基于最优传输的时间超网络的收敛性质

Diego Baptista[1] 和 Caterina De Bacco[1]

[1]*Max Planck Institute for Intelligent Systems, Cyber Valley, Tuebingen, 72076, Germany*

我们提出一种方法，从作为最优传输问题的解而得到的二维函数序列中提取时序超图。我们从超图结构的角度，探究这些解所体现的最优性原理。离散性质呈现出的模式不同于其连续对应物所表征的模式。分析这些模式有助于深入理解所研究的传输原理。我们还根据标准图性质，将这些高阶结构与其对应网络进行比较。我们提供证据表明，某些传输方案可能受益于超网络表示。我们通过分析从真实系统图像中提取的超网络性质，在真实数据上展示了我们的方法。

arXiv:2301.03498v1 [cs.DM] 9 Jan 2023

引言

最优传输（OT）是一种用于比较概率分布的理论框架 [15, 20, 22, 27]。尽管这项任务通常被表述为一个优化问题，近期研究已将其纳入动态偏微分方程框架 [11–14, 25, 26]。在此背景下，传输问题的解通常被视为一族演变函数的收敛状态。

在某些情形下，这些演化族的稳态支撑在网络形结构中 [29–31]。近年来，这一事实引起了网络科学家和图论学家的关注，并推动了将最优传输问题的解转换为实际图结构的方法的发展 [5, 16]。这拓展了可用于理解和解决这些传输问题的工具集。近期研究表明，在原始数学框架和转换后的图结构中，都能揭示出共同模式 [3]。

将这些函数表示为二元关系集合，已被证明在各种应用中具有意义[4, 13]。然而，传统的二元表示在表征真实系统中观察到的质量或信息等数量的流时可能有所局限。在社会传染 [2, 8], 、随机游走 [7, 23] 或非线性共识 [18]等应用中，可以观察到许多相互作用发生在三人或更多人之间的系统实例。理解高阶结构上的结构与动力学之间的关系，是一个活跃的研究领域 [19, 24]。例如，控制动力学的关键要素与高阶表示中超边大小的异质性有关 [19]。因此，描述这些系统的最佳方式是使用超图；超图是网络的推广形式，能够编码任意数量个体之间的结构化关系。有鉴于此，一个自然的问题是：基于最优传输的结构在高阶表示方面表现如何?

为了弥补关于由最优传输解导出的结构的高阶性质这一知识空白，我们进一步阐述了 [3] 中观察到的结果。具体而言，我们提出一种将二维函数族转换为时序超网络的方法。我们通过将观察到的相互作用编码为超边，丰富了与这些函数相关的现有网络结构。我们研究经典超图性质，并将其与传输问题相关的预定义代价泛函进行比较。最后，我们扩展了这一方法和分析，以研究来自真实数据的系统。我们构建了多头绒泡菌 [28] 的超图表示，并分析其拓扑特征。

方法

动态 Monge-Kantorovich 方法

动力学 Monge-Kantorovich 方程组。我们首先回顾用于求解最优传输问题的机制所涉及的基本要素。与其他用于求解该问题的标准优化方法不同，我们采用一种将问题转化为一组动力学偏微分方程的方法。通过这种方式，不断更新初始条件，直至达到收敛状态。Facca 等人 [12–14], 提出的动力学方程组如下所示。我们假设最优传输问题定义在连续二维空间 $\Omega \in \mathbb{R}^2$上，并且起初不存在底层网络结构。这使我们能够探索整个空间，以设计出作为传输问题之解的最优网络拓扑。输入中需要指定的主要量是源分布和目标分布。我们将它们称为源和汇，其中一定质量（例如运输网络中的乘客、供水网络中的水）被注入后再被抽取。我们用“强迫”函数 $f(x) = f^+(x) - f^-(x) \in \mathbb{R}$表示这些量，它描述了流量生成源 $f^+(x)$和汇 $f^-(x)$。为确保质量守恒，规定 $\int_\Omega f(x)dx = 0$。我们假设流动由瞬态 Fick-Poiseuille 通量 $q = -\mu\nabla u$支配，其中 μ, u 和 q 分别称为电导率（或输运密度）、输运势和

通量。直观而言，质量通过源注入，依据电导率在空间中移动，随后经由汇被抽取。质量移动的方式决定了通量，而通量取决于空间不同位置所承受的压力；这种压力由势函数描述。

动力学 *Monge-Kantorovich* (DMK) 方程组如下：

$$-\nabla \cdot (\mu(t,x)\nabla u(t,x)) = f^+(x) - f^-(x)\,, \tag{1}$$

$$\frac{\partial \mu(t,x)}{\partial t} = [\mu(t,x)\nabla u(t,x)]^\beta - \mu(t,x)\,, \tag{2}$$

$$\mu(0,x) = \mu_0(x) > 0\,, \tag{3}$$

其中 $\nabla = \nabla_x$。方程 (1) 描述菲克-泊肃叶通量的空间平衡，并辅以无流动的诺伊曼边界条件。方程 (2) 规定了该系统的动力学，并由所谓的交通率 β控制。它决定运输方案，并塑造解的拓扑：对于 $\beta < 1$，我们得到拥堵运输，此时流量最小化；而 $\beta > 1$ 则产生分支运输，使流量汇集到更小的空间范围内。$\beta = 1$ 的情形则恢复类最短路径结构。最后，方程 (3) 构成系统的初始化，可视为解的初始猜测。

解 $(\mu^*,\ u^*)$ o f 方程 (1) 至 (3) 最小化运输成本函数 $\mathcal{L}(\mu,\ u)$ [12–14], ，其定义如下：

$$\mathcal{L}(\mu,u) := \mathcal{E}(\mu,u) + \mathcal{M}(\mu,u) \tag{4}$$

$$\mathcal{E}(\mu,u) := \frac{1}{2}\int_\Omega \mu|\nabla u|^2 dx,\ \ \mathcal{M}(\mu,u) := \frac{1}{2}\int_\Omega \frac{\mu^{\frac{(2-\beta)}{\beta}}}{2-\beta} dx \quad . \tag{5}$$

$\mathcal{L}$ 可视为 $\mathcal{M}$（运输过程中耗散的总能量，即网络运行成本）与 $\mathcal{E}$（建设网络基础设施的代价，即基础设施成本）的组合。众所周知，该泛函的凸性会随 β 而变化。非凸情形出现在分支方案中，从而形成分形结构[13, 21]。本文考虑的正是这种情形，也是唯一能够提取出有意义的网络结构、进而提取超图的情形 [5]。

超图序列

超图构造。我们如下定义超图（亦称超网络） [6]: 超图 是一个元组H =(V，E)，其中 $V = \{v_1, ..., v_n\}$ 是顶点集，$E = \{e_1, e_2, ..., e_m\}$ 是超边集，其中$e_i \subset V$，$\forall i = 1,\ ...,\ m$，且 $|e_i| > 1$。若 $|e_i| = 2,\ \forall i$，则 H就是一个图。我们将满足$|e_i| = 2$ 的超边 e_i称为三角形，而将满足 $|e_i| = 3$的超边称为边。我们将 H 的1-骨架称为 H的团扩张。该图 G = (V，E_G)由 H的顶点 V 以及成对边构成；这些成对边是通过枚举构成 E中每条超边的节点集所能形成的所有节点对而建立的。

令 μ 为方程 (1) 至 (3) 的解所对应的电导率。如前所述，μ 在收敛时决定质量应如何移动以实现最优运输。与 [3], 类似，我们将这个二维函数转换为另一种数据结构，即超图。具体做法如下：考虑 $G(\mu) = (V_G, E_G)$，即使用 [5]中提出的方法提取的网络。我们将 $H(\mu)$定义为元组 (V_H, E_H)，其中 $V_H = V_G$和 $E_H = E_G \cup T_G$，使得 $T_G = \{(u, v, w) : (u, v), (v, w), (w, u) \in E_G, \}$. 换言之，$H(\mu)$ 是图 $G(\mu)$ 及其所有三角形的集合。这一选择的依据是，[5] 中提出的图提取方法使用三角形对连续空间 Ω 进行离散化，这可能会对提取出的图或超图结构产生显著影响。因此，三角形是构造超图时自然的子结构。本文提出的方法适用于三角形以外的高阶结构。探究这些附加结构如何影响所得超图的性质，将留待未来研究。

