Sufficiency of Local Exchange: Continuation and Admissibility in Distributed Computation

Flyxion
Independent Researcher

November 2026

Abstract

Agreement between locally evolving parts at the moment they exchange information is easily mistaken for the ability to continue together. The essay makes the difference precise as a relation between an exchange and a target. An exchange is sufficient for a target when each part's required output can be decoded from its own state, the messages it received, and its random inputs, and it is insufficient when two joint states give some part identical inputs while requiring different outputs. The same exchange can therefore be sufficient for one target and insufficient for another. The relation is applied to four settings. In coordinate formation on an amorphous computer, local coordinate systems fitted pairwise on their overlaps can match on every overlap and still fail to assemble, because errors accumulate around loops, and a defect that survives better observations must be traced either to the fitted model or to the space. In local gradient steps with averaging, a quadratic example separates two targets: a gradient correction restores the pooled fixed point for every number of local steps, while exact agreement with the pooled trajectory is guaranteed only through two. In tiled denoising, agreement on an overlap is guaranteed to persist on a core whose width is the overlap width minus twice the distance information travels between exchanges, and a shared coarse representation extends the guarantee only if the update factors through it. In growing-point construction, a failure erases both built structure and the state of the processes building it, and repair is possible exactly when the survivors determine an acceptable completion. The essay closes by asking what the smallest sufficient message or record costs for a fixed target, tolerance, and interval, and holds that the cost of an exchange is measured by what it carries and not by how much agreement it displays when it is sent.

Introduction

Amorphous computing begins from a deliberately unpromising substrate. A very large number of small processors are scattered irregularly, interconnected in unknown and possibly changing ways, and individually unreliable. The research program set out to obtain organized behavior from the cooperation of such parts without assuming any precision interconnect or precision geometrical arrangement [1, 5, 6]. The program is usually presented as a collection of techniques: wave propagation, coordinate formation, club formation, activation and inhibition, growing points. Read together, the techniques pose a single question that is more general than any one of them. Once the parts of a system evolve on their own, what must they tell one another so that what each does next still belongs to the same computation?

The question recurs well outside amorphous computing. Workers that take several optimization steps on different local objectives and then average their parameters face it [13, 17]. Image tiles that are denoised separately between exchanges face it [2]. A construction that is interrupted halfway and must be resumed by the surviving parts faces it as well. In each case agreement at one moment is easy to check and is not the same thing as the ability to continue together. The essay uses the term global admissibility for the consistency or correctness of the joint outcome, and examines how it can fail when agreement holds.

Sufficiency of an exchange

The question can be given a definite form. Index the parts by i, let x_i be the local state of part i and r_i its random inputs, and write σ for the joint state. An exchange delivers to each part a message $m_i(\sigma)$. Each part then applies a local rule, whatever algorithm it actually runs, and produces an output $y_i = A_i(x_i, m_i, r_i)$, and the outcome of the joint computation combines the outputs. The output of a given rule is fixed by its inputs, so the rule's behavior does not by itself say whether the computation succeeds. Success is measured against a target F, a map from joint states to a desired outcome or to a class of acceptable outcomes, such as the pooled endpoint, a consistent global chart, or the set of acceptable completed structures. When the target depends on an interval of s local steps between exchanges it is written F_s, and when it is a consistency requirement with no interval, as in assembling coordinates, it carries no subscript. Where the target prescribes what part i must output, write that required output as $\rho_i(\sigma)$.

Definition 1. An exchange is *sufficient* for the target F to tolerance ε if, for every part, there is a decoder D_i with $\|D_i(x_i, m_i, r_i) - \rho_i(\sigma)\| \le \varepsilon$ for every joint state σ in scope. It is *insufficient* if some part receives the same (x_i, m_i, r_i) in two joint states σ, σ' whose required outputs differ by more than 2ε. The factor two is the triangle inequality: a decoder within ε of both required outputs would put them within 2ε of each other. In the exact case $\varepsilon = 0$ the two required outputs must simply differ.

