varying vertex intrinsic fitness. *Phys Rev Lett*, 89(25):258702.

Carlsson, A. E. (2010). Actin dynamics: from nanoscale to microscale. *Annu Rev Biophys*, 39:91–110.

Caviston, J. P. and Holzbaur, E. L. F. (2006). Microtubule motors at the intersection of trafficking and transport. *Trends Cell Biol*, 16(10):530–537.

Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W., and Prasher, D. C. (1994). Green fluorescent protein as a marker for gene expression. *Science*, 263(5148):802–805.

Chen, X., Irani, N. G., and Friml, J. (2011). Clathrin-mediated endocytosis: the gateway into plant cells. *Curr Opin Plant Biol*, 14(6):674–682.

Cheung, A. Y. and de Vries, S. C. (2008). Membrane trafficking: intracellular highways and country roads. *Plant Physiol*, 147(4):1451–1453.

Chevalier-Larsen, E. and Holzbaur, E. L. (2006). Axonal transport and neurodegenerative disease. *Biochim Biophys*, 1762(11):1094–1108.

Claessens, M. M., Bathe, M., Frey, E., and Bausch, A. R. (2006). Actin-binding proteins sensitively mediate F-actin bundle stiffness. *Nat Mater*, 5(9):748–753.

Clark, A. G., Dierkes, K., and Paluch, E. K. (2013). Monitoring actin cortex thickness in live cells. *Biophys J*, 105(3):570–580.

Coen, M. H., Ansari, M. H., and Fillmore, N. (2010). Comparing clusterings in space. In *Proc 27th Internat Conf Mach Learn*, pages 231–238, Madison. Omnipress.

Cohen, A. R., Roysam, B., and Turner, J. N. (1994). Automated tracing and volume measurements of neurons from 3D confocal fluorescence microscopy data. *J Microsc*, 173(2):103–114.

Cohen, E. and Tarsi, M. (1991). NP-completeness of graph decomposition problems. *J Complexity*, 7(2):200–212.

Collings, D. (2008). Crossed-wires: interactions and cross-talk between the microtubule and microfilament networks in plants. In *Plant microtubules*, pages 47–79. Springer, Berlin.

Cooper, G. M. (2000). *The cell: a molecular approach*. ASM Press, Boston, 2nd edition.

Cosgrove, D. J. (1996). Plant cell enlargement and the action of expansins. *Bioessays*, 18(7):533–540.

Courtat, T., Douady, S., and Gloaguen, C. (2011). Centrality maps and the analysis of city street networks. In *Proc 5th Internat ICST Conf Perform Eval Method Tools*, pages 316–321, Brussels. ICST.

Cover, T. M. and Hart, P. E. (1967). Nearest neighbor pattern classification. *IEEE Trans Inf Theory*, 13(1):21–27.

Croton, D. J. (2013). Damn you, little h! (or, real-world applications of the hubble constant using observed and simulated data). *Publ Astron Soc Aust*, 30:e052.

Crowell, E. F., Bischoff, V., Desprez, T., Rolland, A., Stierhof, Y.-D., Schumacher, K., Gonneau, M., Höfte, H., and Vernhettes, S. (2009). Pausing of Golgi bodies on microtubules regulates secretion of cellulose synthase complexes in Arabidopsis. *Plant Cell Online*, 21(4):1141–1154.

Dahl, J. and Vandenberghe, L. (2006). Cvxopt: a Python package for convex optimization. http://abel.ee.ucla.edu/cvxopt/, Date of access: 07/09/2015.

Denœud, L. and Guénoche, A. (2006). Comparison of distance indices between partitions. In *Data Science and Classification*, pages 21–28. Springer, Berlin.

Dhondt, S., Van Haerenborgh, D., Van Cauwenbergh, C., Merks, R. M. H., Philips, W., Beemster, G. T. S., and Inzé, D. (2012). Quantitative analysis of venation patterns of Arabidopsis leaves by supervised image analysis. *Plant J*, 69(3):553–563.

Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. *Numer Math*, 1(1):269–271.

Dodds, P. S., Corson, F., Katifori, E., Szöllősi, G. J., and Magnasco, M. O. (2010). Focus: why leaves aren't trees. *Phys Rev Foc*, 25:4.

Dominguez, R. (2004). Actin-binding proteins: a unifying hypothesis. *Trends Biochem Sci*, 29(11):572–578.

Dorogovtsev, S. N., Mendes, J. F., and Samukhin, A. N. (2000). Structure of growing networks with preferential linking. *Phys Rev Lett*, 85(21):4633–4636.

Dujardin, F. (1835). Recherches sur les organismes inférieurs. *Ann Sci Nat Zool Second Series*, 4:343–377.

Dupree, P. and Sherrier, D. J. (1998). The plant Golgi apparatus. *Biochim Biophys Acta*, 1404(1):259–270.

Eccles, J. C. (1982). The synapse: from electrical to chemical transmission. *Annu Rev Neurosci*, 5(1):325–339.

Edwards, A. M., Phillips, R. A., Watkins, N. W., Freeman, M. P., Murphy, E. J., Afanasyev, V., Buldyrev, S. V., da Luz, M. G. E., Raposo, E. P., Stanley, H. E., and Viswanathan, G. M. (2007). Revisiting Lévy flight search patterns of wandering albatrosses, bumblebees and deer. *Nature*, 449(7165):1044–1048.

Ehrgott, M. (2006). *Multicriteria optimization*. Springer Science, New York, 1st edition.

Ehrhardt, D. W. (2008). Straighten up and fly right: microtubule dynamics and organization of non-centrosomal arrays in higher plants. *Curr Opin Cell Biol*, 20(1):107–116.

Ehrhardt, D. W. and Shaw, S. L. (2006). Microtubule dynamics and organization in the plant cortical array. *Annu Rev Plant Biol*, 57:859–875.

Eliceiri, K. W., Berthold, M. R., Goldberg, I. G., Ibanez, L., Manjunath, B. S., Martone, M. E., Murphy, R. F., Peng, H., Plant, A. L., Roysam, B., Stuurman, N., Stuurmann, N., Swedlow, J. R., Tomancak, P., and Carpenter, A. E. (2012). Biological imaging software tools. *Nat Methods*, 9(7):697–710.

Ellis, R. J. (2001). Macromolecular crowding: an important but neglected aspect of the intracellular environment. *Curr Opin Struct Biol*, 11(1):114–119.

Ermentrout, G. B. and Edelstein-Keshet, L. (1998). Models for the length distributions of actin filaments: II. polymerization and fragmentation by gelsolin acting together. *B Math Biol*, 60(3):477–503.

Even, G., Garg, N., Könemann, J., Ravi, R., and Sinha, A. (2004). Min–max tree covers of graphs. *Op Res Lett*, 32(4):309–315.

Faltenbacher, A., Gottlöber, S., Kerscher, M., and Mueller, V. (2002). Correlations in the orientations of galaxy clusters. *Astron Astrophys*, 395(1):1–9.

Fell, D. A. and Wagner, A. (2000). The small world of metabolism. *Nat Biotechnol*, 18(11):1121–1122.

Felleman, D. J. and Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cereb Cortex*, 1(1):1–47.

Finlay, J. (2005). PyGTK 2.0 Tutorial. http://www.pygtk.org/dist/pygtk2-tut.pdf, Date of access: 07/09/2015.

Fleischer, F., Ananthakrishnan, R., Eckel, S., Schmidt, H., Käs, J., Svitkina, T., Schmidt, V., and Beil, M. (2007). Actin network architecture and elasticity in lamellipodia of melanoma cells. *New J Phys*, 9(11):420.

Fletcher, D. A. and Mullins, R. D. (2010). Cell mechanics and the cytoskeleton. *Nature*, 463(7280):485–492.