图1展示了所研究的超图之一的示例。红色形状表示$H(\mu)$中的不同三角形。请注意，尽管这里为简单起见，我们考虑每条超边 e对应 $|e| \leq 3$ 的情况，但如图右侧各面板所示，这些元素的并集也能很好地表示高阶结构。

由于这种超图构造方法适用于任意二维输运密度，因此我们不仅可以从收敛的 μ中提取超图，也可以从收敛前的任意时间步提取超图。这样，我们便可以将最优输运序列表示为随时间演化的超图，即时序超网络。

超图序列。形式上，设 $\mu(x,t)$ 为通过求解 DMK 模型得到的、同时依赖于时间和空间的输运密度 （或电导率）函数。我们将其记为序列 $\{\mu_t\}_{t=0}^T$，其中索引为 T （通常取收敛状态对应的索引）。每个 μ_t 都是初始猜测 μ_0的第 t 次更新，按照式 (1) 至式 (3) 中所述规则计算得到。由此确定从 $\{\mu_t\}_{t=0}^T$ 中提取的超网络序列 $\{H(\mu_t)\}_{t=0}^T$

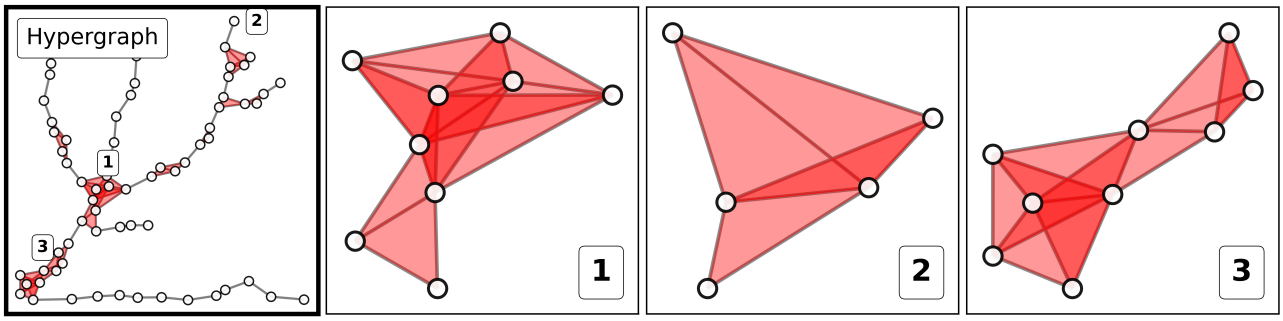

图 1：**超网络构造**。使用边和三角形作为超边来构建高阶结构。最左侧面板展示了所研究的图之一，以及所使用的三角形（红色）。后续面板突出显示了主超图中可见的不同三角形簇。

使用 [5]中提出的提取方法。图 2 展示了采用该方法，在不同时间步从所研究的某个序列$\{\mu_t\}$ 中构建的三个超图。相应的最优传输问题由（填充和空心）圆定义：质量从左下方的圆注入，并必须在标出的目的地提取。上排展示了解的不同更新（即 $t = 12$,18, 26）。它们定义在 $[0, 1]^2$. 的离散化上。最深的颜色表示其支撑集。下排展示了从这些函数中提取的超图。如图所示，只有边（灰色）和三角形（红色）被视为 $H(\mu_t)$ 的一部分。请注意， t 越大，超图越稀疏；对于均匀初始分布 μ_0 和分支最优传输 ($\beta > 1$) [13]，这是预期的结果。

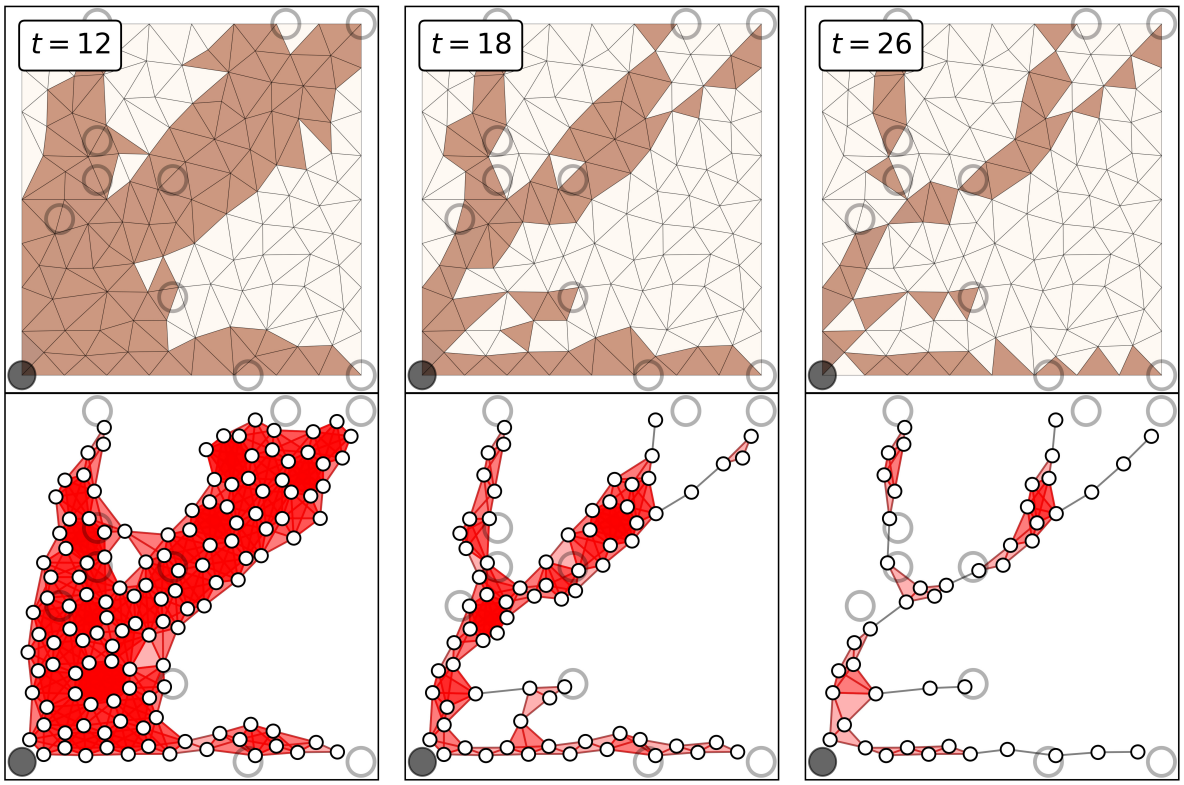

图 2：**时序超图**。顶行：序列 $\{\mu_t\}$ 的不同时间戳；三角形是 $[0, 1]^2$ 的离散化结果。底行：在顶行所示时间步提取的 μ_t 超图；三角形以红色突出显示。在两行中，实心圆和空心圆分别对应 f^+ 和 f^- 的支撑集，即源和汇。该序列由 $\beta = 1.5$ 获得。

图与超图性质

我们将超图序列与其对应网络（定义如上一段所述）进行比较。我们针对序列中的不同元素以及不同序列，分析以下主要网络和超图性质。分别用 $G=(V_G, E_G)$ 和 $H=(V_H, E_H)$ 表示属于某个序列 $\{G(\mu_t)\}_{t=0}^{T}$ and $\{H(\mu_t)\}_{t=0}^{T}$ 的一个待研究图和超图。我们考虑以下网络性质：

1. $|E_G|$，边的总数；
2. 平均度 $d(G)$，即每个节点的邻居数均值；
3. 平均接近中心性 $c(G)$：设 $v \in V_G$， v 的接近中心性定义为 $\sum_{u \in V_G} 1/d(u,v)$，其中 $d(u,v)$ 是 u 与 v 之间的最短路径距离。

借助广义邻接矩阵和线图 [1]，可以轻松根据前面的定义调整超网络性质。设 H 为一个顶点集为 $V=\{1,..,n\}$、边集为 $E=\{e_1,...,e_m\}$ 的超图。我们将 H 的广义节点 s-邻接矩阵 A_s 定义为一个大小为 $n \times n$ 的二进制矩阵，即当 i 和 j 至少属于 s 条共享超边时， $A_s[i][j]=1$；否则为 $A_s[i][j]=0$。我们将 s-线图 L_s 定义为由邻接矩阵 A_s 生成的图。注意， A_1 对应于 H 的骨架（即 L_1）的邻接矩阵。图 3 展示了一个邻接矩阵族，以及利用这些矩阵生成的线图。于是，我们可以按以下方式定义超图性质：

