Electronic Supplementary Material:
Quantitative analyses of the plant cytoskeleton
reveal underlying organizational principles

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

[1]Systems Biology and Mathematical Modeling, Max Planck Institute of Molecular Plant Physiology,
Am Muehlenberg 1, 14476 Potsdam, Germany
[2]Plant Cell Walls, Max Planck Institute of Molecular Plant Physiology,
Am Muehlenberg 1, 14476 Potsdam, Germany
[3]Sainsbury Laboratory, Cambridge University, Bateman Street, Cambridge CB2 1LR, United Kingdom
[4]ARC Centre of Excellence in Plant Cell Walls, School of Botany, University of Melbourne,
Parkville 3010, Victoria, Australia

[*]Electronic address: nikoloski@mpimp-golm.mpg.de

ESM S1: 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 [55]. 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 ref. [55]. 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.

ESM S2: Network reconstruction procedure, different grid topologies and null models

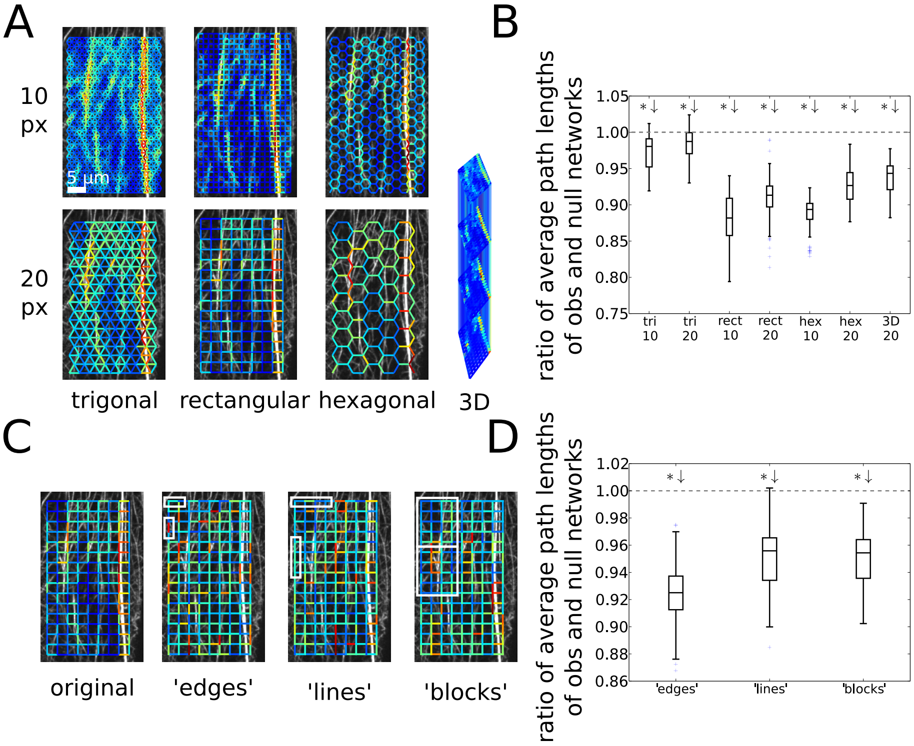


Figure S2: **Extensions of the network reconstruction framework and different null models.** **(A)** Reconstructions of the actin cytoskeleton of the same cell using different sizes (grid spacings of 10 and 20 pixels) of triangular, rectangular, and hexagonal grids, and a three-dimensional grid with 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.

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 [56]: First, the drift of the seedlings under the microscope was corrected using a stack registration algorithm [3]. 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 [61], 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 filaments and, hence, was filtered out in the present study. Finally, photobleaching was compensated by rescaling all images' mean intensities to one.

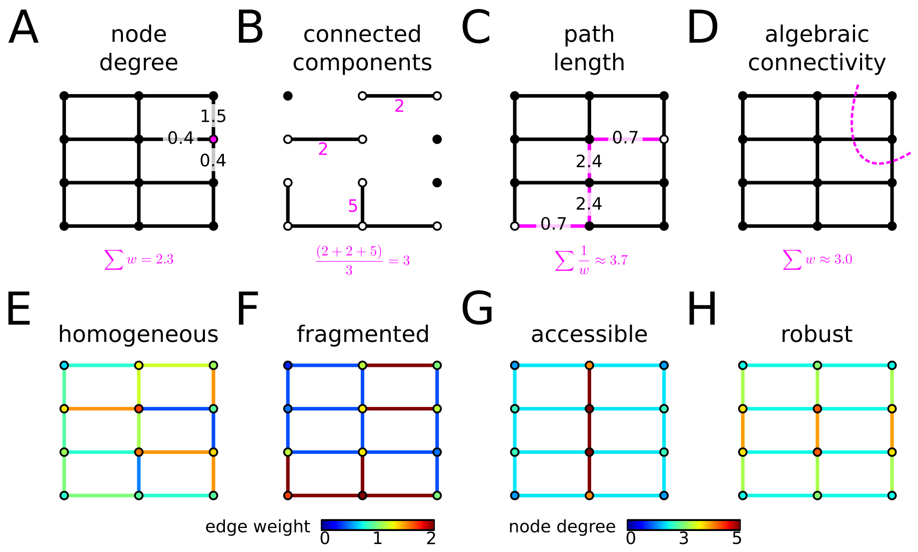
From these preprocessed images we obtained the cytoskeletal components as complex networks through a two-step procedure as described in the main text (*cf.* section "Reconstruction of complex networks from cytoskeletal images"). These steps and all statistical analyses were performed using Python [65]. 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. 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 section "Accessibility and robustness of cytoskeletal networks"), 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. S2A). 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. S2B; 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. S2C, "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 section "Accessibility and robustness of cytoskeletal networks". For example, for the APL, the ratio of observed and null model values falls significantly below one (Fig. S2D, "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. S2C, "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. S2D, "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. S2C, "blocks"; here, the network is composed into three times four blocks). This "block" null model also exhibits longer APLs than its biological counterpart (Fig. S2D, "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 section "Accessibility and robustness of cytoskeletal networks").