Foissner, I. (2004). Microfilaments and microtubules control the shape, motility, and subcellular distribution of cortical mitochondria in characean internodal cells. *Protoplasma*, 224(3-4):145–157.

Forgacs, G. (1995). On the possible role of cytoskeletal filamentous networks in intracellular signaling: an approach based on percolation. *J Cell Sci*, 108(6):2131–2143.

Frangi, A. F., Niessen, W. J., Vincken, K. L., and Viergever, M. A. (1998). Multiscale vessel enhancement filtering. In *Med Image Comput Comput-Ass Interv*, pages 130–137. Springer, Berlin.

Freedman, D. A. (2009). *Statistical models: theory and practice*. Cambridge Univ Press, Cambridge, 1st edition.

Freeman, L. C. (1977). A set of measures of centrality based on betweenness. *Sociometry*, 40(1):35–41.

Fu, Y., Gu, Y., Zheng, Z., Wasteneys, G., and Yang, Z. (2005). Arabidopsis interdigitating cell growth requires two antagonistic pathways with opposing action on cell morphogenesis. *Cell*, 120(5):687–700.

Gadeyne, A., Sánchez-Rodríguez, C., Vanneste, S., Di Rubbo, S., Zauber, H., Vanneste, K., Van Leene, J., De Winne, N., Eeckhout, D., Persiau, G., Van De Slijke, E., Cannoot, B., Vercruysse, L., Mayers, J., Adamowski, M., Kania, U., Ehrlich, M., Schweighofer, A., Ketelaar, T., Maere, S., Bednarek, S., Friml, J., Gevaert, K., Witters, E., Russinova, E., Persson, S., De Jaeger, G., and Van Damme, D. (2014). The TPLATE adaptor complex drives clathrin-mediated endocytosis in plants. *Cell*, 156(4):691–704.

Galkovskyi, T., Mileyko, Y., Bucksch, A., Moore, B., Symonova, O., Price, C. A., Topp, C. N., Iyer-Pascuzzi, A. S., Zurek, P. R., Fang, S., et al. (2012). GiA roots: software for the high throughput analysis of plant root system architecture. *BMC Plant Biol*, 12(1):116.

Ganguly, S., Williams, L. S., Palacios, I. M., and Goldstein, R. E. (2012). Cytoplasmic streaming in Drosophila oocytes varies with kinesin activity and correlates with the microtubule cytoskeleton architecture. *Proc Natl Acad Sci*, 109(38):15109–15114.

Gardel, M. L., Shin, J. H., MacKintosh, F. C., Mahadevan, L., Matsudaira, P., and Weitz, D. A. (2004). Elastic behavior of cross-linked and bundled actin networks. *Science*, 304(5675):1301–1305.

Garey, M. R. and Johnson, D. S. (1979). *Computers and intractability: a guide to NP-completeness*. WH Freeman, New York, 1st edition.

Garey, M. R., Johnson, D. S., and Tarjan, R. E. (1976). The planar Hamiltonian circuit problem is NP-complete. *SIAM J Comput*, 5(4):704–714.

Gazit, Y., Berk, D. A., Leunig, M., Baxter, L. T., and Jain, R. K. (1995). Scale-invariant behavior and vascular network formation in normal and tumor tissue. *Phys Rev Lett*, 75(12):2428–2431.

Geisler, D. A., Sampathkumar, A., Mutwil, M., and Persson, S. (2008). Laying down the bricks: logistic aspects of cell wall biosynthesis. *Curr Opin Plant Biol*, 11(6):647–652.

Gendreau, E., Traas, J., Desnos, T., Grandjean, O., Caboche, M., and Höfte, H. (1997). Cellular basis of hypocotyl growth in Arabidopsis thaliana. *Plant Physiol*, 114(1):295–305.

Gibbon, B. C., Kovar, D. R., and Staiger, C. J. (1999). Latrunculin B has different effects on pollen germination and tube growth. *Plant Cell*, 11(12):2349–2363.

Gibeaut, D. and Carpita, N. (1994). Biosynthesis of plant cell wall polysaccharides. *FASEB J*, 8(12):904–915.

Gibson, S. and Fan, Y. (2006). Coronal prominence structure and dynamics: a magnetic flux rope interpretation. *J Geophys Res-Space*, 111(A12).

Gittes, F., Mickey, B., Nettleton, J., and Howard, J. (1993). Flexural rigidity of microtubules and actin filaments measured from thermal fluctuations in shape. *J Cell Biol*, 120(4):923–934.

Goldberg, M. K., Hayvanovych, M., and Magdon-Ismail, M. (2010). Measuring similarity between sets of overlapping clusters. In *IEEE 2nd Internat Conf Social Comput*, pages 303–308, Minneapolis. IEEE.

Goldstein, R. E., Tuval, I., and van de Meent, J.-W. (2008). Microfluidics of cytoplasmic streaming and its implications for intracellular transport. *Proc Natl Acad Sci*, 105(10):3663–3667.

Goldstein, R. E. and van de Meent, J.-W. (2015). A physical perspective on cytoplasmic streaming. *J R Soc Interface Focus*, 5(4):20150030.

Gonzalez, N., Vanhaeren, H., and Inzé, D. (2012). Leaf size control: complex coordination of cell division and expansion. *Trends Plant Sci*, 17(6):332–340.

Gonzalez, R. C. and Woods, R. E. (2009). *Digital image processing*. Pearson Edu India, New Delhi, 3rd edition.

Goode, B. L., Drubin, D. G., and Barnes, G. (2000). Functional cooperation between the microtubule and actin cytoskeletons. *Curr Opin Cell Biol*, 12(1):63–71.

Granger, C. and Cyr, R. (2001). Spatiotemporal relationships between growth and microtubule orientation as revealed in living root cells of Arabidopsis thaliana transformed with green-fluorescent-protein gene construct GFP-MBD. *Protoplasma*, 216(3):201–214.

Griffing, L. R. (1991). Comparisons of Golgi structure and dynamics in plant and animal cells. *J Electron Microsc Tech*, 17(2):179–199.

Griva, I., Nash, S. G., and Sofer, A. (2009). *Linear and nonlinear optimization*. SIAM, Philadelphia, 2nd edition.

Gross, S. P. (2004). Hither and yon: a review of bi-directional microtubule-based transport. *Phys Biol*, 1(2):1–11.

Guérin, T., Prost, J., Martin, P., and Joanny, J.-F. (2010). Coordination and collective properties of molecular motors: theory. *Curr Opin Cell Biol*, 22(1):14–20.

Guimerà, R., Mossa, S., Turtschi, A., and Amaral, L. N. (2005). The worldwide air transportation network: anomalous centrality, community structure, and cities' global roles. *Proc Natl Acad Sci*, 102(22):7794–7799.

Gunning, P. W., Ghoshdastider, U., Whitaker, S., Popp, D., and Robinson, R. C. (2015). The evolution of compositionally and functionally distinct actin filaments. *J Cell Sci*, 128(11):2009–2019.

Gutierrez, R., Lindeboom, J. J., Paredez, A. R., Emons, A. M. C., and Ehrhardt, D. W. (2009). Arabidopsis cortical microtubules position cellulose synthase delivery to the plasma membrane and interact with cellulose synthase trafficking compartments. *Nat Cell Biol*, 11(7):797–806.

Hagberg, A., Swart, P., and S Chult, D. (2008). Exploring network structure, dynamics, and function using NetworkX. In *Proc 7th Python Sci Conf*, Pasadena. LANL.

Han, J., Kamber, M., and Pei, J. (2011). *Data mining: concepts and techniques*. Elsevier, Waltham, 3rd edition.