1. $|E_H|$，超边总数；
2. $|T|=|\{e \in E_H : \ |e|=3\}|$，三角形总数；
3. $S=\sum_{t \in T} a(t)$，覆盖面积，其中 $a(t)$是三角形 t的面积；
4. 平均度 $d_s(H)$，每个节点所关联的大小大于或等于 s 的超边的平均数量；
5. 平均接近中心性 $c_s(H)$：设 $v \in V_H$， v 的接近中心性定义为其在 L_s中的接近中心性。

S 可以根据超边的任何其他属性来定义，例如定义为其大小 $|e|$的函数。这里我们考虑超边覆盖的面积，以保持几何视角。另一方面，通过适当改变求和中的集合 T，该面积 S 也可以轻松推广到具有 $|e_i|>3$ 的超边，例如考虑包含四个节点的结构。至于中心性度量，我们重点比较 $s>1$ 与 $s=1$的情况，因为后者可追溯至标准图性质，而我们更关注于研究超图所固有的性质。图4显示了收敛超图 H （由 β的不同取值获得）中 $d_s(H)$和 $c_s(H)$的值，以及其对应图版本的度中心性和接近中心性。所考虑的超图显示在图的顶行中。如图所示，对于 β的不同取值，模式存在显著差异。由于 s 控制网络中不同节点之间的最小共享连接数，该数值越高，这一条件就越严格，从而导致更多不连通线图。在 s-度中心性的情况下，我们观察到其值随着 s的增加而降低，其中中心性最高的节点的值远高于中心性较低的节点。对于 $s=2$， 3，我们观察到的值高于 G中的节点。这是因为一旦向 G中添加超边，每个节点的关联数只会增加。中心性分布强烈依赖于 β。当取值较小时——即流量更加分散（$\beta=1.1$）——每个节点的超边数量仍多于与其连接的普通边数量。但如果流量集中在更小的空间内（$\beta=1.9$），则只会发现极少的超边。这表明，从超图中学习到的、不同于图骨架中所包含信息的信息，会受到所选流量机制的影响。

关于接近中心性分布，当 β较小时，它与 G 的分布相似，不受 s影响。当β 较大时，分布则趋向于近乎二元的信号。因此，随着 β 增大，节点往往会变得更加居中，这表明向网络中添加超边 G 会缩短节点之间的距离。在 s 取值最高时出现的信息损失，是因为线图 L_s 会变得不连通，并包含许多较小的连通分量。在这些情况下，若节点孤立，其接近中心性为 0；否则，其接近中心性与该节点所在的小型连通分量的直径成正比。

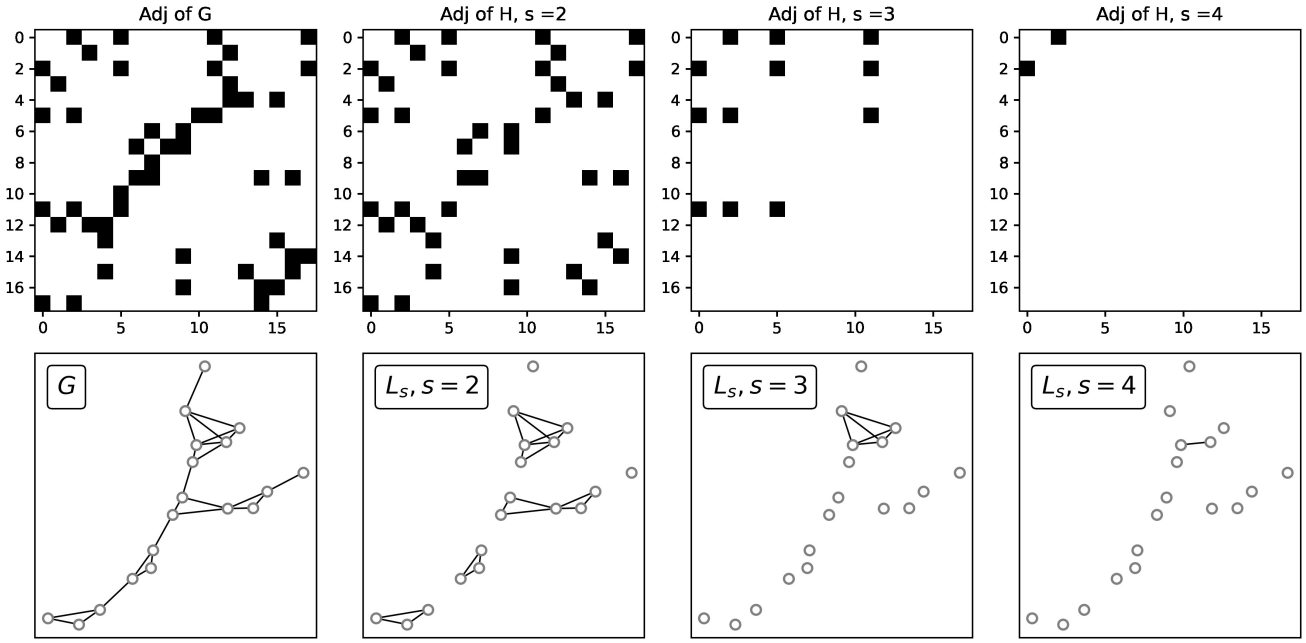

图 3：**邻接矩阵与线图**。上：不同取值下的广义节点 s-邻接矩阵 s 由给定的玩具图 G生成。下排从左到右：参考网络 G，以及适用于 $s = 2$、3和4的 s-线图。

a. 收敛准则。DMK 方程 (1) 至 (3) 的数值收敛通常通过设定一个阈值τ来定义。当与更新相关的代价相较于前一个时间步的代价变化不再超过$\leq \tau$时，即认为更新已足够。通常，当该阈值过小（在我们的实验中为$\tau = 10^{-12}$）时，代价或网络结构可能会在算法收敛之前就稳定在某个常数值。与 [3], 类似，为了有意义地确定何时达到超图最优性，我们将收敛时间定义为运输代价或某一给定网络性质首次达到一个值的时间步，该值小于或等于算法收敛时同一量所达到值的一定比例 p（在此处的实验中，我们使用 $p = 1.05$）。我们分别用 $t_{\mathcal{L}}$ 和 t_P 表示以代价函数或网络性质衡量的收敛时间。

结果

为了检验上一节提出的性质，并理解这些性质与运输最优性的联系，我们通过源和汇的配置合成一组最优传输问题。如 [3], 所做，我们固定一个源的位置，并在集合 $[0,1]^2$ 中采样若干点作为汇的位置。令 $S = \{s_0, s_1, ..., s_M\}$ 为空间 $[0,1]^2$, 中的位置集合，并固定一个正数 $0 < r$。我们将分布f^+和 f^- 定义为 $f^+(x) \propto \mathbb{1}_{R_0}(x)$，以及 $f^-(x) \propto \sum_{i>0} \mathbb{1}_{R_i}(x)$，其中 $\mathbb{1}_{R_i}(x) := 1$，当 $x \in R_i$时成立；否则，$\mathbb{1}_{R_i}(x) := 0$； $R_i = C(s_i,\ r)$表示圆心为 s_i、半径为 r 的圆。 r的取值依据所用的离散化方式确定；如前所述，圆心是均匀随机采样的。符号 $\propto$ 表示成比例关系，用于确保 f^+ 和 f^- 都是概率分布。运输成本取自式（4）。

b. 合成最优传输问题。我们实验中考虑的传输问题由 100 个源点-汇点配置组成。我们将源点 $s_0 = (0,\ 0)$（即 f^+ 在 (0，0) 处的支撑集）设为固定位置，并从规则网格中均匀随机抽取 15 个点 s_1，s_2，..... s_M。由于这些点取自网格的节点，我们确保任意两个不同位置之间都有足够的间距，因此在空间离散化后仍会被视为不同位置。我们将 $\mu_0(x) = 1$，$\forall x$ 初始化为 $[0,1]^2$ 上的均匀分布。这可以理解为对解的一种不带任何先验信息的初始猜测。从 μ_0, 出发，我们最多计算 300 次更新。根据所选的交通率 β，所需迭代次数可能有所不同。我们认为，序列 $\{\mu_t\}_{t=0}^{T}$ 收敛到某个函数 μ^*，其收敛迭代次数为 T，条件是满足 $|\mu_T - \mu_{T-1}| < \tau$，对于容差 $\tau \in (0,1],$，或者 T达到上述最大迭代次数。本文报告的实验中，容差 τ 设为 10^{-12}。鉴于解对交通约束的依赖，我们考虑了 β 的一系列取值。具体而言，我们研究从低流量情形（$\beta = 1.1$，因此交通惩罚较小）到高流量情形（$\beta = 1.9$）下得到的解，这些情形均生成分支运输方案。我们的 100 个问题共对应 900 个超图序列，每个序列