ESM S3: Network properties used for quantifying the cytoskeletal organization



SDD:	0.74	1.42	1.71	0.82
ASC:	8.00	3.00	4.00	4.0
OA:	45.00	50.70	37.96	38.88
APL:	2.36	3.65	2.26	2.42
AC:	0.49	0.27	0.66	0.74
AS:	0.06	0.23	0.10	0.21

Figure S3: **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) and, accordingly, displays a high AC.

In the manuscript, 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.*,

$$d_n = \sum_{\substack{e \in \mathcal{E} \\ n \in e}} w_e \tag{1}$$

(*cf.* Fig. S3A for the degree of a node of the network in panel E; color-coded node degrees in Figs. S3E-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,

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

In particular, the SDD does not measure the heterogeneity in filament thicknesses or numbers but a combination thereof. Comparing Figs. S3E 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. S3B 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 per non-trivial connected component (ASC) is a measure for the fragmentation of the network.

$$\mathrm{ASC} \;=\; C^{-1}\sum_{c=1}^{C} N_c, \tag{3}$$

where $\mathcal{C}$ is the set of $C = |\mathcal{C}|$ non-trivial components and N_c is the number of nodes in component $c \in \mathcal{C}$. Comparing Figs. S3E 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. (12) and its derivation is explained in ESM4. Two networks with different OA are shown in Figs. S3F and G with overall horizontal ($\mathrm{OA} > 45°$) and vertical ($\mathrm{OA} < 45°$) orientations, respectively, as confirmed visually.

Average path length: A path $\mathcal{P}$ 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. S3C 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.

$$\mathrm{APL} \;=\; 2^{-1}N^{-1}(N-1)^{-1}\sum_{n=1}^{N}\sum_{\substack{m=1\\ m>n}}^{N}\min_{\mathcal{P}\in\mathcal{P}_{n,m}}\sum_{e\in\mathcal{P}} w_e^{-1}, \tag{4}$$

where $\mathcal{P}_{n,m}$ 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. 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. S3G 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

$$\mathrm{AC} \;\equiv\; \lambda_2 \tag{5}$$

of the graph Laplacian L,

$$L_{n,m} \;=\; \begin{cases} d_n & \text{if } n = m \\ -w_{(n,m)} & \text{if } (m,n) \in \mathcal{E} \\ 0 & \text{otherwise,} \end{cases} \tag{6}$$

with $n, m \in \mathcal{N}$. 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.* [48]). 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. S3D). We solved a semi-definite optimization problem described in the discussion of Fig. 4 to construct a network with a fixed sum of edge weights that maximizes the AC. This network is shown in Fig. S3H 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. S3G 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

$$\mathrm{AS} \; = \; \frac{1}{2E} \sum_{n=1}^{N} \sum_{m=1}^{N} \left(w_{(n,m)} - \frac{d_n d_m}{2E} \right) d_n d_m. \tag{7}$$

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. S3E 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.

ESM S4: Method for determining angles and filament orientations from network structure