Handl, J., Kell, D. B., and Knowles, J. (2007). Multiobjective optimization in bioinformatics and computational biology. *IEEE Trans Comp Biol Bioinformatics*, 4(2):279–292.

Haralick, R. M., Sternberg, S. R., and Zhuang, X. (1987). Image analysis using mathematical morphology. *IEEE Trans Pattern Anal Mach Intell*, 4:532–550.

Hargreaves, A. J., Goodbody, K. C., and Lloyd, C. W. (1989). Reconstitution of intermediate filaments from a higher plant. *Biochem J*, 261:679–682.

Harris, T. and Ross, F. (1955). *Fundamentals of a method for evaluating rail net capacities*. NTIS, Alexandria, 1st edition.

Hartmann, L. and Börner, H. (2009). Precision polymers: monodisperse, monomer-sequence-defined segments to target future demands of polymers in medicine. *Adv Mater*, 21(32-33):3425–3431.

Haviv, L., Gov, N., Ideses, Y., and Bernheim-Groswasser, A. (2008). Thickness distribution of actin bundles in vitro. *Eur Biophys J*, 37(4):447–454.

Hawes, C. (2004). Cell biology of the plant Golgi apparatus. *New Phytol*, 165:29–44.

Hawes, C. and Satiat-Jeunemaitre, B. (2005). The plant Golgi apparatus: going with the flow. *Biochim Biophys Acta*, 1744(3):466–480.

Hawkins, R. J., Tindemans, S. H., and Mulder, B. M. (2010). Model for the orientational ordering of the plant microtubule cortical array. *Phys Rev E*, 82(1):011911.

Hayashi, Y. (2006). A review of recent studies of geographical scale-free networks. *IPSJ Digital Courier*, 2:155–164.

Heaton, L., Obara, B., Grau, V., Jones, N., Nakagaki, T., Boddy, L., and Fricker, M. D. (2012). Analysis of fungal networks. *Fungal Biol Rev*, 26(1):12–29.

Hell, S. W. and Wichmann, J. (1994). Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. *Opt Lett*, 19(11):780–782.

Hemminger, R. L. and Beineke, L. W. (1978). Line graphs and line digraphs. *STGT*, 1:291–305.

Henty-Ridilla, J. L., Li, J., Blanchoin, L., and Staiger, C. J. (2013). Actin dynamics in the cortical array of plant cells. *Curr Opin Plant Biol*, 16(6):678–687.

Himsolt, M. (1997). GML: a portable graph file format. http://www.fmi.uni-passau.de/graphlet/gml/gml-tr.html, Date of access: 07/09/2015.

Holweg, C. and Nick, P. (2004). Arabidopsis myosin XI mutant is defective in organelle movement and polar auxin transport. *Proc Natl Acad Sci*, 101(28):10488–10493.

Horak, P. and McAvaney, K. (2008). On covering vertices of a graph by trees. *Discrete Math*, 308(19):4414–4418.

Houtman, D., Pagonabarraga, I., Lowe, C., Esseling-Ozdoba, A., Emons, A., and Eiser, E. (2007). Hydrodynamic flow caused by active transport along cytoskeletal elements. *Europhys Lett*, 78(1):18001.

Hubert, L. and Arabie, P. (1985). Comparing partitions. *J Classif*, 2(1):193–218.

Humphries, N. E. and Sims, D. W. (2014). Optimal foraging strategies: Lévy walks balance searching and patch exploitation under a very broad range of conditions. *J Theor Biol*, 358:179–193.

Humphries, N. E., Weimerskirch, H., Queiroz, N., Southall, E. J., and Sims, D. W. (2012). Foraging success of biological Lévy flights recorded in situ. *Proc Natl Acad Sci*, 109(19):7169–7174.

Hunter, J. D. (2007). Matplotlib: a 2D graphics environment. *Comput Sci Eng*, 9(3):90–95.

Hush, J., Wadsworth, P., Callaham, D., and Hepler, P. (1994). Quantification of microtubule dynamics in living plant cells using fluorescence redistribution after photobleaching. *J Cell Sci*, 107(4):775–784.

Hussey, P. J., Ketelaar, T., and Deeks, M. J. (2006). Control of the actin cytoskeleton in plant cell growth. *Annu Rev Plant Biol*, 57:109–125.

Ingber, D. E. (1998). The architecture of life. *Sci Am*, 278(1):48–57.

Inoue, K. (2000). Functional dendrimers, hyperbranched and star polymers. *Prog Polym Sci*, 25(4):453–571.

Ivakov, A. and Persson, S. (2013). Plant cell shape: modulators and measurements. *Front Plant Sci*, 4:439.

Jacques, E., Buytaert, J., Wells, D. M., Lewandowski, M., Bennett, M. J., Dirckx, J., Verbelen, J.-P., and Vissenberg, K. (2013a). Microfilament Analyzer, an image analysis tool for quantifying fibrillar orientation, reveals changes in microtubule organization during gravitropism. *Plant J*, 74(6):1045–1058.

Jacques, E., Lewandowski, M., Buytaert, J., Fierens, Y., Verbelen, J.-P., and Vissenberg, K. (2013b). Microfilament Analyzer identifies actin network organizations in epidermal cells of Arabidopsis thaliana roots. *Plant Signal Behav*, 8(7):e24821.

Jaqaman, K., Loerke, D., Mettlen, M., Kuwata, H., Grinstein, S., Schmid, S. L., and Danuser, G. (2008). Robust single-particle tracking in live-cell time-lapse sequences. *Nat Methods*, 5(8):695–702.

Jensen, P. A. and Barnes, J. W. (1980). *Network flow programming*. Wiley, New York, 1st edition.

Jeong, H., Tombor, B., Albert, R., Oltvai, Z., and Barabási, A.-L. (2000). The large–scale organization of metabolic networks. *Nature*, 407:651–654.

Jiang, B., Zhao, S., and Yin, J. (2008). Self-organized natural roads for predicting traffic flow: a sensitivity study. *J Stat Mech*, 2008(07):P07008.

Jordano, P., Bascompte, J., and Olesen, J. M. (2003). Invariant properties in coevolutionary networks of plant–animal interactions. *Ecol Lett*, 6(1):69–81.

Kachar, B. and Reese, T. S. (1988). The mechanism of cytoplasmic streaming in characean algal cells: sliding of endoplasmic reticulum along actin filaments. *J Cell Biol*, 106(5):1545–1552.

Kandel, E., Schwartz, J., and Jessell, T. (2000). *Principles of Neural Science*. McGraw-Hill, New York, 4 edition.

Kapitein, L. C. and Hoogenraad, C. C. (2011). Which way to go? Cytoskeletal organization and polarized transport in neurons. *Mol Cell Neurosci*, 46(1):9–20.

Karp, R. M. (1972). *Reducibility among combinatorial problems*. Springer, New York, 1st edition.

Kasahara, M., Kagawa, T., Oikawa, K., Suetsugu, N., Miyao, M., and Wada, M. (2002). Chloroplast avoidance movement reduces photodamage in plants. *Nature*, 420(6917):829–832.

Kasza, K. E., Broedersz, C. P., Koenderink, G. H., Lin, Y. C., Messner, W., Millman, E. A., Nakamura, F., Stossel, T. P., Mackintosh, F. C., and Weitz, D. A. (2010). Actin filament length tunes elasticity of flexibly cross-linked actin networks. *Biophys J*, 99(4):1091–1100.

Katifori, E. and Magnasco, M. O. (2012). Quantifying loopy network architectures. *PLoS One*, 7(6):e37994.

Katifori, E., Szöllősi, G. J., and Magnasco, M. O. (2010). Damage and fluctuations induce loops in optimal transport networks. *Phys Rev Lett*, 104(4):048704.