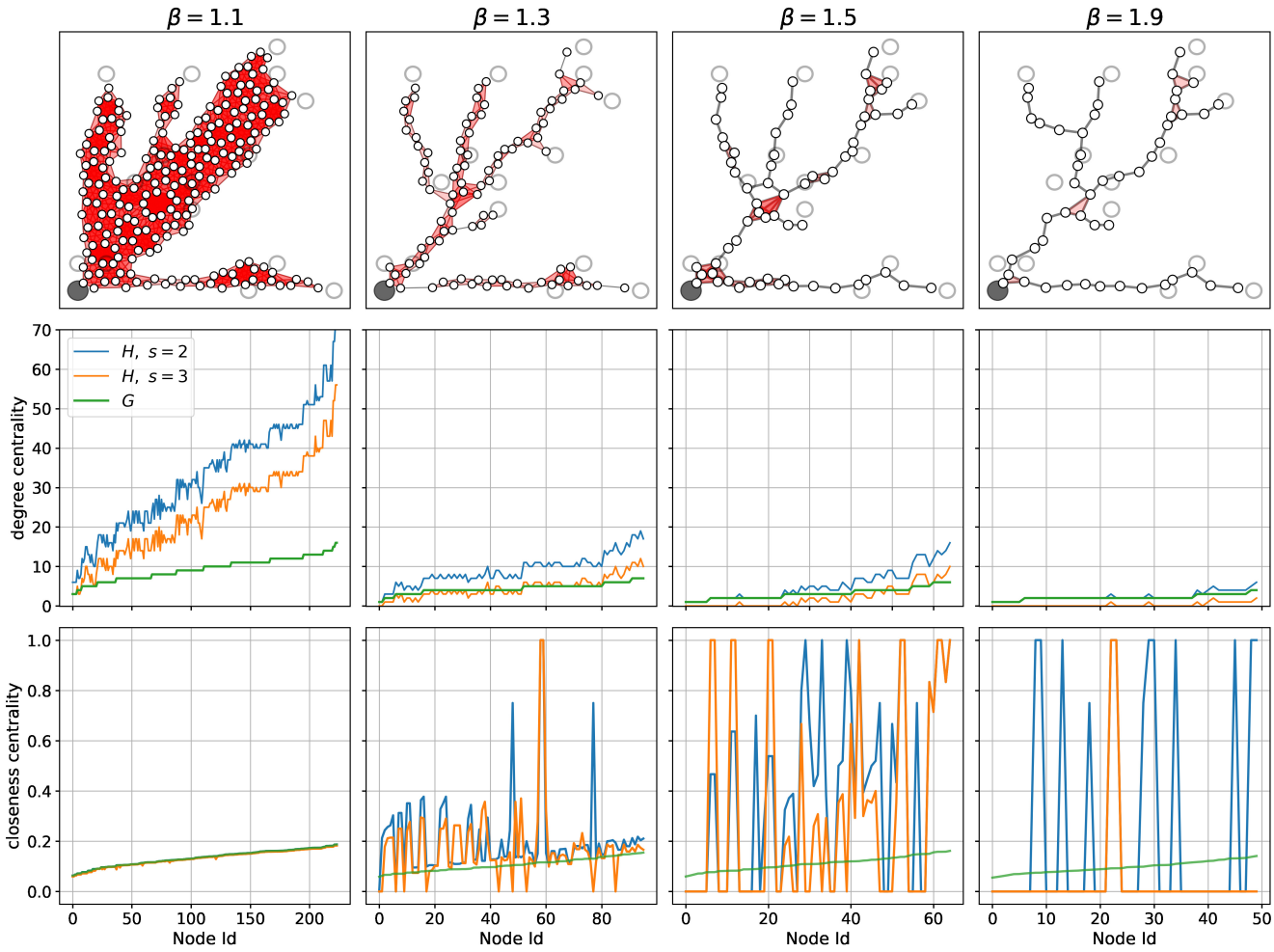

图 4：**图与超图性质**。顶行：采用不同交通率得到的最优超网络。中间行和底行：顶行所示超网络及其 1-骨架的度分布和接近中心性分布。中间行和底行 x 轴上的节点标签按其中心性度数值排序。

包含 50 至 80 个高阶结构。

c. 收敛：运输代价与超网络性质。如 [3], 所示，我们比较了超网络性质与动力学所最小化的代价函数，并突出显示了收敛时间（图 5）。我们重点考察性质 S，即 H中三角形覆盖的曲面面积。该量同时受三角形（进而也受超边）数量及其空间分布的影响。因此，它是衡量超图性质如何随迭代时间变化以及随 β 调整而变化的良好代理指标。我们观察到，在所有情况下，$t_P > t_{\mathcal{L}}$，也就是说，运输代价先于拓扑性质收敛。在其他 $\beta \in [1.1, 1.9]$ 值下以及其他网络性质中也能看到类似的行为（见附录）。与基于DMK的网络性质类似，覆盖面积在 β 取最小值时衰减更快。这符合预期，因为 $\mathcal{L}$ [12–14]具有凸性。然而，运输代价的衰减更快，因此在拥堵运输情况下（ β 较小）， S 的值仍远未收敛。

请注意， S 在最初几次迭代后保持稳定，随后开始以不同的速率（取决于 β）下降，直到达到收敛值。这表明，在演化初期，动力学倾向于形成覆盖大面积的粗分支，随后再逐渐将其压缩，直至达到最优拓扑。

S 和 $\mathcal{L}$ 的这些不同收敛速率可能会阻碍收敛超网络拓扑结构的构建：如果求解器在 $t_{\mathcal{L}} < t_P$时停止，所得超图 $H(\mu_t)$ $t = t_{\mathcal{L}}$ 会错误地覆盖比其收敛对应图更大的曲面（$H(\mu_t)$，对于 $t \geq t_P$）。当 β取值较大时，这种情况的影响较小；不过在这些情况下，由于高阶结构的数量较少， H 与正则图更为相似