Insufficiency is witnessed by a fiber: the set of joint states that look identical to a part contains states that call for different behavior from it. The test is independent of the particular rule the part runs. A weaker statement is relative to a class of rules, for example K local

gradient steps, and says that no member of the class achieves the target. Where a target constrains only a combination of outputs, such as an average, required outputs are not unique and the argument is made directly on the combination, as in Section 4.

Sufficiency is therefore a relation between an exchange and a target and not a property of the exchange alone. The central observation of the essay is that an exchange can be sufficient for one target and insufficient for another, and that confusing the two is the usual way continuation fails. Section 1 describes how countdown waves give processors approximate coordinates. Section 2 treats the reconciliation of local coordinate systems as an assembly problem with a consistency target. Section 3 sets reaction–diffusion beside generative diffusion and asks where the information that organizes the result resides; the contrast it draws between a bounded and an unbounded dependence radius per update is what Section 5 uses. Section 4 works a minimal example in which one exchange is sufficient for one target and not another. Section 5 asks what a denoising tile needs from its neighbors. Section 6 turns to growing points and to what a construction must retain in order to be resumed. Section 7 collects the results and asks what the smallest sufficient exchange costs.

1 Coordinate Formation

Consider processors scattered over a surface, each able to exchange messages with the processors within a fixed radius and none holding any coordinate. The simplest structure such a collection can build is a gradient. One processor, the source, emits a message carrying an integer n. Each processor that receives a message with value k keeps the largest value it has heard and, if $k > 0$, relays $k - 1$ to its neighbors. After the wave has passed, the value held at a processor x is $g_s(x) = n - h_s(x)$, where $h_s(x)$ is the number of communication hops from the source s to x. The processors have constructed a scalar field over themselves using nothing but local messages and a counter [6].

A single wave gives one number per processor, which locates the processor on a circle-like level set around the source. Two waves emitted from different sources give a pair $\big(g_{s_1}(x), g_{s_2}(x)\big)$, and where the level sets of the two fields cross transversally this pair determines position up to the accuracy of the hop count. The slides describe the result as crude local coordinate systems obtained by intersecting countdown waves radiating from several loci [6], and Nagpal's work on coordinate formation develops the idea of forming several gradients and triangulating a coordinate system from them [14]. Two features of the construction matter for what follows. The coordinates are only approximate, because hop count overestimates Euclidean distance in sparse deployments and fluctuates with local density. And they are relative to the chosen sources: a region that uses one pair of sources obtains a coordinate system that is, in general, unrelated to the one obtained by a region that uses another pair.

The approximation can be improved by relaxation. If boundary values are fixed and every interior processor repeatedly replaces its value with the mean of its neighbors' values, the values converge to a discrete harmonic function, the solution of Laplace's equation with the given boundary condition. The slides use this to interpolate coordinates from boundary data

by successive averaging on an amorphous computer [6]. The iteration is the discrete heat equation run to its steady state, and the number of sweeps needed grows with the square of the diameter of the region measured in hops, a standard property of Jacobi iteration on a graph. Refinement is therefore affordable on a region of modest extent and expensive on a large one. That asymmetry is the practical reason for building many local coordinate systems and joining them, rather than one large one.

2 Compatibility on Overlaps

Suppose the surface is covered by regions $U_1, \dots, U_m$, and region U_i has built a coordinate system $\varphi_i : U_i \to \mathbb{R}^2$. Two regions that share processors can compare the coordinates those processors received in each system. Write $U_{ij} = U_i \cap U_j$ and

$$T_{ij} = \varphi_i \circ \varphi_j^{-1},$$

which takes φ_j-coordinates to φ_i-coordinates on $\varphi_j(U_{ij})$. The slide presenting this step states that a global picture can be constructed by working out the mappings at the overlaps among the local coordinate systems [6]. The reconciliation is an assembly problem, and its target is a consistency condition with no interval of local steps attached: the local systems should be pieces of one description. The question for the exchange is what the regions must send one another for that target to be achievable.