Kazerani, A. and Winter, S. (2009). Can betweenness centrality explain traffic flow? In *Proc 12th AGILE Int Conf GIS*, Hannover. Leibniz Universität Hannover.

Kim, H., Park, M., Kim, S. J., and Hwang, I. (2005). Actin filaments play a critical role in vacuolar trafficking at the Golgi complex in plant cells. *Plant Cell*, 17(3):888–902.

Kim, S. and Coulombe, P. A. (2010). Emerging role for the cytoskeleton as an organizer and regulator of translation. *Nat Rev Mol Cell Biol*, 11(1):75–81.

Kitano, H. (2002). Systems biology: a brief overview. *Science*, 295(5560):1662–1664.

Klemm, D., Heublein, B., Fink, H.-P., and Bohn, A. (2005). Cellulose: fascinating biopolymer and sustainable raw material. *Angew Chem Int Ed*, 44(22):3358–3393.

Klumpp, S. and Lipowsky, R. (2005). Cooperative cargo transport by several molecular motors. *Proc Natl Acad Sci*, 102(48):17284–17289.

Koch, C. and Laurent, G. (1999). Complexity and the nervous system. *Science*, 284(5411):96–98.

Kost, B. and Chua, N.-H. (2002). The plant cytoskeleton: vacuoles and cell walls make the difference. *Cell*, 108(1):9–12.

Kuhn, H. W. (1955). The Hungarian method for the assignment problem. *Nav Res Logist Q*, 2(1):83–97.

Kumar, S., Maxwell, I. Z., Heisterkamp, A., Polte, T. R., Lele, T. P., Salanga, M., Mazur, E., and Ingber, D. E. (2006). Viscoelastic retraction of single living stress fibers and its impact on cell shape, cytoskeletal organization, and extracellular matrix mechanics. *Biophys J*, 90(10):3762–3773.

Kurant, M. and Thiran, P. (2006). Layered complex networks. *Phys Rev Lett*, 96(13):138701.

Lancichinetti, A. and Fortunato, S. (2009a). Benchmarks for testing community detection algorithms on directed and weighted graphs with overlapping communities. *Phys Rev E*, 80(1):016118.

Lancichinetti, A. and Fortunato, S. (2009b). Community detection algorithms: a comparative analysis. *Phys Rev E*, 80(5):056117.

Lancichinetti, A., Fortunato, S., and Kertész, J. (2009). Detecting the overlapping and hierarchical community structure in complex networks. *New J Phys*, 11(3):033015.

Landrein, B., Lathe, R., Bringmann, M., Vouillot, C., Ivakov, A., Boudaoud, A., Persson, S., and Hamant, O. (2013). Impaired cellulose synthase guidance leads to stem torsion and twists phyllotactic patterns in Arabidopsis. *Curr Biol*, 23(10):895–900.

Langford, G. M. (1995). Actin- and microtubule-dependent organelle motors: interrelationships between the two motility systems. *Curr Opin Cell Biol*, 7(1):82–88.

Langville, A. N. and Meyer, C. D. (2005). A survey of eigenvector methods for web information retrieval. *SIAM Rev*, 47(1):135–161.

Le, J., Vandenbussche, F., De Cnodder, T., Van Der Straeten, D., and Verbelen, J.-P. (2005). Cell elongation and microtubule behavior in the arabidopsis hypocotyl: responses to ethylene and auxin. *J Plant Growth Regul*, 24(3):166–178.

Leandro, J. J. G., Cesar-Jr, R. M., and Costa, L. d. (2009). Automatic contour extraction from 2D neuron images. *J Neurosci Methods*, 177(2):497–509.

Lee, M. C. S., Miller, E. A., Goldberg, J., Orci, L., and Schekman, R. (2004). Bi-directional protein transport between the ER and Golgi. *Annu Rev Cell Dev Biol*, 20:87–123.

Lew, R. R. (2005). Mass flow and pressure-driven hyphal extension in Neurospora crassa. *MicroBiol*, 151(8):2685–2692.

Li, D., Shao, L., Chen, B.-C., Zhang, X., Zhang, M., Moses, B., Milkie, D. E., Beach, J. R., Hammer, J. A., Pasham, M., Kirchhausen, T., Baird, M. A., Davidson, M. W., Xu, P., and Betzig, E. (2015). Extended-resolution structured illumination imaging of endocytic and cytoskeletal dynamics. *Science*, 349(6251):aab3500.

Lichtman, J., Livet, J., and Sanes, J. (2008). A technicolour approach to the connectome. *Nat Rev Neurosci*, 9:417–422.

Lieleg, O., Claessens, M., Heussinger, C., Frey, E., and Bausch, A. (2007). Mechanics of bundled semiflexible polymer networks. *Phys Rev Lett*, 99(8):088102.

Lin, C. and Shyu, T.-W. (1996). A necessary and sufficient condition for the star decomposition of complete graphs. *J Graph Theor*, 23(4):361–364.

Lin, G., Cai, Z., and Lin, D. (2006). Vertex covering by paths on trees with its applications in machine translation. *Inform Process Lett*, 97(2):73–81.

Lin, R., Olariu, S., and Pruesse, G. (1995). An optimal path cover algorithm for cographs. *Comput Math Appl*, 30(8):75–83.

Lindeboom, J. J., Nakamura, M., Hibbel, A., Shundyak, K., Gutierrez, R., Ketelaar, T., Emons, A. M. C., Mulder, B. M., Kirik, V., and Ehrhardt, D. W. (2013). A mechanism for reorientation of cortical microtubule arrays driven by microtubule severing. *Science*, 342(6163):1245533.

Linderoth, J. T. and Ralphs, T. K. (2005). Noncommercial software for mixed-integer linear programming. *Int Prog Theor Pract*, 3:253–303.

Lippincott-Schwartz, J. and Patterson, G. H. (2003). Development and use of fluorescent protein markers in living cells. *Science*, 300(5616):87–91.

Liu, B. (2010). *The plant cytoskeleton*. Springer, New York, 2nd edition.

Lobet, G., Pound, M. P., Diener, J., Pradal, C., Draye, X., Godin, C., Javaux, M., Leitner, D., Meunier, F., Nacry, P., et al. (2015). Root System Markup Language: toward a unified root architecture description language. *Plant Physiol*, 167(3):617–627.

Lodish, H. F., , and Zipursky, S. L. (2000). *Molecular cell biology*. WH Freeman, New York, 4th edition.

Longair, M. H., Baker, D. A., and Armstrong, J. D. (2011). Simple neurite tracer: open source software for reconstruction, visualization and analysis of neuronal processes. *Method Biochem Anal*, 27(17):2453–2454.

Louf, R., Jensen, P., and Barthélemy, M. (2013). Emergence of hierarchy in cost-driven growth of spatial networks. *Proc Natl Acad Sci*, 110(22):8824–8829.

Lu, L., Oswald, S. J., Ngu, H., and Yin, F. C.-P. (2008). Mechanical properties of actin stress fibers in living cells. *Biophys J*, 95(12):6060–6071.

Luby-Phelps, K. (2000). Cytoarchitecture and physical properties of cytoplasm: volume, viscosity, diffusion, intracellular surface area. *Int Rev Cytol*, 192:189–221.

Luo, Y., Scholl, S., Doering, A., Zhang, Y., Irani, N. G., Rubbo, S. D., Neumetzler, L., Krishnamoorthy, P., Houtte, I. V., Mylle, E., Bischoff, V., Vernhettes, S., Winne, J., Friml, J., Stierhof, Y.-D., Schumacher, K., Persson, S., and Russinova, E. (2015). V-ATPase activity in the TGN/EE is required for exocytosis and recycling in Arabidopsis. *Nat Plants*, 1(7).