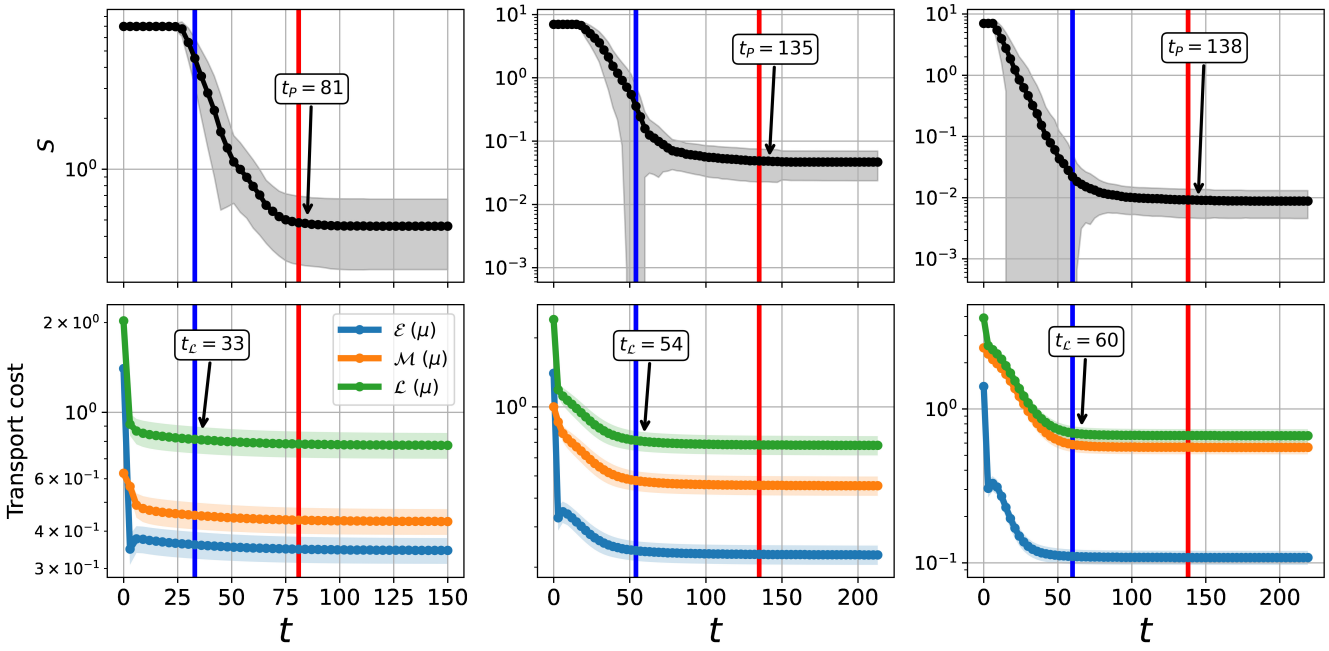

图5：**覆盖面积和李雅普诺夫代价**。覆盖面积 S （上图）以及李雅普诺夫代价、能量耗散 $\mathcal{E}$ 和结构代价 $\mathcal{M}$ （下图）随时间 t变化的均值（标记）和标准差（标记周围的阴影）。均值和标准差根据合成最优传输问题段落中描述的集合计算得出。从左到右： $\beta = 1.2$，1。5 和 1。8。红线和蓝线分别表示 t_P 和 $t_{\mathcal{L}}$。

结构 。收敛后的超网络之间的拓扑差异可见于图4。
最后，我们观察到 $t_{\mathcal{L}}(\beta)$ 和 $t_P(\beta)$ 都是 β上的递增函数。这是预期的，因为交通率越高，序列收敛所需的时间就越长。这一特定行为与 $t_{\mathcal{L}}$情形下 [3]所示的情况相符，但 $t_P(\beta)$ 并非如此：在网络情形中观察到了非单调行为。

d. 超网络性质的收敛行为。图6展示了各种网络性质如何随交通率变化。对于固定的β值，均值和标准差是跨时间计算得出的。如图所示，超边数量、三角形数量、覆盖面积和平均1度均随 t增大而呈现下降趋势。因此，传输最优性可视为使上述超网络性质达到最小状态。值得注意的是，这些函数的另一项显著特征与实际收敛值有关： β 越大，这些指标就越小。这是因为代价函数会越来越倾向于将路径合并到更少的边上。还请注意，这些收敛值之间的差距表明，动力学输出与这些性质之间存在非线性依赖关系。例如，对于 $\beta = 1.1$. 得到的收敛超网络，如果随后将交通率设为1.2，超边数量会大幅减少；而如果将 $\beta = 1.2$ 增至1.3，这种减少就不会那么明显。这些差距的性质因具体性质而异，差别相当显著。这也表明，某些性质更能揭示不同最优交通状态之间的差异。

接近中心性的变化与其他性质明显不同。虽然对于 β 的所有取值，其初始值都相同（与前述性质类似），但随着时间推移，并未呈现出明显趋势。对于 $s = 1$，平均而言， $\beta = 1.1$ 生成的序列往往会先升高再降低，最终恢复到初始值。对于其他交通率，我们观察到了不同的模式。请注意，对于 $s = 1$ ，超图上的 $s-$ 接近中心性与其骨架上的经典接近中心性相同。因此，这些颇为嘈杂的模式并非由添加超边造成。另一方面，对于 $s = 2$ ，平均中心性呈现上升曲线。这可能是因为 L_s 日益分裂为规模较小的连通分量。因此， s 越大，节点看起来就越接近（见图 3）。此外，在这种情况下，较小的 β 会带来更稳定的接近中心性值，体现了 β 对构建高阶结构的影响。虽然 β 的不同取值会导致超图性质呈现不同变化（例如，随着 β 增大，度和超边数量会减少），但我们认为， β 的取值应取决于具体应用。这里的分析展示了这一选择可能如何影响最终形成的拓扑结构。这有助于从业者直观了解其对底层基础设施运输属性下游分析可能造成的影响。

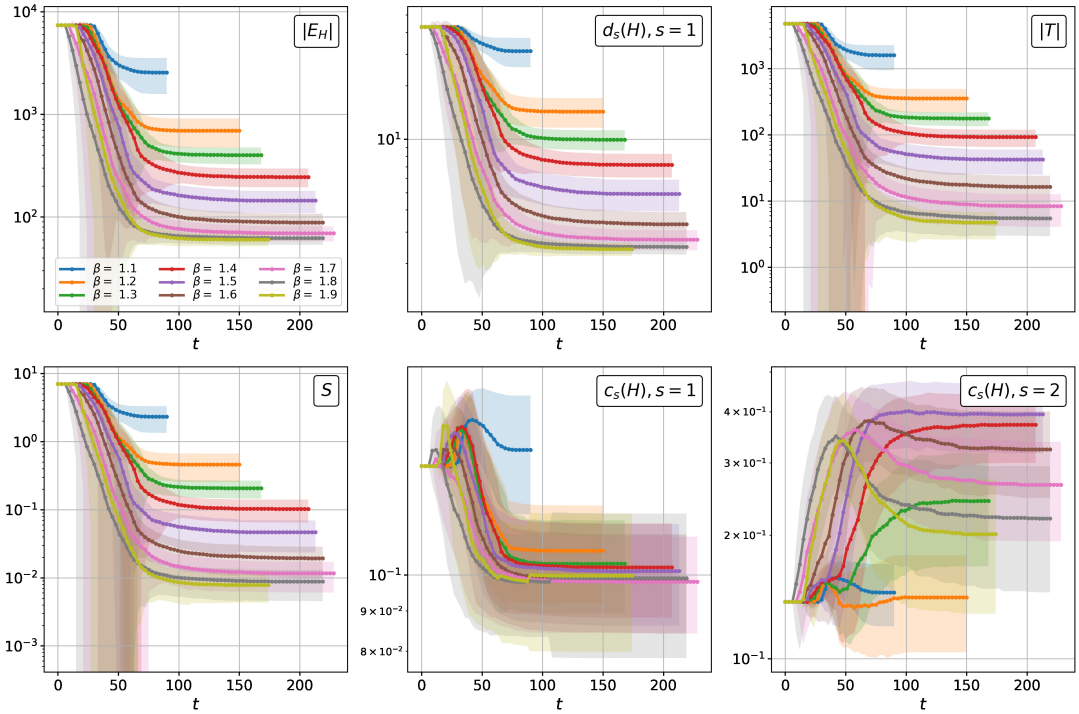

图 6：**超网络性质的演化**。对于不同的 β 值，计算超边数量 $|E_H|$（左上）、三角形数量 $|T|$（上中）、覆盖面积 S（H）（右上）、平均二阶度 d_2（H）（左下）、平均一阶接近中心性 c_1（H）（下中）和二阶接近中心性 c_2（H）（右下）随时间的变化。均值以标记表示，标准差以标记周围的阴影表示。

多头绒泡菌 **超网络**

现在，我们分析从真实数据图像中提取的超网络。我们关注由真实黏菌图像中提取的序列 $\{H(\mu_t)\}_{t=0}^{T}$ 中三角形覆盖面积的演变。多头绒泡菌这种生物的行为，启发了文中所述 DMK 方程的建模思路。研究表明，这些黏菌在二维表面 [17, 25, 26] 上觅食时，会遵循与 DMK 动力学所描述的相似优化策略。我们按照文中所述思路从图像中提取超网络，但不同于应用 [5] 来获得网络，我们采用 [4] 提出的方法，该方法以图像作为输入。此流程利用不同图像像素的颜色强度构建图，并连接相邻的有效节点。我们重点研究来自黏菌图存储库 [10] 的 4 个图像序列。这些序列展现了置于矩形培养皿中的多头绒泡菌 的活动过程。每张图像，也即每个超网络，都是对该生物在 120 秒时间段内运动情况的一个快照。

