

THE MESH EQUILIBRIUM

*How Heterogeneous Specialized Agents Self-Organize
to Exceed Centralized Provision After the Crossing Point*

Jon Smirl

Independent Researcher

February 2026

WORKING PAPER

Abstract

Smirl (2026a) models how concentrated AI infrastructure investment finances the learning curves that enable distributed inference. That paper ends at the crossing point—the moment the distributed architecture becomes cost-competitive. This paper answers: what happens after the crossing? The answer is not billions of isolated devices running local inference. It is a self-organizing mesh of heterogeneous specialized agents whose collective capability exceeds centralized provision once the mesh reaches critical mass.

Four contributions are new. First, I formalize the mesh equilibrium as a three-layer model—percolation-based critical mass, CES-aggregated heterogeneous specialization, and Laplacian knowledge diffusion—and prove the existence, uniqueness, and local asymptotic stability of the mesh equilibrium (Theorem 1). Second, I show that the phase transition from centralized to distributed provision is *first-order* (discontinuous) when the number of specialization types exceeds two, using the Fortuin-Kasteleyn representation of the q -state Potts model. The mesh does not form gradually—it crystallizes. Third, I connect the crossing point $x(t) = 0$ from the predecessor paper to an inverse Bose-Einstein condensation in the Bianconi-Barabási network fitness model: the packaging learning curve broadens the fitness distribution, dissolving the centralized

condensate. Fourth, I derive that the mesh’s routing and compensation requirements endogenously generate the need for a programmable settlement layer—connecting to a separate analysis of monetary infrastructure (Smirl 2026b, forthcoming). The model generates six falsifiable predictions with timing and failure conditions.

Keywords: mesh equilibrium, phase transition, Bose-Einstein condensation, heterogeneous specialization, CES aggregation, Potts model, percolation, self-organizing networks, distributed AI inference

JEL: D85, L14, O33, C62, C73

1. Introduction

Smirl (2026a) formalizes endogenous decentralization: concentrated capital investment in centralized AI infrastructure finances the component learning curves—particularly in 3D memory stacking and advanced packaging ($\alpha = 0.23$)—that enable distributed alternatives to replicate datacenter-class inference on consumer hardware. The model’s state variable $x(t) = \bar{Q}_{\text{eff}} - Q(t)$ measures remaining cumulative production until the crossing threshold at which distributed architecture becomes cost-competitive. When $x(t) = 0$, the basic reproduction number R_0 of the distributed ecosystem crosses unity, and the transition becomes self-sustaining and irreversible.

That paper ends at the crossing point. This paper begins there.

A natural but incorrect conjecture is that post-crossing dynamics are simply “inference runs locally.” Isolated devices running local models do not constitute an economic equilibrium. A single consumer device, however capable, cannot match the breadth of a centralized provider serving millions of queries across every domain. The question is whether there exists an organizational form—emergent, not designed—in which distributed devices collectively exceed centralized provision.

This paper argues that such a form exists: a *mesh* of heterogeneous specialized agents that self-organize through local interactions into a division of labor whose aggregate capability exceeds any centralized system. The mesh equilibrium has three defining properties. First, *connectivity*: agents form a giant connected component through which queries can be routed to specialists. Second, *heterogeneous specialization*: agents develop complementary capabilities through adaptive threshold dynamics, so that the collective covers a broader capability space than any single agent. Third, *knowledge diffusion*: information propagates across the network at a rate governed by the graph Laplacian’s spectral gap, and is self-sustaining on scale-free topologies at any nonzero transmission rate.

The formal model has three composable layers. Layer 1 (Section 3) uses percolation theory to establish the existence of a connected mesh when $R_0^{\text{mesh}} > 1$, where R_0^{mesh} inherits directly from the predecessor paper’s generalized crossing condition. Layer 2 (Section 4) introduces heterogeneous agents with CES-aggregated capabilities and adaptive specialization dynamics following the Bonabeau-Theraulaz fixed response threshold model. Layer 3 (Section 5) characterizes knowledge diffusion on the mesh using the graph Laplacian and establishes self-sustaining propagation on scale-free networks via the Pastor-Satorras–Vespignani vanishing epidemic threshold.

The central theorem (Section 6) unifies the three layers: for $R_0^{\text{mesh}} > 1$ and CES substitution parameter $\rho < 1$, there exists a finite critical mass N^* such that for all $N > N^*$, the

mesh equilibrium exists, is unique, is locally asymptotically stable, and $C_{\text{mesh}}(N) > C_{\text{cent}}$. The critical mass N^* is decreasing in the diversity of the agent population.

Three mathematical results from network science, statistical mechanics, and theoretical ecology provide the formal foundations.

The first is the Bianconi-Barabási (2001) fitness model, which exhibits two phases depending on the shape of the node fitness distribution $\rho(\eta)$: a winner-takes-all Bose-Einstein condensation when $\rho(\eta)$ is sharply peaked (the centralized equilibrium), and a fit-get-rich distributed phase when $\rho(\eta)$ is broad (the mesh equilibrium). The predecessor paper’s packaging learning curve broadens the fitness distribution by enabling edge devices to achieve fitness levels previously restricted to datacenter-scale nodes. The crossing point $x(t) = 0$ corresponds to the system reaching the BEC phase boundary. This is *inverse* Bose-Einstein condensation: the condensate dissolves as the technology parameter θ improves.

The second is the Fortuin-Kasteleyn (1972) representation, which reveals that percolation and specialization are the same mathematical object evaluated at different values of the Potts model state parameter q . For $q > 2$ specialization types, the phase transition is first-order (discontinuous). The mesh does not form gradually—it crystallizes abruptly once conditions are met. This is a testable prediction distinguishable from smooth adoption curves.

The third is the Pastor-Satorras–Vespignani (2001) result: on scale-free networks with degree exponent $\gamma \leq 3$, the epidemic threshold vanishes. Any nonzero rate of knowledge sharing sustains itself indefinitely on a heterogeneous network, because $\langle k^2 \rangle$ diverges and $\lambda_c = \langle k \rangle / \langle k^2 \rangle \rightarrow 0$. The topology does the work.