Mackay, D., Karpen, J., Ballester, J., Schmieder, B., and Aulanier, G. (2010). Physics of solar prominences: II. magnetic structure and dynamics. *Space Sci Rev*, 151(4):333–399.

MacKintosh, F., Käs, J., and Janmey, P. (1995). Elasticity of semiflexible biopolymer networks. *Phys Rev Lett*, 75(24):4425.

Marieb, E. N. and Hoehn, K. (2007). *Human anatomy and physiology*. Pearson Edu, San Francisco, 7th edition.

Masland, R. H. (2001). The fundamental plan of the retina. *Nat Neurosci*, 4(9):877–886.

Mathur, J. (2004). Plant cytoskeleton: reinforcing lines of division in plant cells. *Curr Biol*, 14(7):R287–R289.

Matthäus, F., Jagodic, M., and Dobnikar, J. (2009). E. coli superdiffusion and chemotaxis-search strategy, precision, and motility. *Biophys J*, 97(4):946–957.

Mayerich, D., Bjornsson, C., Taylor, J., and Roysam, B. (2012). NetMets: software for quantifying and visualizing errors in biological network segmentation. *BMC Bioinformatics*, 13(8):S7.

Mayerich, D. M. and Keyser, J. (2008). Filament tracking and encoding for complex biological networks. In *Proc 2008 ACM Symp Solid Phys Model*, pages 353–358, New York. ACM.

McCarthy, M. I., Abecasis, G. R., Cardon, L. R., Goldstein, D. B., Little, J., Ioannidis, J. P., and Hirschhorn, J. N. (2008). Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat Rev Gen*, 9(5):356–369.

McGrath, J. L., Osborn, E. A., Tardy, Y. S., Dewey, C. F., and Hartwig, J. H. (2000). Regulation of the actin cycle in vivo by actin filament severing. *Proc Natl Acad Sci*, 97(12):6532–6537.

McGrath, J. L., Tardy, Y., Dewey, C. F., Meister, J. J., and Hartwig, J. H. (1998). Simultaneous measurements of actin filament turnover, filament fraction, and monomer diffusion in endothelial cells. *Biophys J*, 75(4):2070–2078.

Meijering, E. (2010). Neuron tracing in perspective. *Cytometry A*, 77(7):693–704.

Meilă, M. (2003). Comparing clusterings by the variation of information. In *Learning Theory and Kernel Machines*, pages 173–187. Springer, Berlin.

Meilă, M. (2005). Comparing clusterings: an axiomatic view. In *Proc 22nd Internat Conf Mach Learn*, pages 577–584, New York. ACM.

Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., and Alon, U. (2002). Network motifs: Simple building blocks of complex networks. *Science*, 298:824–827.

Mitchell-Olds, T. (2010). Complex-trait analysis in plants. *Genome Biol*, 11(113):423.

Moon, R. J., Martini, A., Nairn, J., Simonsen, J., and Youngblood, J. (2011). Cellulose nanomaterials review: structure, properties and nanocomposites. *Chem Soc Rev*, 40(7):3941–3994.

Moseley, J. B. (2013). An expanded view of the eukaryotic cytoskeleton. *Mol Biol Cell*, 24(11):1615–1618.

Müller, M. J. I., Klumpp, S., and Lipowsky, R. (2008). Tug-of-war as a cooperative mechanism for bidirectional cargo transport by molecular motors. *Proc Natl Acad Sci*, 105(12):4609–4614.

Nebenführ, A., Gallagher, L. A., Dunahay, T. G., Frohlick, J. A., Mazurkiewicz, A. M., Meehl, J. B., and Staehelin, L. A. (1999). Stop-and-go movements of plant Golgi stacks are mediated by the acto-myosin system. *Plant Physiol*, 121(4):1127–1141.

Nebenführ, A. and Staehelin, L. A. (2001). Mobile factories: Golgi dynamics in plant cells. *Trends Plant Sci*, 6(4):160–167.

Neri, I., Kern, N., and Parmeggiani, A. (2013). Modeling cytoskeletal traffic: an interplay between passive diffusion and active transport. *Phys Rev Lett*, 110(9):098102.

Newman, M. (2009). *Networks: an introduction*. Oxford Univ Press, Oxford, 1st edition.

Newman, M. (2012). Communities, modules and large-scale structure in networks. *Nat Phys*, 8(1):25–31.

Nick, P. (2007). The plant cytoskeleton - new jobs for a versatile network. *Protoplasma*, 230(3):125–127.

North, A. (2006). Seeing is believing? A beginners' guide to practical pitfalls in image acquisition. *J Cell Biol*, 172(1):9–18.

Obara, B., Fricker, M., Gavaghan, D., and Grau, V. (2012a). Contrast-independent curvilinear structure detection in biomedical images. *IEEE Trans Image Process*, 21(5):2572–2581.

Obara, B., Grau, V., and Fricker, M. D. (2012b). A bioimage informatics approach to automatically extract complex fungal networks. *Bioinformatics*, 28(18):2374–2381.

Oliphant, T. E. (2006). *A Guide to NumPy*. Trelgol Pub, Spanish Fork, 1st edition.

Olivier, B., Rohwer, J., and Hofmeyr, J.-H. (2002). Modelling cellular processes with Python and Scipy. *Mol Biol Rep*, 29(1):249–254.

Onoda, Y., Westoby, M., Adler, P. B., Choong, A. M., Clissold, F. J., Cornelissen, J. H., Díaz, S., Dominy, N. J., Elgart, A., Enrico, L., et al. (2011). Global patterns of leaf mechanical properties. *Eco Lett*, 14(3):301–312.

Osunbayo, O., Butterfield, J., Bergman, J., Mershon, L., Rodionov, V., and Vershinin, M. (2015). Cargo transport at microtubule crossings: evidence for prolonged tug-of-war between kinesin motors. *Biophys J*, 108(6):1480–1483.

Otsu, N. (1975). A threshold selection method from gray-level histograms. *Automatica*, 11(285):23–27.

Pak-Ken, W. (1999). Optimal path cover problem on block graphs. *Lect Notes Comput Sc*, 225(1):163–169.

Paradez, A., Wright, A., and Ehrhardt, D. W. (2006). Microtubule cortical array organization and plant cell morphogenesis. *Curr Opin Plant Biol*, 9(6):571–578.

Paredez, A. R., Somerville, C. R., and Ehrhardt, D. W. (2006). Visualization of cellulose synthase demonstrates functional association with microtubules. *Science*, 312(5779):1491–1495.

Pastor-Satorras, R., Vázquez, A., and Vespignani, A. (2001). Dynamical and correlation properties of the internet. *Phys Rev Lett*, 87(25):258701.

Peng, H., Hawrylycz, M., Roskams, J., Hill, S., Spruston, N., Meijering, E., and Ascoli, G. A. (2015a). BigNeuron: large-scale 3D neuron reconstruction from optical microscopy images. *Neuron*, 87(2):252–256.

Peng, H., Meijering, E., and Ascoli, G. A. (2015b). From DIADEM to BigNeuron. *Neuroinformatics*, 13(3):259–260.

Peremyslov, V. V., Mockler, T. C., Filichkin, S. A., Fox, S. E., Jaiswal, P., Makarova, K. S., Koonin, E. V., and Dolja, V. V. (2011). Expression, splicing, and evolution of the myosin gene family in plants. *Plant Physiol*, 155(3):1191–1204.

Peremyslov, V. V., Morgun, E. A., Kurth, E. G., Makarova, K. S., Koonin, E. V., and Dolja, V. V. (2013). Identification of myosin XI receptors in Arabidopsis defines a distinct class of transport vesicles. *Plant Cell*, 25(8):3022–3038.