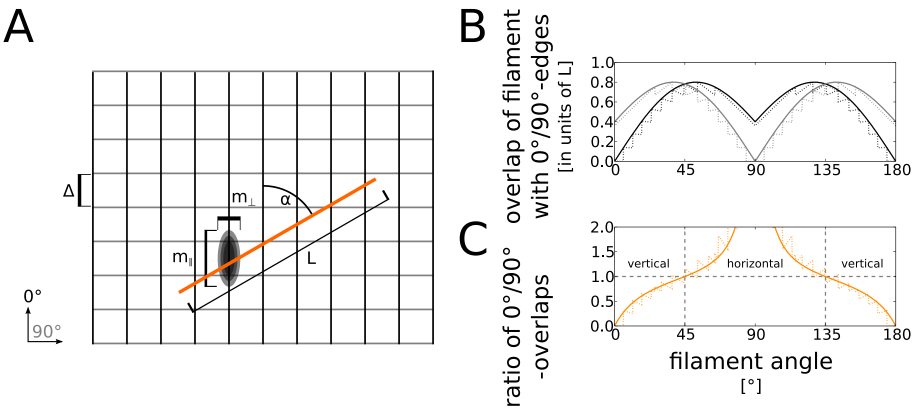


Figure S4: **Inferring orientation of filament from network structure.** **(A)** Schematic filament of length L at an angle α, placed on a rectangular grid with spacing Δ, vertical (0°, black), and horizontal edges (90°, gray). The shape of the edges' convolution kernels determines the contribution of the filament to each edge, with the two extremes $m_{\|}$ 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.

We now present a method to evaluate the orientation of AFs and MTs by exploiting their network structures. Starting from a given grid (Fig. S4A), we placed a stiff rod of length L that is rotated by an angle α and 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,

$$n_\gamma \;=\; \Delta^{-1} L \left|\sin\left(\alpha-\gamma\right)\right| . \tag{8}$$

The overlap m of the rod and a γ-edge was computed via the convolution kernel of that edge (see section "Network reconstruction procedure, different grid topologies and null models") and was approximated as

$$m_\gamma \;=\; m_\perp + \left(m_\| - m_\perp\right) \left|\cos\left(\alpha-\gamma\right)\right| , \tag{9}$$

where $m_\perp$ and $m_\|$ 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. (8) and (9))

$$w_\gamma \;=\; n_\gamma m_\gamma = \Delta^{-1} L \left|\sin\left(\alpha-\gamma\right)\right| \left[m_\perp + \left(m_\| - m_\perp\right)\left|\cos\left(\alpha-\gamma\right)\right|\right] . \tag{10}$$

Furthermore, we may include that only integer numbers of edges may be crossed by a filament. Then, $L \sin$ in Eq. (10) is replaced by $\lfloor L\sin\rfloor$ and the contribution of a filament of finite length becomes discontinuous and approaches the continuous curve for long filaments (Fig. S4B). For simplicity, we considered small grid spacings or long filaments and work with Eq. (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 $m_\|$ 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,

$$\frac{w_\gamma}{w_{\gamma'}} \;=\; \frac{\left|\sin\left(\alpha-\gamma\right)\right| \left[m_\perp + \left(m_\| - m_\perp\right)\left|\cos\left(\alpha-\gamma\right)\right|\right]}{\left|\sin\left(\alpha-\gamma'\right)\right| \left[m_\perp + \left(m_\| - m_\perp\right)\left|\cos\left(\alpha-\gamma'\right)\right|\right]} =: r, \tag{11}$$

where α may be interpreted as the overall orientation of the cytoskeletal filaments (Fig. S4C). For $\gamma = 0°$ and $\gamma^{'} = 90°$, Eq. (11) yields

$$\alpha \;=\; \arctan\left(\frac{m_\perp + r\left(m_\parallel - m_\perp\right)}{m_\perp r + \left(m_\parallel - m_\perp\right)}\right). \tag{12}$$

In our analysis, we refer to $\alpha \in [0°, 45°)$ and $\alpha \in (45°, 90°]$ as an overall vertical and horizontal orientation, respectively. See section "The reconstructed networks capture biologically relevant features of the actin and microtubule cytoskeletal components" for results on the orientation of the cytoskeletal components under different conditions.

ESM S5: Reconstruction and analysis of the German autobahn network

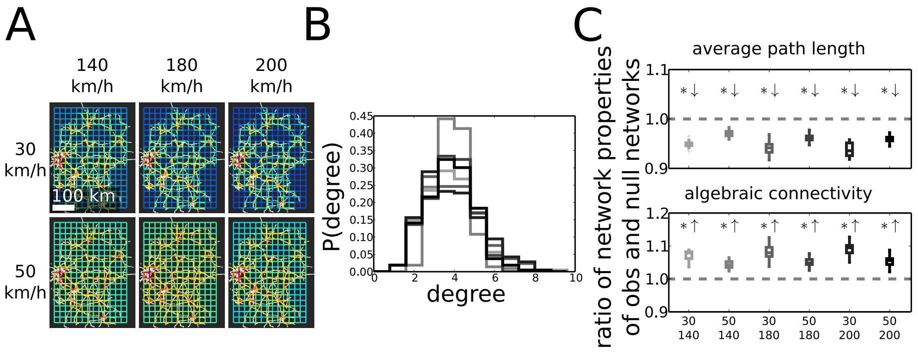


Figure S5: **The efficiency of the autobahn network is pertained for different choices of speed limits. (A)** Reconstructed autobahn networks with different maximum high speed limits (140, 180, and 200 km/h) and different off-highway speed limits (30 and 50 km/h). **(B)** The degree distributions of 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).

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. S5A).

The results for speed limits of 200 km/h and 50 km/h (Fig. 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. S5B 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.