A methodological observation unifies these results. The self-consistency equations across percolation, epidemiology, Ising/Potts models, ecology, and network economics all reduce to the same form: $m = f(m; \lambda)$ with a transcritical bifurcation at the critical parameter value. The R_0 from the predecessor paper *is* the percolation threshold *is* the Ising critical temperature. One equation, six notations. The crossing point $x(t) = 0$ is the moment R_0^{mesh} crosses unity.

The paper proceeds as follows. Section 2 reviews the predecessor paper’s terminal conditions. Section 3 develops the percolation layer. Section 4 develops heterogeneous specialization. Section 5 develops knowledge diffusion. Section 6 states and proves the central theorem. Section 7 characterizes the post-crossing dynamics in three phases. Section 8 derives the settlement layer requirement. Section 9 discusses frameworks considered and rejected. Section 10 presents falsifiable predictions. Section 11 concludes. Appendix A provides a remark on renormalization group universality that justifies the mean-field modeling approach.

2. Terminal Conditions from the Predecessor Paper

This section summarizes the results from Smirl (2026a) that serve as initial conditions for the present analysis. No new results are presented; the purpose is notational continuity.

2.1 The Crossing Point

The state variable $x(t) = \bar{Q}_{\text{eff}}(\eta(t)) - Q(t)$ measures remaining cumulative production until the effective crossing threshold. The threshold \bar{Q}_{eff} is approached from two directions: hardware costs decline from below along the packaging learning curve ($\alpha = 0.23$), and algorithmic efficiency gains reduce the threshold from above. The crossing event occurs at T^* , the first date at which $x(T^*) = 0$.

At T^* , the following conditions hold:

- (i) *Cost parity*: distributed inference cost per token equals centralized cost per token for the marginal user.
- (ii) $R_0 \geq 1$: the distributed ecosystem’s basic reproduction number, $R_0 = \beta(c, \lambda) \cdot \gamma / (\kappa + \mu)$, crosses unity. Self-sustaining adoption begins.
- (iii) *Irreversibility*: cumulative production $Q(t)$ is monotonically increasing; once Q crosses \bar{Q} , reversion requires cumulative production to decrease, which is impossible.
- (iv) *Training persistence*: frontier model training remains centralized because the binding constraint is topological (synchronization bandwidth), not cost-based.

2.2 Initial State of the Mesh

At $t = T^*$, the “mesh” consists of a population of heterogeneous devices capable of running inference workloads locally but not yet organized into a connected network with routing, specialization, or coordination. The number of capable devices $N(T^*)$ is positive but below critical mass. The predecessor paper’s $R_0 > 1$ condition ensures that the population grows; the present paper characterizes what it grows *into*.

The technology parameter $\theta(t)$ continues to evolve after the crossing. As θ increases, additional device types become inference-capable, broadening the fitness distribution—the mechanism that drives inverse BEC condensation in Layer 1.

3. Layer 1: Phase Transition and Critical Mass

3.1 The R_0^{mesh} Framework

The R_0 framework from Smirl (2026a) provides the boundary condition for the mesh. Define the mesh reproduction number:

$$R_0^{\text{mesh}} = N \cdot \beta \cdot v / D \quad (1)$$

where N is the number of active nodes, β is the connection probability per node pair, v is the expected value per interaction, and D is the attrition rate (encompassing coordination friction κ and churn μ from the predecessor paper). The mesh is self-sustaining when $R_0^{\text{mesh}} > 1$.

3.2 Giant Component Existence

The fraction of nodes belonging to the giant connected component satisfies:

$$S_\infty = 1 - \exp(-R_0^{\text{mesh}} \cdot S_\infty) \quad (2)$$

Proposition 1 (Giant Component Existence and Uniqueness). *Equation (2) has:*

- (a) *the trivial solution $S_\infty = 0$ for all R_0^{mesh} ;*
- (b) *a unique positive solution $S_\infty^* \in (0, 1)$ if and only if $R_0^{\text{mesh}} > 1$;*
- (c) *S_∞^* is locally asymptotically stable as a fixed point of the iteration $S_\infty \mapsto 1 - \exp(-R_0^{\text{mesh}} \cdot S_\infty)$.*

Proof. Define $g(s) = 1 - \exp(-R_0^{\text{mesh}} \cdot s) - s$. At $s = 0$: $g(0) = 0$ and $g'(0) = R_0^{\text{mesh}} - 1 > 0$ when $R_0^{\text{mesh}} > 1$, so g is initially increasing. At $s = 1$: $g(1) = -\exp(-R_0^{\text{mesh}}) < 0$. By the intermediate value theorem, g has a root $S_\infty^* \in (0, 1)$.

For uniqueness: $g''(s) = -R_0^{\text{mesh}2} \exp(-R_0^{\text{mesh}} s) < 0$, so g is strictly concave. A strictly concave function crossing zero from above can have at most one positive root.

For stability: the iteration map $h(s) = 1 - \exp(-R_0^{\text{mesh}} s)$ satisfies $h'(S_\infty^*) = R_0^{\text{mesh}} \exp(-R_0^{\text{mesh}} S_\infty^*) = R_0^{\text{mesh}}(1 - S_\infty^*)$. At the positive fixed point with $R_0^{\text{mesh}} > 1$, $0 < h'(S_\infty^*) < 1$ (since $S_\infty^* > 0$ implies $1 - S_\infty^* < 1$ and $R_0^{\text{mesh}}(1 - S_\infty^*) = -g'(S_\infty^*)/1 < 1$ by concavity), confirming local asymptotic stability by the contraction mapping principle. \square

3.3 The Fortuin-Kasteleyn Unification

The giant component result establishes *connectivity*. The question of whether the mesh develops *specialization*—a division of labor among heterogeneous agents—might appear to require a separate model. The Fortuin-Kasteleyn (1972) representation reveals that these are the same mathematical object.

The partition function of the q -state Potts model on graph $G = (V, E)$ has the exact cluster expansion:

$$Z_{\text{Potts}} = \sum_{A \subseteq E} p^{|A|} (1 - p)^{|E| - |A|} \cdot q^{c(A)} \quad (3)$$

where $p = 1 - e^{-\beta_T J}$ is the bond occupation probability (with β_T the inverse temperature and J the coupling), A is a subset of edges, and $c(A)$ is the number of connected components in the subgraph (V, A) .

At $q = 1$, this reduces to the generating function for bond percolation. The Potts model at general q describes a system in which nodes adopt one of q states (specializations) and prefer to match their neighbors. The FK representation shows that percolation and specialization are controlled by the same cluster structure evaluated at different q .