Peremyslov, V. V., Prokhnevsky, A. I., Avisar, D., and Dolja, V. V. (2008). Two class XI myosins function in organelle trafficking and root hair development in Arabidopsis. *Plant Physiol*, 146(3):1109–1116.

Peremyslov, V. V., Prokhnevsky, A. I., and Dolja, V. V. (2010). Class XI myosins are required for development, cell expansion, and F-actin organization in Arabidopsis. *Plant Cell*, 22(6):1883–1897.

Pinski, G. and Narin, F. (1976). Citation influence for journal aggregates of scientific publications: theory, with application to the literature of physics. *Inf Proc Manage*, 12(5):297–312.

Pound, M. P., French, A. P., Atkinson, J., Wells, D. M., Bennett, M. J., and Pridmore, T. P. (2013). RootNav: navigating images of complex root architectures. *Plant Physiol*, 162(4):1802–1814.

Prokhnevsky, A. I., Peremyslov, V. V., and Dolja, V. V. (2008). Overlapping functions of the four class XI myosins in Arabidopsis growth, root hair elongation, and organelle motility. *Proc Natl Acad Sci*, 105(50):19744–19749.

Qiu, J. and Li, F.-F. (2014). Quantitative morphological analysis of curvilinear network for microscopic image based on individual fibre segmentation. *J Microsc*, 256(3):153–165.

Racine, V., Sachse, M., Salamero, J., Fraisier, V., Trubuil, A., and Sibarita, J.-B. (2007). Visualization and quantification of vesicle trafficking on a three-dimensional cytoskeleton network in living cells. *J Microsc*, 225(3):214–228.

Rao, A. S. and C., P. R. (1990). Linear algorithm for optimal path cover problem on interval graphs. *Inform Process Lett*, 35(3):149–153.

Recho, P., Joanny, J.-F., and Truskinovsky, L. (2014). Optimality of contraction-driven crawling. *Phys Rev Lett*, 112(21):218101.

Reggiani, A. and Nijkamp, P. (2009). *Complexity and spatial networks: in search of simplicity*. Springer, Berlin, 1st edition.

Riedl, J., Crevenna, A. H., Kessenbrock, K., Yu, J. H., Neukirchen, D., Bista, M., Bradke, F., Jenne, D., Holak, T. A., Werb, Z., et al. (2008). Lifeact: a versatile marker to visualize F-actin. *Nat Methods*, 5(7):605–607.

Rigort, A., Günther, D., Hegerl, R., Baum, D., Weber, B., Prohaska, S., Medalia, O., Baumeister, W., and Hege, H.-C. (2012). Automated segmentation of electron tomograms for a quantitative description of actin filament networks. *J Struct Biol*, 177(1):135–144.

Rogers, S. L. and Gelfand, V. I. (2000). Membrane trafficking, organelle transport, and the cytoskeleton. *Curr Opin Cell Biol*, 12(1):57–62.

Ross, J. L., Ali, M. Y., and Warshaw, D. M. (2008). Cargo transport: molecular motors navigate a complex cytoskeleton. *Curr Opin Cell Biol*, 20(1):41–47.

Sack, L. and Scoffoni, C. (2013). Leaf venation: structure, function, development, evolution, ecology and applications in the past, present and future. *New Phytologist*, 198(4):983–1000.

Samaj, J., Baluška, F., Voigt, B., Schlicht, M., Volkmann, D., and Menzel, D. (2004). Endocytosis, actin cytoskeleton, and signaling. *Plant Physiol*, 135(3):1150–1161.

Sambade, A., Pratap, A., Buschmann, H., Morris, R. J., and Lloyd, C. (2012). The influence of light on microtubule dynamics and alignment in the Arabidopsis hypocotyl. *Plant Cell Online*, 24(1):192–201.

Sampathkumar, A., Gutierrez, R., McFarlane, H. E., Bringmann, M., Lindeboom, J., Emons, A.-M., Samuels, L., Ketelaar, T., Ehrhardt, D. W., and Persson, S. (2013). Patterning and lifetime of plasma membrane-localized cellulose synthase is dependent on actin organization in Arabidopsis interphase cells. *Plant Physiol*, 162(2):675–688.

Sampathkumar, A., Lindeboom, J. J., Debolt, S., Gutierrez, R., Ehrhardt, D. W., Ketelaar, T., and Persson, S. (2011). Live cell imaging reveals structural associations between the actin and microtubule cytoskeleton in Arabidopsis. *Plant Cell*, 23(6):2302–2313.

Saporta, G. and Youness, G. (2002). Comparing two partitions: some proposals and experiments. In *Compstat*, pages 243–248. Springer, Berlin.

Sato, Y., Nakajima, S., Shiraga, N., Atsumi, H., Yoshida, S., Koller, T., Gerig, G., and Kikinis, R. (1998). Three-dimensional multi-scale line filter for segmentation and visualization of curvilinear structures in medical images. *Med Image Anal*, 2(2):143–168.

Schich, M., Song, C., Ahn, Y.-Y., Mirsky, A., Martino, M., Barabási, A.-L., and Helbing, D. (2014). Quantitative social science: a network framework of cultural history. *Science*, 345(6196):558–562.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nat Methods*, 9(7):676–682.

Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012). Nih image to imagej: 25 years of image analysis. *Nat Methods*, 9(7):671–675.

Schrijver, A. (1998). *Theory of Linear and Integer Programming*. Wiley, New York, 1st edition.

Shafrir, Y., ben Avraham, D., and Forgacs, G. (2000). Trafficking and signaling through the cytoskeleton: a specific mechanism. *Sci Signal*, 113(15):2747.

Shafrir, Y. and Forgacs, G. (2002). Mechanotransduction through the cytoskeleton. *J Cell Physiol*, 282(3):C479–C486.

Sheahan, M. B., Staiger, C. J., Rose, R. J., and McCurdy, D. W. (2004). A green fluorescent protein fusion to actin-binding domain 2 of Arabidopsis fimbrin highlights new features of a dynamic actin cytoskeleton in live plant cells. *Plant Physiol*, 136(4):3968–3978.

Shen-Orr, S., Milo, R., Mangan, S., and Alon, U. (2002). Network motifs in the transcriptional regulation network of Escherichia coli. *Nat Genet*, 31(1):64–68.

Shih, Y.-L. and Rothfield, L. (2006). The bacterial cytoskeleton. *Microbiol Mol Biol Rev*, 70(3):729–754.

Shimmen, T. (2007). The sliding theory of cytoplasmic streaming: fifty years of progress. *J Plant Res*, 120(1):31–43.

Shimmen, T., Ridge, R., Lambiris, I., Plazinski, J., Yokota, E., and Williamson, R. (2000). Plant myosins. *Protoplasma*, 214(1-2):1–10.

Shimmen, T. and Yokota, E. (2004). Cytoplasmic streaming in plants. *Curr Opin Cell Biol*, 16(1):68–72.

Shoval, O., Sheftel, H., Shinar, G., Hart, Y., Ramote, O., Mayo, A., Dekel, E., Kavanagh, K., and Alon, U. (2012). Evolutionary trade-offs, Pareto optimality, and the geometry of phenotype space. *Science*, 336(6085):1157–1160.

Smith, L. G. and Oppenheimer, D. G. (2005). Spatial control of cell expansion by the plant cytoskeleton. *Annu Rev Cell Dev Biol*, 21:271–295.

Smith, M. B., Li, H., Shen, T., Huang, X., Yusuf, E., and Vavylonis, D. (2010). Segmentation and tracking of cytoskeletal filaments using open active contours. *Cytoskeleton*, 67(11):693–705.

Snowman, B. N., Kovar, D. R., Shevchenko, G., Franklin-Tong, V. E., and Staiger, C. J. (2002). Signal-mediated depolymerization of actin in pollen during the self-incompatibility response. *Plant Cell*, 14(10):2613–2626.