我们研究了 4 个序列各自的覆盖面积，并在图 7 中绘出了其中一个序列（即图像集 *motion12*；见附录）的结果。我们选取属性序列中的 4 个时间点，展示所用图像及其对应的超网络。左下角的图展示了所研究快照中的一个子区域。如图所示，该子超网络拓扑包含大量维度为 3 的超边，主要分布在黏菌最粗的部分周围。另一方面，右下角的图显示，S 总体上随时间推移而减少（其他序列也得到了类似结果，见附录）。这表明，随着多头绒泡菌 演化至巩固状态，其较厚的身体部分往往会变薄。这种模式与上文合成数据所展示的情况相似，即覆盖面积会随着时间推移而减小，行为与基于 DMK 的超网络序列类似。这表明，DMK 模型能够真实地反映真实黏菌 [10] 朝最优性巩固的阶段。

结论

我们提出了一种从 OT序列 构建高阶结构的方法。该方法将序列中的每个成员映射为一个超图，从而生成时序超网络。我们分析了这些时序族的标准超图性质，并将其与连续对应物进行了比较。我们发现，按运输成本衡量的收敛往往比按超网络覆盖面积衡量的收敛更快。这表明

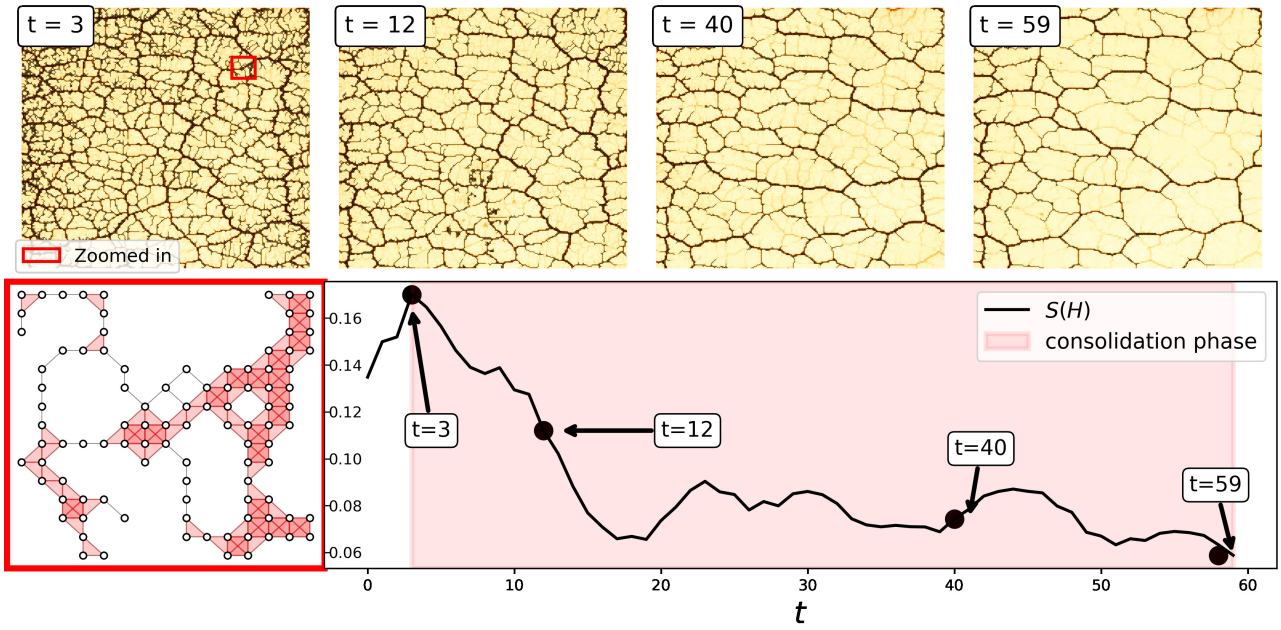

图 7：多头绒泡菌 **超图**。 上方：多头绒泡菌 图像及从中提取的超网络。左下：上方红色矩形框内所示超图的局部放大图。右下：覆盖面积随时间的变化。红色阴影表示向最优性趋近时可能出现的巩固阶段。

用于求解最优传输问题的动力学会将转移质量集中到主分支中；完成这项任务后，主分支覆盖的面积会略微减小。我们研究了这一性质及其他超图性质，并将其与网络版本的相应性质进行了比较。在某些情况下，超网络能够揭示更多关于收敛时拓扑的信息。这表明，对于某些传输方案，超网络或许是表示最优传输问题解的一种更优选择。本研究得出的结论或可进一步加深我们对最优传输解的理解，以及对这一领域与超图领域之间联系的认识。

e. 致谢 作者感谢马克斯·普朗克智能系统国际研究生院 (IMPRS-IS) 对 Diego Baptista 的支持。

[1] Aksoy, S.G., Joslyn, C., Marrero, C.O., Praggastis, B., Purvine, E.：通过高阶超图随机游走研究超网络科学。EPJ Data Science 9(1), 16 (2020) [2] de Arruda, G.F., Petri, G., Moreno, Y.：超图上的社会传染模型。Phys. Rev. Research 2, 023032 (2020年4月 2020 10 1103 2 023032)，https://link.aps.org/doi/ . / [3] PhysRevResearch. . Baptista, D., Bacco, C.D.：基于最优传输的时序网络的收敛性质。载于：International Conference on Complex Networks and Their Applications。第 578–593 页。Springer (2021年) [4] Baptista, D., De Bacco, C.：基于原理的图像网络提取。R. Soc. open sci. (2021年) [5] Baptista, D., Leite, D., Facca, E., Putti, M., De Bacco, C.：通过路径优化进行网络提取。Scientific reports 10(1), 1–13 (2020) [6] Battiston, F., Cencetti, G., Iacopini, I., Latora, V., Lucas, M., Patania, A., Young, J.G., Petri, G.：超越成对相互作用的网络：结构与动力学。Physics Reports 874, 1–92 (2020) [7] Carletti, T., Battiston, F., Cencetti, G., Fanelli, D.：超图上的随机游走。Phys. Rev. E 101, 022308 (2020年2月), https://link.aps.org/doi/ 10 1103 101 022308 . /PhysRevE. . Chowdhary, S., Kumar, A., Cencetti, G., [8] Iacopini, I., Battiston, F.：时序高阶网络中的单纯形传染。Journal of Physics: Complexity 2(3), 035019 (2021年) [9] Cuturi, M.：Sinkhorn距离：以闪电般速度计算最优传输。Advances in Neural Information Processing Systems 26 (2013) [10] Dirnberger, M., Mehlhorn, K., Mehlhorn, T.：介绍黏菌图存储库。Journal of Physics D: Applied Physics 50(26), 264001 (2017) [11] Evans, L.C., Gangbo, W.：Monge-Kantorovich质量传输问题的微分方程方法。第 653 号，美国数学学会 (1999)