Proposition 2 (First-Order Crystallization). *For $q > 2$ specialization types on a graph with mean degree $\langle k \rangle > 1$, the specialization transition is first-order: the order parameter (fraction of nodes in the dominant specialization cluster) jumps discontinuously from zero to a positive value at the critical coupling $\beta_T J = \beta_c(q)$.*

This is a classical result in statistical mechanics (Baxter 1982; Wu 1982). The economic content is that the mesh does not form gradually. When the parameter p (interpretable as the probability that two neighboring agents successfully coordinate on a task division) crosses the threshold, division of labor crystallizes abruptly. Early mesh growth appears stochastic and fragile; the crystallization event is sudden.

For $q = 2$ (two specialization types), the transition is second-order (continuous)—the standard Ising universality class. This means the first-order prediction is specific to settings with $q \geq 3$ distinct specialization roles, which is the empirically relevant case for AI inference (coding, creative writing, mathematical reasoning, multimodal processing, domain-specific knowledge, real-time translation, etc.).

3.4 Inverse Bose-Einstein Condensation

The Bianconi-Barabási (2001) model assigns each node i a fitness η_i drawn from distribution $\rho(\eta)$. The degree (number of connections) of node i evolves as:

$$k_i(t) \sim \left(\frac{t}{t_i} \right)^{\eta_i/C} \quad (4)$$

where t_i is the node's arrival time and C satisfies the self-consistency equation:

$$\int \frac{\rho(\eta)}{C/\eta - 1} d\eta = 1 \quad (5)$$

This has the identical mathematical structure to the Bose gas number equation, with C playing the role of the fugacity.

The phase structure depends on $\rho(\eta)$ near its maximum η_{\max} . If $\rho(\eta) \sim (\eta_{\max} - \eta)^{\alpha_B}$ as $\eta \rightarrow \eta_{\max}$:

- $\alpha_B \leq 0$ (sharply peaked): *Bose-Einstein condensation*. The single fittest node captures a macroscopic fraction of all connections. This is the centralized equilibrium—AWS, Google Cloud, and Azure dominate because they are the only nodes with high fitness (massive compute, low latency via CDN, broad model coverage).
- $\alpha_B > 0$ (broad distribution): *Fit-get-rich* phase. Many nodes share traffic proportional to their fitness. No single node dominates. This is the mesh equilibrium.

Proposition 3 (Learning-Curve-Driven Phase Transition). *Let the technology parameter $\theta(t)$ index the packaging learning curve output from Smirl (2026a). Suppose the fitness of an edge device of type j is $\eta_j(\theta) = \eta_j^0 + g_j(\theta)$, where g_j is increasing in θ and $g_j(0) = 0$. If $\theta(0)$ produces a fitness distribution $\rho_0(\eta)$ with $\alpha_B \leq 0$, and $\theta(\bar{t})$ produces $\rho_{\bar{t}}(\eta)$ with $\alpha_B > 0$ for some finite \bar{t} , then the system undergoes a BEC-to-FGR phase transition at the critical θ^* where α_B crosses zero. This is inverse Bose-Einstein condensation: the centralized condensate dissolves.*

Proof. The learning curve increases the fitness of previously low-fitness edge devices. Initially, only datacenter nodes have η near η_{\max} , so $\rho(\eta)$ is sharply peaked at the maximum—the BEC phase. As θ increases, the support of ρ broadens: more device types achieve fitness levels closer to η_{\max} . The exponent α_B characterizing the behavior of ρ near η_{\max} transitions from ≤ 0 to > 0 at $\theta = \theta^*$. By the Bianconi-Barabási classification, this is the BEC phase boundary.

The mapping to the predecessor paper is direct: θ^* corresponds to $x(t) = 0$. At T^* , the fitness distribution has broadened sufficiently that the centralized condensate—the macroscopic concentration of traffic at a single node—dissolves. Traffic distributes across the mesh proportional to fitness. \square

Remark 1. *The BEC framework describes the competitive dynamics of traffic allocation on the network. It does not describe the physical connectivity (Layer 1’s percolation) or the specialization structure (Layer 2’s CES aggregation). The three layers compose: percolation ensures the mesh is connected, BEC dynamics govern how traffic flows within it, and CES aggregation determines whether the mesh’s collective capability exceeds the centralized alternative.*

4. Layer 2: Heterogeneous Specialization

4.1 Agent Capabilities and CES Aggregation

Each agent $i \in \{1, \dots, N\}$ has a capability vector $\mathbf{c}_i = (c_{i1}, \dots, c_{iJ})$ across J task types. The effective capability of the mesh for task type j is the sum of agent capabilities: $C_j = \sum_i c_{ij}$. The aggregate mesh capability is the CES function:

$$C_{\text{eff}}(N) = \left(\sum_{j=1}^J C_j^\rho \right)^{1/\rho}, \quad 0 < \rho < 1 \quad (6)$$

The parameter $\rho < 1$ implies imperfect substitutability across task types. This is the critical modeling choice: because task types are complements ($\rho < 1$), the aggregate rewards *diversity*. Two agents with different specializations contribute more to C_{eff} than two identical agents. Formally, the CES function is strictly concave in each C_j , so transferring capability from an over-represented task type to an under-represented one increases C_{eff} .

Lemma 1 (Diversity Premium). *Fix total capability $\bar{C} = \sum_j C_j$. For $\rho < 1$, C_{eff} is maximized when $C_j = \bar{C}/J$ for all j (equal coverage across task types). More precisely:*

$$C_{\text{eff}}|_{\text{equal}} = J^{(1-\rho)/\rho} \cdot C_{\text{eff}}|_{\text{concentrated}} \quad (7)$$

where the concentrated case places all capability in a single task type. The diversity premium $J^{(1-\rho)/\rho}$ is increasing in J and decreasing in ρ .

Proof. With equal allocation: $C_{\text{eff}} = (J \cdot (\bar{C}/J)^\rho)^{1/\rho} = J^{1/\rho-1} \cdot \bar{C}$. With concentration in one type: $C_{\text{eff}} = \bar{C}$. The ratio is $J^{(1-\rho)/\rho}$, which exceeds 1 for $J \geq 2$ and $\rho < 1$. \square

This is the Becker-Murphy (1992) division of labor result in CES form. The mesh’s advantage over centralized provision does not come from superior individual capability—each edge device is weaker than the datacenter—but from the breadth of specialized coverage that heterogeneous agents collectively provide.