Somerville, C., Bauer, S., Brininstool, G., Facette, M., Hamann, T., Milne, J., Osborne, E., Paredez, A., Persson, S., Raab, T., Vorwerk, S., and Youngs, H. (2004). Toward a systems approach to understanding plant cell walls. *Science*, 306(5705):2206–2211.

Sousbie, T., Pichon, C., Colombi, S., Novikov, D., and Pogosyan, D. (2008a). The 3D skeleton: tracing the filamentary structure of the universe. *Mon Not R Astron Soc*, 383(4):1655–1670.

Sousbie, T., Pichon, C., Courtois, H., Colombi, S., and Novikov, D. (2008b). The three-dimensional skeleton of the SDSS. *Astrophys J Lett*, 672(1):L1.

Sporns, O. and Kötter, R. (2004). Motifs in brain networks. *PLoS Biol*, 2:e369.

Sporns, O., Tononi, G., and Kötter, R. (2005). The human connectome: a structural description of the human brain. *PLoS Comput Biol*, 1(4):e42.

Staehelin, L. A. and Hepler, P. K. (1996). Cytokinesis in higher plants. *Cell*, 84(6):821–824.

Staiger, C. J. (2000). Signaling to the actin cytoskeleton in plants. *Annu Rev Plant Biol*, 51(1):257–288.

Staiger, C. J., Baluška, F., Volkmann, D., and Barlow, P. (2000). *Actin: a dynamic framework for multiple plant cell functions*. Springer, Amsterdam, 1st edition.

Staiger, C. J. and Blanchoin, L. (2006). Actin dynamics: old friends with new stories. *Curr Opin Plant Biol*, 9(6):554–562.

Staiger, C. J. and Lloyd, C. W. (1991). The plant cytoskeleton. *Curr Opin Cell Biol*, 3(1):33–42.

Staiger, C. J., Sheahan, M. B., Khurana, P., Wang, X., McCurdy, D. W., and Blanchoin, L. (2009). Actin filament dynamics are dominated by rapid growth and severing activity in the Arabidopsis cortical array. *J Cell Biol*, 184(2):269–280.

Stamm, A. J. (1964). *Wood and cellulose science*. Ronald Press Co, New York, 1st edition.

Steinmetz, L. M., Sinha, H., Richards, D. R., Spiegelman, J. I., Oefner, P. J., McCusker, J. H., and Davis, R. W. (2002). Dissecting the architecture of a quantitative trait locus in yeast. *Nature*, 416(6878):326–330.

Stephens, D. J. and Allan, V. J. (2003). Light microscopy techniques for live cell imaging. *Science*, 300(5616):82–86.

Sternberg, S. R. (1983). Biomedical image processing. *Computer*, 16(1):22–34.

Stoica, R. S., Martinez, V. J., Mateu, J., and Saar, E. (2005). Detection of cosmic filaments using the Candy model. *Astron Astrophys*, 434(2):423–432.

Sun, J., Boyd, S., Xiao, L., and Diaconis, P. (2006). The fastest mixing Markov process on a graph and a connection to a maximum variance unfolding problem. *SIAM Rev*, 48(4):681–699.

Supowit, K. J. (1983). The relative neighborhood graph, with an application to minimum spanning trees. *J ACM*, 30(3):428–448.

Swedlow, J. R. and Eliceiri, K. W. (2009). Open source bioimage informatics for cell biology. *Trends Cell Biol*, 19(11):656–660.

Swedlow, J. R., Goldberg, I. G., Eliceiri, K. W., and Consortium, O. M. E. (2009). Bioimage informatics for experimental biology. *Annu Rev Biophys*, 38:327–346.

Szymanski, D. B. (2009). Plant cells taking shape: new insights into cytoplasmic control. *Curr Opin Plant Biol*, 12(6):735–744.

Tarsi, M. (1981). On the decomposition of a graph into stars. *Discrete Math*, 36(3):299–304.

Tero, A., Takagi, S., Saigusa, T., Ito, K., Bebber, D. P., Fricker, M. D., Yumiki, K., Kobayashi, R., and Nakagaki, T. (2010). Rules for biologically inspired adaptive network design. *Science*, 327(5964):439–442.

Thévenaz, P., Ruttimann, U. E., and Unser, M. (1998). A pyramid approach to subpixel registration based on intensity. *IEEE Trans Image Process*, 7(1):27–41.

Thomas, C., Tholl, S., Moes, D., Dieterle, M., Papuga, J., Moreau, F., and Steinmetz, A. (2009). Actin bundling in plants. *Cell Motil Cytoskeleton*, 66(11):940–957.

Tindemans, S. H., Hawkins, R. J., and Mulder, B. M. (2010). Survival of the aligned: ordering of the plant cortical microtubule array. *Phys Rev Lett*, 104(5):058103.

Tindemans, S. H. and Mulder, B. M. (2010). Microtubule length distributions in the presence of protein-induced severing. *Phys Rev E*, 81(3 Pt 1):031910.

Tomalia, D. A. and Frechet, J. M. (2001). *Dendrimers and other dendritic polymers*. Wiley, New York, 1st edition.

Toussaint, G. T. (1980). The relative neighbourhood graph of a finite planar set. *Lect Notes Comput Sci*, 12(4):261–268.

Tully, R. B., Courtois, H., Hoffman, Y., and Pomarède, D. (2014). The Laniakea supercluster of galaxies. *Nature*, 513(7516):71–73.

Vale, R. D. (2003). The molecular motor toolbox for intracellular transport. *Cell*, 112(4):467–480.

van den Ent, F., Amos, L. A., and Löwe, J. (2001). Prokaryotic origin of the actin cytoskeleton. *Nature*, 413(6851):39–44.

van der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, J. D., Yager, N., Gouillart, E., and Yu, T. (2014). scikit-image: image processing in Python. *PeerJ*, 2:e453.

van Mameren, J., Vermeulen, K. C., Gittes, F., and Schmidt, C. F. (2009). Leveraging single protein polymers to measure flexural rigidity. *J Phys Chem B*, 113(12):3837–3844.

Van Rossum, G. and Drake, F. L. (2011). *Python Language Reference Manual*. Network Theory Ltd, Godalming, 1st edition.

Vazirani, V. V. (2001). *Approximation algorithms*. Springer, Berlin, 1st edition.

Vaziri, A. and Gopinath, A. (2008). Cell and biomolecular mechanics in silico. *Nat Mat*, 7(1):15–23.

Versteeg, H. K. and Malalasekera, W. (2007). *An introduction to computational fluid dynamics*. Pearson Edu, Essex, 2nd edition.

Verwer, R. W. and van Pelt, J. (1983). A new method for the topological analysis of neuronal tree structures. *J Neurosci Meth*, 8(4):335–351.

Viswanathan, G. M., Buldyrev, S. V., Havlin, S., da Luz, M. G., Raposo, E. P., and Stanley, H. E. (1999). Optimizing the success of random searches. *Nature*, 401(6756):911–914.

Volkmann, D. and Baluska, F. (1999). Actin cytoskeleton in plants: from transport networks to signaling networks. *Microsc Res Tech*, 47(2):135–154.

Wagner, B., Tharmann, R., Haase, I., Fischer, M., and Bausch, A. (2006). Cytoskeletal polymer networks: the molecular structure of cross-linkers determines macroscopic properties. *Proc Natl Acad Sci*, 103(38):13974–13978.

Wakatsuki, T., Schwab, B., Thompson, N. C., and Elson, E. L. (2001). Effects of cytochalasin D and latrunculin B on mechanical properties of cells. *J Cell Sci*, 114(5):1025–1036.