[12] Facca, E., Cardin, F., Putti, M.：迈向稳态蒙日–康托罗维奇动力学：多头绒泡菌的经验。SIAM 应用数学杂志 78(2), 651–676 (2018) [13] Facca, E., Cardin, F., Putti, M.：连续最优传输模型中涌现的分支结构。计算物理学杂志 447, 110700 (2021) [14] Facca, E., Daneri, S., Cardin, F., Putti, M.：通过动态公式求解蒙日–康托罗维奇方程的数值方法。J Sci Comput 82(68), 1–26 (2020年2月) [15] Kantorovich, L.：论质量转移（俄文）37, 227–229 (1942) [16] Leite, D., De Bacco, C.：揭示城市交通网络与基于最优传输的基础设施之间的相似性。arXiv 预印本 arXiv:2209.06751 (2022)[17] Nakagaki, T., Yamada, H., T´oth, ´A.：变形虫状生物的迷宫求解。《自然》407(6803), 470–470 (2000) [18] Neuh¨auser, L., Mellor, A., Lambiotte, R.：网络化系统中的多体相互作用与非线性共识动力学。Phys. Rev. E 101, 032310 2020 10 1103 101 032310 (3月), [19] 6 1 1 6 2017 https://link.aps.org/doi/ . /PhysRevE. . Patania, A., Vaccarino, F., Petri, G.：数据拓扑分析。EPJ Data Science ( ), – ( ) [20] Peyr´e, G., Cuturi, M., 等：计算最优传输：及其在数据科学中的应用。Foundations and Trends® in Machine Learning 11(5-6), 355–607 (2019) [21] Santambrogio, F.：最优通道网络、景观函数与分支传输。《界面与自由边界》9(1), 149–169 (2007) [22] Santambrogio, F.：应用数学家的最优传输。Birk¨auser, NY 55, 58–63 (2015) [23] Schaub, M.T., Benson, A.R., Horn, P., Lippner, G., Jadbabaie, A.：单纯复形上的随机游走与归一化霍奇 1-拉普拉斯算子。《SIAM 评论》62(2), 353–391 (2020) [24] Taylor, D., Klimm, F., Harrington, H.A., Kram´ar, M., Mischaikow, K., Porter, M.A., Mucha, P.J.：传染图的拓扑数据分析：考察网络上的传播过程。《自然·通讯》6(1), 1–11 (2015) [25] Tero, A., Kobayashi, R., Nakagaki, T.：真正的黏菌通过路径寻找进行自适应传输网络构建的数学模型。《理论生物学杂志》244(4), 553–564 (2007) [26] Tero, A., Takagi, S., Saigusa, T., Ito, K., Bebber, D.P., Fricker, M.D., Yumiki, K., Kobayashi, R., Nakagaki, T.：受生物启发的自适应网络设计规则。《科学》327(5964), 439–442 (2010) [27] Villani, C.：最优传输：旧与新，第338卷。Springer (2009) [28] Westendorf, C., Gruber, C., Grube, M.：不同黏菌原质团网络的定量比较。载于：第9届 EAI 生物启发信息与通信技术国际会议（原 BIONETICS）论文集。第611–612页 (2016) [29] Xia, Q.：与传输问题相关的最优路径。《当代数学通讯》5(02), 251–279 (2003) [30] Xia, Q.：与传输路径相关的景观函数。Discrete Contin. Dyn. Syst 34(4), 1683–1700 (2014) [31] Xia, Q.：分枝最优传输的动机、思想与应用。ESAIM Math. Model. Numer. Anal. 49(6), 1791–1832 (2015)

附录

β 的其他取值对应的覆盖面积

本节中，我们针对其他 β. 值绘制了与图 5 类似的图，比较覆盖面积和代价函数。正如其中所述， S 呈现出递减趋势，$t_P > t_{\mathcal{L}}$ 成立（见图 8）。

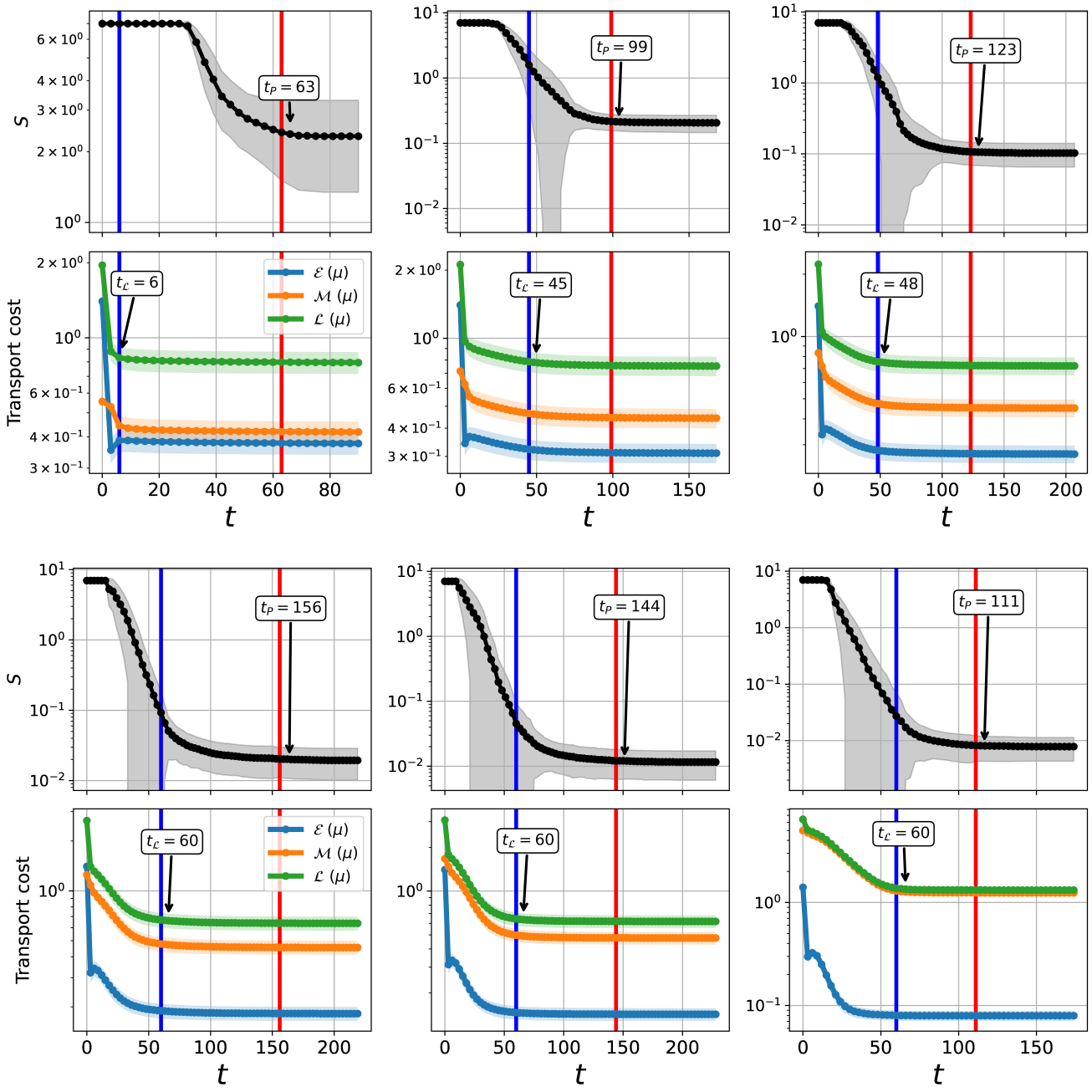

图 8： S **与李雅普诺夫代价**。从上至下第一、二行：从左到右分别为 $\beta = 1.1$、1。3 和 1。4。从上至下第三、四行：从左到右分别为 $\beta = 1.6$、1。7 和 1。9。从上至下第一、三行： S 随时间 t 变化的均值和标准差；从上至下第二、四行：输运密度的李雅普诺夫代价 $\mathcal{L}$、能量耗散 $\mathcal{E}$ 和结构代价 $\mathcal{M}$ 的均值和标准差。红线和蓝线分别表示 $p = 1.05$ 的 t_P 和 $t_{\mathcal{L}}$。