4.2 Centralized Capability Benchmark

A centralized provider operates M identical high-capability units, each with capability \bar{c} spread uniformly across all J task types: $c_j^{\text{cent}} = \bar{c}/J$ per unit. Total centralized capability for task j is $C_j^{\text{cent}} = M\bar{c}/J$, giving:

$$C_{\text{cent}} = \left(J \cdot \left(\frac{M\bar{c}}{J} \right)^\rho \right)^{1/\rho} = J^{(1-\rho)/\rho} \cdot M\bar{c} \quad (8)$$

The centralized provider has fixed capacity $M\bar{c}$ (determined by datacenter investment). The mesh’s aggregate capability grows with N and with the diversity of specialists.

4.3 Specialization Dynamics: The Fixed Response Threshold Model

Agents do not arrive pre-specialized. Specialization emerges endogenously through local interactions. The mechanism follows the Bonabeau-Theraulaz (1998) fixed response threshold model, originally developed for division of labor in social insect colonies.

Agent i has a vector of response thresholds $\theta_i = (\theta_{i1}, \dots, \theta_{iJ})$ for each task type. When demand signal s_j for task j arrives, agent i performs the task with probability:

$$P_{ij}(s_j) = \frac{s_j^n}{s_j^n + \theta_{ij}^n} \quad (9)$$

where $n \geq 2$ is a steepness parameter. The response probability is a sigmoid: high when $s_j \gg \theta_{ij}$, low when $s_j \ll \theta_{ij}$.

Thresholds adapt through reinforcement:

$$\dot{\theta}_{ij} = -\xi \cdot \mathbf{1}[i \text{ performs task } j] + \varphi \cdot \mathbf{1}[i \text{ does not perform task } j] \quad (10)$$

where $\xi > 0$ is the reinforcement rate (performing a task lowers the threshold, increasing future responsiveness) and $\varphi > 0$ is the decay rate (not performing a task raises the threshold).

Proposition 4 (Emergent Specialization). *Under the threshold dynamics (10) with heterogeneous initial thresholds $\theta_{ij}(0)$, agents self-sort into specialist roles: for each agent i , there*

exists $j^*(i) = \arg \max_j c_{ij}(t)$ such that $c_{ij^*}(t) \rightarrow \bar{c}_i$ and $c_{ij}(t) \rightarrow 0$ for $j \neq j^*$ as $t \rightarrow \infty$, where \bar{c}_i is agent i 's maximum achievable capability.

The proof follows from the reinforcement dynamics: an agent that performs task j frequently sees θ_{ij} decrease, making it more responsive to future demand for j , which further increases frequency of performance—a positive feedback loop. Simultaneously, thresholds for other tasks rise. The dynamics converge to a fixed point where each agent responds primarily to one task type. This is the Becker-Murphy division of labor emerging from local interactions without central coordination.

Remark 2 (Connection to Potts Crystallization). *The specialization dynamics of Proposition 4 are the micro-level mechanism underlying the Potts model crystallization of Proposition 2. The Potts “state” of each node is its dominant specialization $j^*(i)$. The “coupling” J in the Potts Hamiltonian corresponds to the task-sharing benefit between neighboring agents with compatible specializations. The first-order crystallization at $q > 2$ means that when the number of specialization types exceeds two, the transition from unspecialized to specialized is abrupt—consistent with the reinforcement dynamics exhibiting a bifurcation from mixed to specialized response profiles.*

5. Layer 3: Knowledge Diffusion

5.1 Laplacian Dynamics

Let $\mathbf{u}(t) \in \mathbb{R}^N$ represent the knowledge state of each node (e.g., model weights, fine-tuning updates, or capability parameters). Knowledge diffusion on the network follows:

$$\frac{\partial \mathbf{u}}{\partial t} = -L \cdot \mathbf{u} \quad (11)$$

where $L = D_{\text{deg}} - A$ is the graph Laplacian, with D_{deg} the diagonal degree matrix and A the adjacency matrix. The convergence rate to the consensus state is governed by $\lambda_2(L)$, the Fiedler eigenvalue (algebraic connectivity).

Lemma 2 (Convergence Rate). *The knowledge state $\mathbf{u}(t)$ converges to the average $\bar{\mathbf{u}} = N^{-1} \sum_i \mathbf{u}_i(0)$ at rate $\lambda_2(L)$:*

$$\|\mathbf{u}(t) - \bar{\mathbf{u}}\mathbf{1}\|_2 \leq \|\mathbf{u}(0) - \bar{\mathbf{u}}\mathbf{1}\|_2 \cdot e^{-\lambda_2(L)t} \quad (12)$$

For connected graphs, $\lambda_2(L) > 0$.

5.2 Bandwidth Scaling

The total bandwidth available for knowledge diffusion in the mesh scales as $B_{\text{mesh}} = O(N \cdot \langle k \rangle)$, where $\langle k \rangle$ is the mean degree. The centralized hub has fixed bandwidth B_{hub} determined by datacenter interconnect capacity. Once:

$$N \cdot \langle k \rangle > B_{\text{hub}} \quad (13)$$

the mesh serves more total queries per unit time than the centralized provider. This is a necessary condition for capability dominance, complementing the CES capability comparison.

5.3 Vanishing Epidemic Threshold on Scale-Free Networks

Pastor-Satorras and Vespignani (2001) established that the SIS epidemic threshold on networks with degree distribution $P(k) \sim k^{-\gamma}$ is:

$$\lambda_c = \frac{\langle k \rangle}{\langle k^2 \rangle} \quad (14)$$

For scale-free networks with $\gamma \leq 3$: $\langle k^2 \rangle$ diverges in the thermodynamic limit, so $\lambda_c \rightarrow 0$.

Proposition 5 (Self-Sustaining Knowledge Propagation). *If the mesh has a scale-free degree distribution with $\gamma \leq 3$ —which MoE routing produces endogenously through preferential specialization—then any nonzero rate of knowledge sharing sustains itself indefinitely. The topology ensures propagation without requiring a minimum transmission rate.*

The economic content is that knowledge diffusion need not be modeled as a separate mechanism with its own threshold. Once the mesh achieves a fat-tailed degree distribution (which the specialization dynamics of Layer 2 produce through preferential attachment to high-quality specialists), capability propagation is guaranteed. Hub agents—the most capable specialists—emerge endogenously and serve as conduits for knowledge transfer.