Waller, F. and Nick, P. (1997). Response of actin microfilaments during phytochrome-controlled growth of maize seedlings. *Protoplasma*, 200(3-4):154–162.

Wang, P. and Hussey, P. J. (2015). Interactions between plant endomembrane systems and the actin cytoskeleton. *Front Plant Sci*, 6:422.

Wasteneys, G. O. (2000). The cytoskeleton and growth polarity. *Curr Opin Plant Biol*, 3(6):503–511.

Wasteneys, G. O. and Ambrose, J. C. (2009). Spatial organization of plant cortical microtubules: close encounters of the 2D kind. *Trends Cell Biol*, 19(2):62–71.

West, D. B. (2001). *Introduction to graph theory*. Prentice Hall, Upper Saddle River, 2nd edition.

West, G. B., Brown, J. H., and Enquist, B. J. (1997). A general model for the origin of allometric scaling laws in biology. *Science*, 276(5309):122–126.

West, G. B., Brown, J. H., and Enquist, B. J. (1999). A general model for the structure and allometry of plant vascular systems. *Nature*, 400(6745):664–667.

Wickstead, B. and Gull, K. (2011). The evolution of the cytoskeleton. *J Cell Biol*, 194(4):513–525.

Wightman, R. and Turner, S. (2010). Trafficking of the plant cellulose synthase complex. *Plant Physiol*, 153(2):427–432.

Wightman, R. and Turner, S. R. (2008). The roles of the cytoskeleton during cellulose deposition at the secondary cell wall. *Plant J*, 54(5):794–805.

Williams, R. J. and Martinez, N. D. (2000). Simple rules yield complex food webs. *Nature*, 404(6774):180–183.

Wintrebert, P. (1931). La rotation immédiate de l'oeuf pondu et la e rotation d'activation chez Discoglossus pictus otth. *C R Soc Biol*, 106:439–442.

Wolkowicz, H., Saigal, R., and Vandenberghe, L. (2012). *Handbook of semidefinite programming: theory, algorithms, and applications*. Springer Science, New York, 1st edition.

Wolsey, L. A. and Nemhauser, G. L. (1999). *Integer and Combinatorial Optimization*. Wiley-Interscience, New York, 1st edition.

Wood, S. T., Dean, B. C., and Dean, D. (2013). A linear programming approach to reconstructing subcellular structures from confocal images for automated generation of representative 3D cellular models. *Med Image Anal*, 17(3):337–347.

Woodhouse, F. G. and Goldstein, R. E. (2013). Cytoplasmic streaming in plant cells emerges naturally by microfilament self-organization. *Proc Natl Acad Sci*, 110(35):14132–14137.

Wu, T.-H. (1997). A note on a global approach for general 0–1 fractional programming. *Eur J Oper Res*, 101(1):220–223.

Wymer, C. and Lloyd, C. (1996). Dynamic microtubules: implications for cell wall patterns. *Trends Plant Sci*, 1(7):222–228.

Xu, T., Li, H., Shen, T., Ojkic, N., Vavylonis, D., and Huang, X. (2011). Extraction and analysis of actin networks based on open active contour models. *Proc IEEE Int Symp Biomed Imaging*, 2011:1334–1340.

Xu, T., Vavylonis, D., and Huang, X. (2014). 3D actin network centerline extraction with multiple active contours. *Med Image Anal*, 18(2):272–284.

Xu, T., Vavylonis, D., Tsai, F.-C., Koenderink, G. H., Nie, W., Yusuf, E., I-Ju Lee, Wu, J.-Q., and Huang, X. (2015). SOAX: a software for quantification of 3D biopolymer networks. *Sci Rep*, 5:9081.

Yarmola, E. G., Somasundaram, T., Boring, T. A., Spector, I., and Bubb, M. R. (2000). Actin-latrunculin A structure and function differential modulation of actin-binding protein function by latrunculin A. *J Biol Chem*, 275(36):28120–28127.

Yue, D., Guillén-Gosálbez, G., and You, F. (2013). Global optimization of large-scale mixed-integer linear fractional programming problems: a reformulation-linearization method and process scheduling applications. *AIChE J*, 59(11):4255–4272.

Zhou, D., Li, J., and Zha, H. (2005). A new mallows distance based metric for comparing clusterings. In *Proc 22nd Internat Conf Mach Learn*, pages 1028–1035, New York. ACM.

Zhu, C., Byrd, R. H., Lu, P., and Nocedal, J. (1997). Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization. *ACM Trans Math Software*, 23(4):550–560.

Zhu, J., Ingram, P. A., Benfey, P. N., and Elich, T. (2011). From lab to field, new approaches to phenotyping root system architecture. *Curr Opin Plant Biol*, 14(3):310–317.

Ziemann, F., Rädler, J., and Sackmann, E. (1994). Local measurements of viscoelastic moduli of entangled actin networks using an oscillating magnetic bead micro-rheometer. *Biophys J*, 66(6):2210–2216.

Zumdieck, A., Cosentino Lagomarsino, M., Tanase, C., Kruse, K., Mulder, B., Dogterom, M., and Jülicher, F. (2005). Continuum description of the cytoskeleton: ring formation in the cell cortex. *Phys Rev Lett*, 95(25):258103.

8 Acknowledgements

Since actions speak louder than words, let me simply list here the names of people whose help and support I would like to acknowledge.

I would like to thank my brilliant supervisor Zoran Nikoloski.

I would like to thank the former and current members of the “Systems Biology and Mathematical Modeling” group, Semidán Robaina Estévez, Jeanne Marie Onana-Eloundou Mbebi, Anika Küken, Kevin Schwahn, Max Sajitz-Hermstein, Nooshin Omranian, Sabrina Kleessen, Nadine Töpfer, Sergio Grimbs, Anne Arnold, Jost Neigenfind, Jacqueline Nowak, Juraj Labant, Robert Heise, Hao Tong, Michael Scheunemann, Marina Leer, Andreas Krug, Alberto Castellini, and Alessio Milanese.

I would like to thank my collaborators, Staffan Persson, Alexander Ivakov, René Schneider, Praveen Krishnamoorthy, Clara Sánchez Rodríguez, Federico Apelt, Fritz Kragler, Mark Stitt, Kamil Sklodowski, and Marc Timme.

I would like to thank the International Max Planck Research School “Primary Metabolism and Plant Growth” for funding, Ina Talke, and my fellow scholarship holders.

I would like to thank my PhD advisory committee Joachim Selbig, Lothar Willmitzer, Staffan Persson, and Marc Timme.

I would like to thank for the technical support by Jost Neigenfind and Andreas Donath.

I would like to thank my fellow image analysis enthusiasts Christian Kappel, Alexander Leow, Federico Apelt, and Alexander Ivakov.

I would like to thank my dear friends Alex, Matthias, Lars, and Lisa and my unnamed friends from school, Parkour, Lesekreis, Soliküche, and other contexts.

I would like to thank my familiy, Bärbel, Sonja, Anneliese aka Omi, and most of all Anna.

Finally, I would like to acknowledge the work of countless anonymous and relentless idealists who push forward open-source software such as Linux, Python, Lyx, and Inkscape, for the benefit of everyone.

9 Author contributions

This is to certify that David Breuer significantly contributed as a first author to each of the publications presented in this thesis. In particular, his contributions encompassed design of the studies, execution of microscopic experiments, development and implementation of the computational methods, data analysis, interpretation of the results, and writing of the presented publications.

Potsdam, 04.01.2016

Prof. Dr. Joachim Selbig

10 Statement of authorship

I hereby declare that the present work has not been previously submitted for another degree at any university, has been carried out by myself, and employed only the cited references and resources.

Potsdam, 04.01.2016

David Breuer