其他超网络性质

本节中，我们进一步比较由动力学最小化的代价函数与超网络性质（见图9）。正如正文所述，在这些情况下也可以观察到类似的单调变化。

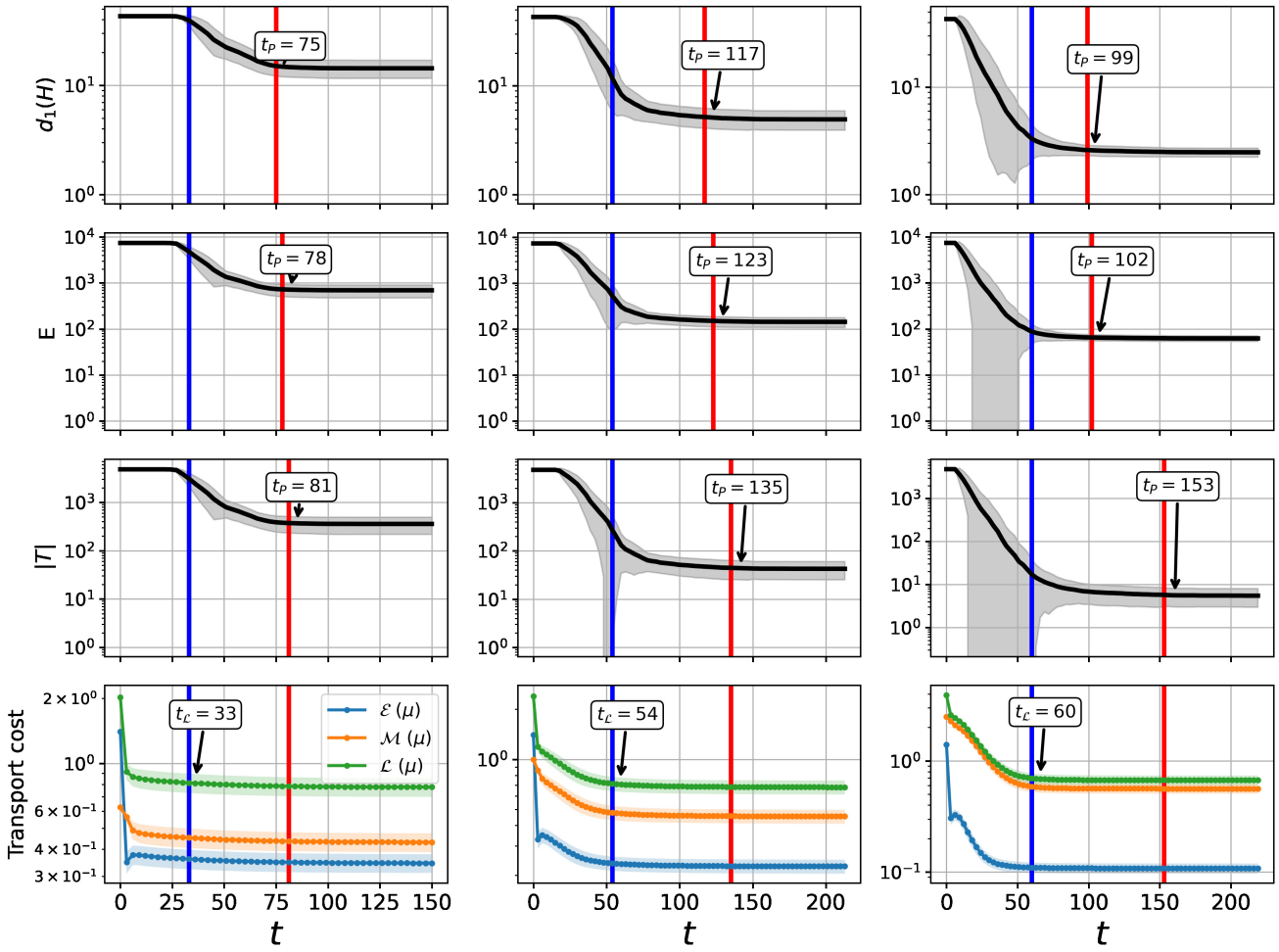

图 9：**其他超网络性质与李雅普诺夫代价**。从左至右： $\beta = 1.2$, 1.5 和 1.8。 从上至下：平均度 $d_1(H)$ 的均值和标准差、超边数量 $|E_H|$，三角形数量 $|T|$，以及李雅普诺夫代价 $\mathcal{L}$、能量耗散 $\mathcal{E}$ 和结构代价 $\mathcal{M}$。红线和蓝线分别表示 t_P和 $t_{\mathcal{L}}$ 对应 $p = 1.05$的值。

多头绒泡菌 **超网络**

f. 数据信息。 本节进一步说明所分析真实数据的详细信息。
这些图像取自黏菌图存储库 [10]，如主文稿中所述。我们研究了4 $\{H_i\}_i^T$不同长度的序列。序列长度（T）取决于序列中包含的图像数量。这是因为不同的实验需要的拍摄次数有所不同。正如代码仓库文档所述，这些实验是在培养皿中放入黏菌进行的，培养皿内铺有一层薄琼脂，且不提供食物。据创建者解释，由于没有食物，黏菌会不偏不倚地探索培养皿。由于这种生物最初被放置在矩形培养皿的一条短边上，因此当原生质团到达另一条短边时，实验便告结束。此后不再拍摄照片。

g. 超图提取。我们使用代码仓库中的图像集*motion12*、*motion24*、*motion40*和$motion79$构建所研究的超网络。这些图像集包含的图像数量从 60 张到 150 张不等。随后使用$Img2net$ 算法提取超网络，算法如 [4] 中所述，主文稿中也有提及；所用配置与 [3]中描述的相同。

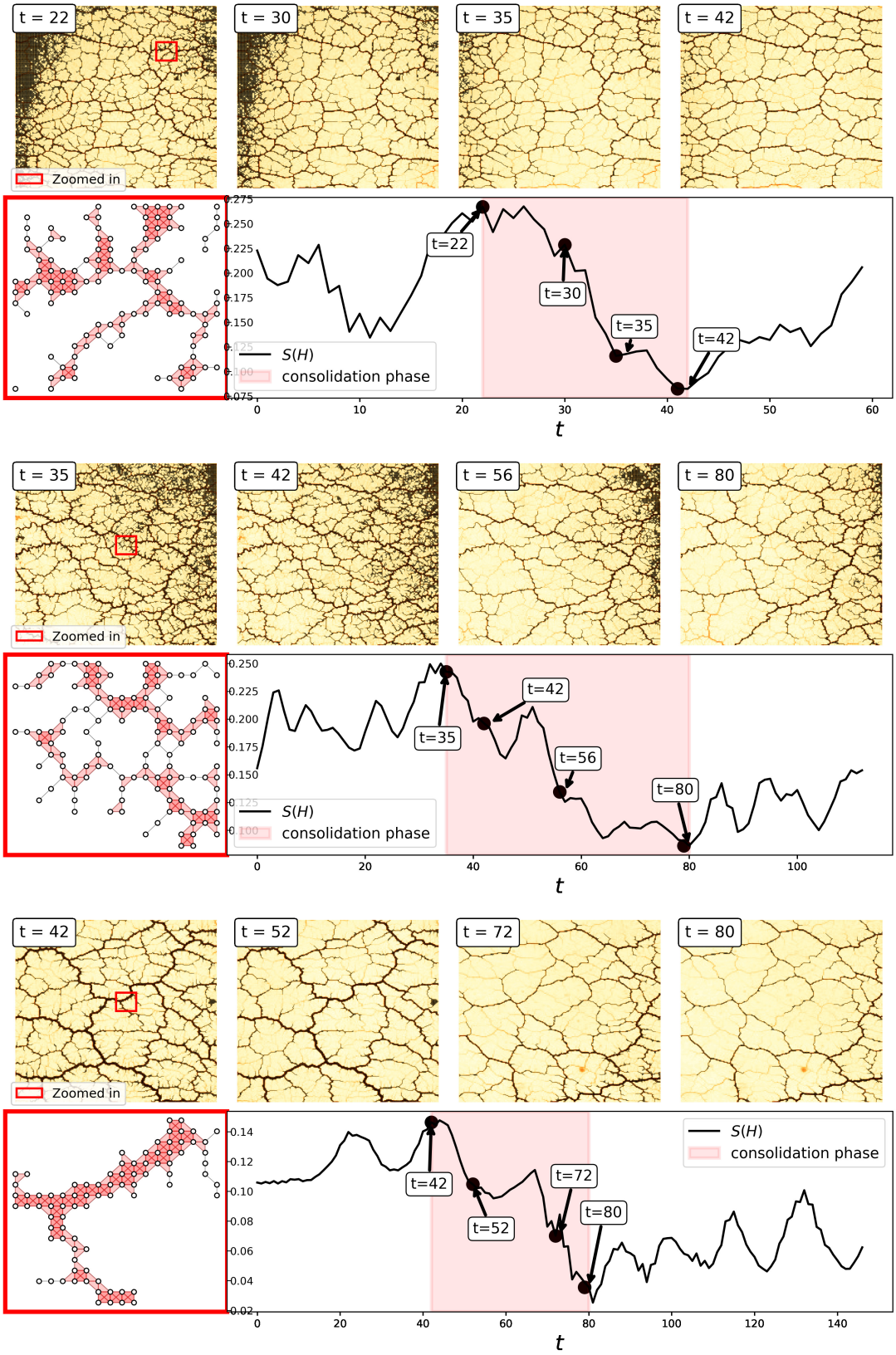

图 10：多头绒泡菌 S **演化**。从上到下：*motion24*、*motion40* 和*motion79*。图表成对排列。每对图表中，上方的图展示了多头绒泡菌图像及从中提取的超网络。左下方图表中的网络是上方红色矩形内所示超图的一个子部分。底部图表展示了覆盖面积随时间的变化。图中红色阴影部分标示了朝最优性迈进的初步巩固阶段。