5.4 Combined Dynamics

The three layers interact as follows. Layer 1 ensures the mesh is connected ($S_\infty > 0$). Layer 2 ensures agents specialize (C_{eff} grows with diversity). Layer 3 ensures knowledge propagates (the Fiedler eigenvalue is positive, and on scale-free topologies, propagation is self-sustaining). The combined effect is a mesh whose aggregate capability $C_{\text{mesh}}(N)$ is superlinear in N over the relevant range, because additional diverse specialists both increase the CES aggregate (Layer 2) and increase the rate at which existing knowledge diffuses to new entrants (Layer 3).

6. The Central Theorem

Theorem 1 (Mesh Equilibrium Existence, Uniqueness, and Dominance). *For $R_0^{\text{mesh}} > 1$ and $\rho < 1$, there exists a finite N^* such that for all $N > N^*$:*

- (a) *The mesh equilibrium exists: a positive fraction $S_\infty^* > 0$ of agents form a connected component with specialized roles covering all J task types.*
- (b) *The mesh equilibrium is unique among equilibria with $S_\infty > 0$.*
- (c) *The mesh equilibrium is locally asymptotically stable.*
- (d) *$C_{\text{mesh}}(N) > C_{\text{cent}}$: the mesh’s aggregate capability exceeds centralized provision.*
- (e) *N^* is decreasing in the diversity of the agent population, measured by the entropy of the fitness distribution $H(\rho) = - \int \rho(\eta) \ln \rho(\eta) d\eta$.*

Proof. The proof proceeds in four steps, one for each component.

Step 1: Existence via percolation (Proposition 1). For $R_0^{\text{mesh}} > 1$, equation (2) has a unique positive solution S_∞^* . The giant component contains $S_\infty^* \cdot N$ agents. For N sufficiently large, this exceeds the minimum number of agents required to cover all J task types (at least one specialist per type), establishing existence.

Step 2: Uniqueness via supermodularity. The mesh participation game—in which each agent decides whether to join the mesh and which task type to specialize in—is a supermodular game. Agent i ’s payoff from joining increases when more agents join (network effect via the CES aggregate and knowledge diffusion) and when agents specialize in complementary types (CES complementarity with $\rho < 1$). By Tarski’s (1955) fixed point theorem, the game has a greatest and least equilibrium. By the strict concavity of the CES function in each C_j , the greatest equilibrium is unique among equilibria with $S_\infty > 0$: any equilibrium with different specialization allocations is payoff-dominated by the efficient allocation.

Step 3: Stability via Lyapunov analysis. Consider the Lyapunov function $V = -C_{\text{eff}}(N) + \sum_j \phi_j(C_j)$ where ϕ_j is the potential function of the specialization dynamics. At the equilibrium, $\dot{V} \leq 0$ with equality only at the fixed point, by the dissipative property of the threshold dynamics (10) and the concavity of C_{eff} . The equilibrium is locally asymptotically stable by LaSalle’s invariance principle.

Step 4: Capability dominance via CES growth. The mesh capability is:

$$C_{\text{mesh}}(N) = \left(\sum_{j=1}^J \left(\sum_{i \in \text{mesh}} c_{ij} \right)^\rho \right)^{1/\rho} \quad (15)$$

Under specialization, each agent i concentrates capability in task $j^*(i)$, so $C_j \approx |\{i : j^*(i) = j\}| \cdot \bar{c}$ where \bar{c} is the mean specialist capability. If agents distribute approximately uniformly across J types (which the specialization dynamics select for, since unmet demand signals attract entrants), then $C_j \approx S_\infty^* N \bar{c} / J$ and:

$$C_{\text{mesh}}(N) \approx J^{(1-\rho)/\rho} \cdot S_\infty^* N \bar{c} / J = J^{1/\rho-2} \cdot S_\infty^* N \bar{c} \quad (16)$$

This exceeds $C_{\text{cent}} = J^{(1-\rho)/\rho} \cdot M \bar{c}_{\text{cent}}$ when $S_\infty^* N \bar{c} / J > M \bar{c}_{\text{cent}} / J$, i.e., when $N > N^* \equiv M \bar{c}_{\text{cent}} / (S_\infty^* \bar{c})$. Since M and \bar{c}_{cent} are fixed (centralized capacity is bounded by datacenter investment), N^* is finite.

Step 5: N^ decreasing in diversity.* Higher entropy $H(\rho)$ of the fitness distribution implies broader coverage of the capability space for any given N . The CES diversity premium (Lemma 1) increases with the number of effectively distinct specialization types. Higher diversity means fewer agents are needed to achieve full task coverage, reducing N^* . \square

Corollary 1 (Centralized Market Share Decline). *For $N > N^*$, the centralized provider’s market share is strictly decreasing in N . In the Bianconi-Barabási framework, the centralized “condensate” fraction declines continuously as the fitness distribution broadens, transitioning from the BEC phase (macroscopic condensate) to the FGR phase (distributed traffic).*

Remark 3 (Self-Consistency Across Fields). *The critical condition $R_0^{\text{mesh}} = 1$ in Theorem 1 is a universal transcritical bifurcation. Table 1 shows the correspondence across fields. All produce the same self-consistency equation $m = f(m; \lambda)$ with the identical phase structure.*

Table 1: Universal self-consistency across fields.

| Field | Order parameter m | Control parameter λ | Critical condition | Source |
|-------------------|--|---------------------------------------|---|--------------------------|
| Percolation | Giant component S_∞ | Mean degree $\langle k \rangle$ | $\langle k \rangle = 1$ | Erdős-Rényi (1960) |
| Epidemiology | Infected fraction | R_0 | $R_0 = 1$ | Kermack-McKendrick |
| Ising/Potts | Magnetization | $\beta_T J q$ | $\beta_T J q = 1$ | Ising (1925); Potts (19) |
| Ecology | Productivity | Species richness S | Minimum S for resilience | Loreau-Hector (2001) |
| Network econ. | Market participation | Transaction benefit | Critical liquidity | Katz-Shapiro (1985) |
| This paper | Mesh fraction S_∞ | R_0^{mesh} | $R_0^{\text{mesh}} = 1$ | — |

7. Post-Crossing Dynamics

The path from $x(t) = 0$ to mesh dominance is not monotonic. Three phases, distinguished by the value of R_0^{mesh} and the state of the specialization structure, characterize the transition.

7.1 Phase 1: Nucleation ($R_0^{\text{mesh}} \approx 1$)

Immediately after crossing, R_0^{mesh} is only marginally above unity. The giant component is small ($S_\infty^* \approx 0$ for R_0^{mesh} near 1, since $S_\infty^* \sim 2(R_0^{\text{mesh}} - 1)/R_0^{\text{mesh}^2}$ to leading order). Growth is slow and stochastic. Small specialist clusters form around high-fitness agents—the enthusiast-tier hardware users running quantized models—but the clusters are fragile. Exogenous shocks (model-release events, API pricing changes, hardware supply disruptions) can temporarily push R_0^{mesh} below unity, collapsing nascent clusters.

The mesh first achieves capability dominance on the *long tail* of niche queries that centralized systems underserve. Centralized providers optimize for the highest-volume query types (general chat, code generation, summarization). Specialized queries—domain-specific technical reasoning, low-resource language translation, real-time edge processing for robotics—are underserved because the revenue per query does not justify dedicated model fine-tuning. The mesh’s heterogeneous agents, each fine-tuned for a niche, collectively cover the long tail.

This is the Christensen (1997) pattern: disruption begins in markets the incumbent rationally ignores.

7.2 Phase 2: Rapid Growth ($R_0^{\text{mesh}} \gg 1$)

As additional device types become inference-capable (driven by the continuing packaging learning curve) and coordination infrastructure matures (κ declines), R_0^{mesh} accelerates well above unity. Network effects dominate. Each new specialist joining the mesh increases C_{eff} superlinearly (by the CES diversity premium) and increases the Fiedler eigenvalue $\lambda_2(L)$ (by adding connectivity), which accelerates knowledge diffusion to subsequent entrants.

In this phase, the Potts crystallization occurs: the division of labor among mesh agents transitions from fragmented proto-specialization to a structured, self-reinforcing configuration. By Proposition 2, this transition is first-order for $q > 2$ specialization types. The crystallization is observable as a sudden increase in the concentration of agent capabilities around distinct specialization types, accompanied by the emergence of routing hub agents that handle disproportionate query traffic.

Centralized providers lose market share in progressively more mainstream query types, beginning with the long-tail niches of Phase 1 and expanding to higher-volume categories as mesh coverage broadens.

7.3 Phase 3: Maturity

Growth saturates as the mesh’s J task types are fully covered. The CES aggregate C_{eff} approaches its maximum for the given device population. Competition shifts from mesh

growth to mesh composition: which specialists are included, the quality of their fine-tuning, and the efficiency of the routing layer.

Centralized providers retain structural advantage in two domains that the mesh cannot replicate:

- (i) *Frontier model training*: As established in Smirl (2026a, Section 4), training requires tightly synchronized GPU clusters at scales incompatible with distributed architecture. The mesh depends on centralized training for the base models it fine-tunes.
- (ii) *Capabilities requiring single-device scale beyond any edge device*: Tasks requiring the full activation of 1T+ parameter dense models in a single forward pass remain centralized. This is the inference analog of the training constraint, but applies to a shrinking fraction of queries as MoE architectures reduce the active parameter requirement.

The mature equilibrium is coexistence: centralized providers dominate training and frontier-capability inference; the mesh dominates the long tail, latency-sensitive applications, and the broad middle of the query distribution where specialized, fine-tuned models outperform general-purpose frontier models.

8. The Settlement Layer

8.1 Derivation from Routing Incentives

The mesh requires agents to route queries to specialists and compensate them. Consider agent A receiving a query for which agent B is the best specialist. The routing decision requires:

- (i) *Discovery*: A must identify B as the appropriate specialist. In the mesh, this is accomplished through the degree distribution: high-quality specialists accumulate connections (the Bianconi-Barabási preferential attachment), making them discoverable through short random walks on the network.
- (ii) *Incentive compatibility*: B must be compensated for serving A ’s query. Without compensation, B has no incentive to allocate compute to external queries rather than local tasks. The compensation must exceed B ’s opportunity cost.
- (iii) *Settlement*: The compensation from A to B must be transferred. At the scale the mesh operates—millions of micro-transactions per second between arbitrary pairs of agents, with millisecond latency requirements—this requires a settlement mechanism

that is (a) peer-to-peer (no centralized clearinghouse, which would reintroduce the centralization the mesh eliminates), (b) programmable (routing decisions and payments must execute atomically), and (c) low-latency.

Proposition 6 (Settlement Layer Necessity). *Any mesh equilibrium with $N > N^*$ and $C_{mesh} > C_{cent}$ requires a settlement layer capable of processing $O(N \cdot \langle k \rangle)$ transactions per second at $O(1)$ ms latency between arbitrary node pairs. No existing payment system satisfies these requirements simultaneously.*

The proof is immediate from the bandwidth scaling result (equation 13): the number of routed queries scales as $N \cdot \langle k \rangle$, and each routed query requires compensation.

8.2 The Hayek Insight

The price system in the mesh plays exactly the role Hayek (1945) described for the market price system: it aggregates dispersed information into sufficient statistics for decentralized decision-making. The bid-ask spread between agents encodes:

- *Current demand* for each query type (high bids for underserved specializations signal entry opportunities).
- *Current supply* of each specialization (narrow spreads indicate competitive specialist markets).
- *Optimal routing* (the lowest-ask specialist for a given query type is the efficient allocation).

No central coordinator computes these allocations. Agents respond to prices, and the price system achieves efficient routing as a fixed point of bilateral negotiations. This is the Hayek mechanism operating at machine speed.

8.3 Connection to Monetary Infrastructure

The settlement layer requirements—programmable, peer-to-peer, high-throughput, low-latency—describe the functional specification of a programmable monetary system. The detailed analysis of how existing and emerging monetary infrastructure maps to these requirements is the subject of Smirl (2026b, *The Monetary Productivity Gap*, forthcoming). This paper notes the connection without developing it: the mesh equilibrium’s viability depends on the existence of adequate settlement infrastructure, and the mesh’s growth may be constrained by the settlement layer before it is constrained by device capability or network connectivity.

Prediction 6 (Section 10) formalizes this as a testable claim. The mesh level is one of N_{eff} endogenously determined hierarchy levels (Smirl 2026d, Theorem 4.3): the effective hierarchy depth is set by the technology’s timescale distribution, not structurally fixed at four.

9. Frameworks Considered and Rejected

Several candidate frameworks were evaluated for the formal model and rejected for specific technical reasons. Documenting these decisions clarifies the modeling choices.

Mean Field Games (Lasry-Lions 2007). MFG assumes a continuum of exchangeable (identical) agents whose individual optimization depends on the population distribution. The mesh’s agents are heterogeneous specialists—heterogeneity is the source of the CES diversity premium that drives Theorem 1. Replacing heterogeneous agents with a continuum of identical agents eliminates the mechanism. The supermodular game framework (Topkis 1998; Milgrom-Roberts 1990) handles heterogeneity naturally through lattice-theoretic monotone comparative statics.

Spin glasses (Edwards-Anderson 1975; Sherrington-Kirkpatrick 1975). Spin glass models require frustrated interactions—a mix of positive and negative couplings. In the mesh, all interactions are positive: each agent benefits from others joining the network (network effect) and from others specializing in complementary tasks (CES complementarity). There is no frustration. The appropriate statistical mechanics model is the random-field Ising/Potts model (positive couplings, heterogeneous external fields representing device-specific capabilities), not a spin glass.

Ecological niche models (Tilman 1982; Loreau-Hector 2001). The conceptual analogy is precise: diverse specialist communities outperform monocultures, exactly as in the mesh. The formal ecological models, however, are calibrated to plant biomass dynamics with resource-competition mechanics (light, nutrients, water) that do not transfer to inference economics. The CES aggregation function captures the identical qualitative result—diversity premium from imperfect substitutability—while being native to the economics literature and directly comparable to standard production function analysis.

Immune system / clonal selection (Burnet 1959). The adaptive immune system provides a vivid metaphor for the mesh: a diverse repertoire of specialized agents (antibodies/T-cells) that collectively cover a vast antigen space through local adaptation. However, the formal ODE models of clonal expansion, affinity maturation, and immune memory are calibrated to lymphocyte population dynamics with proliferation rates, apoptosis, and antigen-presentation mechanisms that have no meaningful economic analog. The metaphor is instructive; the math does not transfer.

10. Falsifiable Predictions

The model generates six predictions with timing and failure conditions, extending the predecessor paper’s predictions into the post-crossing regime. If these fail, the mesh equilibrium theory is wrong.

Prediction 1: First-Order Crystallization, Not Gradual Adoption. Mesh formation exhibits a discontinuous jump in adoption metrics—distributed inference share, number of active mesh participants, routing volume—rather than smooth logistic growth. Specifically, the transition from $<5\%$ to $>25\%$ distributed inference share occurs within 18 months, a pace inconsistent with Bass diffusion ($p + q < 0.5$) but consistent with first-order Potts crystallization. *Timing:* 2030–2033, following the hardware crossing and $R_0 > 1$ events predicted in Smirl (2026a). *Evidence against:* distributed inference share growing smoothly at <5 percentage points per year through 2035.

Prediction 2: Specialization Precedes Generalization. Early mesh agents are narrow specialists (fine-tuned for specific domains: legal reasoning, medical coding, code review in specific languages). General-purpose mesh capability—matching centralized providers on breadth—emerges only after the specialization structure crystallizes and routing infrastructure connects specialists. *Evidence against:* early mesh participants predominantly running general-purpose base models without fine-tuning.

Prediction 3: Long-Tail Niche Dominance First. The mesh achieves capability dominance first on long-tail queries (rare languages, domain-specific technical problems, real-time edge processing) before competing on mainstream tasks (general chat, standard code generation, summarization). Market share data should show distributed inference share exceeding 50% for long-tail categories while remaining below 20% for high-volume mainstream categories. *Timing:* within 2 years of $R_0 > 1$ crossing. *Evidence against:* mesh competing first on mainstream query types.

Prediction 4: Endogenous Hub Emergence. The mesh’s degree distribution becomes fat-tailed ($\gamma \leq 3$) within 3 years of crystallization, with $<1\%$ of nodes handling $>30\%$ of routing traffic. These hub agents emerge from the Bianconi-Barabási preferential attachment dynamics, not from central design. *Evidence against:* the degree distribution remaining thin-tailed (exponential or Gaussian) through 2036.

Prediction 5: Nonlinear Knowledge Acceleration. The rate of capability improvement across the mesh (measured by benchmark scores, task coverage, or response quality) accelerates nonlinearly once the degree distribution becomes fat-tailed, consistent with the vanishing epidemic threshold (Proposition 5). Specifically, the time between successive capability doublings decreases, rather than remaining constant as in standard learning curves.

Evidence against: capability improvement following a constant exponential rate through 2036.

Prediction 6: Settlement Layer as Binding Constraint. The settlement layer (routing compensation and micro-transaction infrastructure) becomes the binding constraint on mesh growth before device capability, network connectivity, or model quality bind. Observable as: mesh growth stalling despite available device capacity, with growth resuming when settlement infrastructure improves. *Timing:* 2031–2034. *Evidence against:* mesh growth constrained by device capability or bandwidth through 2035, with settlement infrastructure remaining unused capacity.

11. Conclusion

This paper has formalized what happens after the crossing point identified in Smirl (2026a). The answer is not isolated devices running local inference. It is a self-organizing mesh of heterogeneous specialized agents whose collective capability exceeds centralized provision once the mesh reaches critical mass.

The mesh equilibrium emerges from three composable mechanisms. Percolation theory establishes that a connected mesh forms when $R_0^{\text{mesh}} > 1$. CES aggregation with $\rho < 1$ ensures that heterogeneous specialists collectively outperform homogeneous centralized provision—the diversity premium. Laplacian knowledge diffusion with a vanishing epidemic threshold on scale-free topologies ensures that capability improvements propagate self-sustainingly across the mesh.

Three results distinguish this analysis from standard network economics. First, the Fortuin-Kasteleyn unification reveals that connectivity and specialization are the same mathematical object at different Potts parameter values, and predicts first-order (discontinuous) crystallization for $q > 2$ specialization types. Second, the Bianconi-Barabási framework connects the predecessor paper’s learning curve to an inverse Bose-Einstein condensation: the packaging learning curve broadens the fitness distribution, dissolving the centralized condensate. Third, the model endogenously derives the need for a programmable settlement layer from the routing incentive structure, connecting to the monetary infrastructure analysis of Smirl (2026b).

The central theorem establishes that the mesh equilibrium exists, is unique, is stable, and dominates centralized provision above a finite critical mass N^* that decreases with population diversity. The post-crossing dynamics proceed through three phases: slow nucleation on long-tail niches, rapid crystallization with first-order specialization, and maturity with centralized coexistence on training and frontier inference.

What the model predicts unambiguously is that the distributed ecosystem does not remain a collection of isolated devices. The same economic forces that drive specialization and trade in human economies—comparative advantage, network effects, complementary capabilities—drive the formation of a mesh whose collective intelligence exceeds what any centralized system can provide. The organizational form is not designed; it is the equilibrium.

The mesh equilibrium results—critical mass N^* , diversity premium, settlement layer endogeneity—are structural features of CES aggregation with $\rho < 1$ and do not depend on the specific number of hierarchy levels. The companion paper (Smirl 2026d) proves that the four-level structure assumed across the paper sequence ($N_{\text{eff}} = 4$) is endogenously determined by the technology’s timescale distribution and may evolve as the AI transition matures.

A. Renormalization Group Universality

The modeling choices in this paper—supermodular game theory for equilibrium existence, mean-field percolation for the giant component, CES production functions for capability aggregation—may appear to be convenient approximations rather than rigorous characterizations. The renormalization group (RG) provides a justification.

Remark 4 (Mean-Field Exactness Above the Upper Critical Dimension). *For systems on networks with spectral dimension $d_s > 4$, mean-field theory is exact—not an approximation, but a rigorous result (see e.g. Dorogovtsev et al. 2008 for a review). Real-world networks, including the internet, social networks, and infrastructure graphs, generically have $d_s > 4$ due to the small-world property (high clustering with short path lengths).*

This means that the mean-field percolation result (Proposition 1), the mean-field Potts transition (Proposition 2), and the Katz-Shapiro network goods model that underlies the supermodular game are not approximations for the mesh. They are exact characterizations of the phase structure. Corrections to mean-field scaling are suppressed by powers of $1/(d_s - 4)$ and are negligible.

The practical implication is that the paper’s use of mean-field models requires no apology. The RG result licenses the framework.

References

- [1] Bianconi, G., & Barabási, A.-L. (2001). Bose-Einstein condensation in complex networks. *Physical Review Letters*, 86(24), 5632–5635.
- [2] Baxter, R. J. (1982). *Exactly Solved Models in Statistical Mechanics*. Academic Press.
- [3] Becker, G. S., & Murphy, K. M. (1992). The division of labor, coordination costs, and knowledge. *Quarterly Journal of Economics*, 107(4), 1137–1160.
- [4] Bonabeau, E., Theraulaz, G., & Deneubourg, J.-L. (1998). Fixed response thresholds and the regulation of division of labor in insect societies. *Bulletin of Mathematical Biology*, 60(4), 753–807.
- [5] Bresnahan, T. F., & Trajtenberg, M. (1995). General purpose technologies: Engines of growth? *Journal of Econometrics*, 65(1), 83–108.
- [6] Burnet, F. M. (1959). *The Clonal Selection Theory of Acquired Immunity*. Cambridge University Press.
- [7] Christensen, C. M. (1997). *The Innovator's Dilemma*. Harvard Business School Press.
- [8] Dorogovtsev, S. N., Goltsev, A. V., & Mendes, J. F. F. (2008). Critical phenomena in complex networks. *Reviews of Modern Physics*, 80(4), 1275–1335.
- [9] Edwards, S. F., & Anderson, P. W. (1975). Theory of spin glasses. *Journal of Physics F: Metal Physics*, 5(5), 965–974.
- [10] Erdős, P., & Rényi, A. (1960). On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, 5, 17–61.
- [11] Fortuin, C. M., & Kasteleyn, P. W. (1972). On the random-cluster model: I. Introduction and relation to other models. *Physica*, 57(4), 536–564.
- [12] Hayek, F. A. (1945). The use of knowledge in society. *American Economic Review*, 35(4), 519–530.
- [13] Katz, M. L., & Shapiro, C. (1985). Network externalities, competition, and compatibility. *American Economic Review*, 75(3), 424–440.
- [14] Lasry, J.-M., & Lions, P.-L. (2007). Mean field games. *Japanese Journal of Mathematics*, 2(1), 229–260.

- [15] Loreau, M., & Hector, A. (2001). Partitioning selection and complementarity in biodiversity experiments. *Nature*, 412, 72–76.
- [16] Milgrom, P., & Roberts, J. (1990). Rationalizability, learning, and equilibrium in games with strategic complementarities. *Econometrica*, 58(6), 1255–1277.
- [17] Pastor-Satorras, R., & Vespignani, A. (2001). Epidemic spreading in scale-free networks. *Physical Review Letters*, 86(14), 3200–3203.
- [18] Potts, R. B. (1952). Some generalized order-disorder transformations. *Mathematical Proceedings of the Cambridge Philosophical Society*, 48(1), 106–109.
- [19] Sherrington, D., & Kirkpatrick, S. (1975). Solvable model of a spin-glass. *Physical Review Letters*, 35(26), 1792–1796.
- [20] Smirl, J. (2026a). Endogenous decentralization: How concentrated capital investment finances the learning curves that enable distributed alternatives. Working Paper.
- [21] Smirl, C. (2026b). The monetary productivity gap. Working paper, Tufts University. Forthcoming.
- [22] Smirl, J. (2026d). Complementary heterogeneity in hierarchical economies. Working Paper.
- [23] Tarski, A. (1955). A lattice-theoretical fixpoint theorem and its applications. *Pacific Journal of Mathematics*, 5(2), 285–309.
- [24] Tilman, D. (1982). *Resource Competition and Community Structure*. Princeton University Press.
- [25] Topkis, D. M. (1998). *Supermodularity and Complementarity*. Princeton University Press.
- [26] Wu, F. Y. (1982). The Potts model. *Reviews of Modern Physics*, 54(1), 235–268.