Two kinds of transition have to be kept apart. Exact chart transitions, those of a single atlas, satisfy the cocycle condition $T_{ij}T_{jk} = T_{ik}$ wherever the composition is defined, and this is an identity between compositions of functions. Around a triple overlap the composite $T_{13}T_{32}T_{21}$ is therefore exactly the identity on $\varphi_1(U_{123})$, whatever the geometry of the underlying surface. Curvature does not violate it. Curvature appears instead in the form of the exact transitions, which for charts of a curved surface are not rigid motions, and in the holonomy of parallel transport around a loop. What the processors actually obtain are fitted transformations $\hat{T}_{ij}$ drawn from a chosen family G, for instance the rigid motions, selected to match the paired coordinates of shared processors. These are group elements, defined as maps of the whole plane, and they need not coincide with the exact transitions anywhere.

Two consistency tests follow, and they test different things.

Definition 2. For three regions with a common processor, the *domain defect* is

$$\delta_{123}^{\mathrm{dom}} = \sup_{q \in \varphi_1(U_{123})} \|\hat{T}_{13}\hat{T}_{32}\hat{T}_{21}(q) - q\|,$$

which measures consistency of the fits on an actual common domain. For any loop of regions, with or without a common processor, the *holonomy defect* is the distance of the composite $H = \hat{T}_{13}\hat{T}_{32}\hat{T}_{21} \in G$ from the identity, for example the largest displacement $\|H(p) - p\|$ over a region of interest. It measures consistency of the group-valued estimates extended beyond the overlaps that determined them.

The domain defect asks whether the fits agree where observations exist. The holonomy defect remains defined around a ring of regions in which no three have a common processor, and there it is the only test available. An *admissible assembly* is a single chart Φ together with transformations $g_i \in G$ such that $\hat{T}_{ij} = g_i g_j^{-1}$ for every overlapping pair. For G-valued transition data this requires the holonomy defect to vanish on every loop. When the nerve of the cover is simply connected, vanishing on triple overlaps suffices: choose a spanning tree of regions, define each g_i by composing the $\hat{T}_{ij}$ along the tree path from a base region, and note that the cocycle condition makes the result independent of the path. Where the nerve has loops that do not bound, as for a ring of regions around a hole, those loops have to be checked separately.

Proposition 1. *Let three charts differ by planar rotations about a common point, and let the fitted rotation angles be $\hat{\alpha}_{ij} = \alpha^*_{ij} + e_{ij}$, where the exact angles satisfy $\alpha^*_{12} + \alpha^*_{23} + \alpha^*_{31} \equiv 0 \pmod{2\pi}$. If $|e_{ij}| \le \varepsilon$ for each pair, the composite around the loop is a rotation by $e_{12}+e_{23}+e_{31}$, whose magnitude is bounded by 3ε, with equality when the errors share a sign, and the extended composite displaces a location at distance ρ from the center by $2\rho\,|\sin\frac{1}{2}(e_{12}+e_{23}+e_{31})|$. Around a ring of n regions the corresponding bound is $n\varepsilon$.*

Proof. Rotations about a common point commute and their angles add. The exact angles contribute a multiple of 2π, so the composite angle is the sum of the errors. The displacement of a point at radius ρ under rotation by θ is the chord $2\rho|\sin(\theta/2)|$. □

The proposition concerns a simplified case, and its content is general, though the clean bound is not: once translations are included, rigid motions do not commute, the composite depends on the order and on the positions of the centers, and the $n\varepsilon$ bound holds only to first order in the errors. A tolerance on each overlap does not give a tolerance on the loop, because the errors add. A procedure that fits each overlap separately has no mechanism for noticing this, since every individual fit is good. Pairwise fits are a sufficient exchange for a pairwise target, namely that each overlap be matched, and an insufficient one for the assembly target, because two joint configurations can give a region the same pairwise fits and require different corrections. Nothing local reveals the difference, and the loop is the smallest object on which it appears.

When the defect persists as observations improve, its cause can be read from which of the two tests fails. If the residuals of the fits on the overlaps themselves do not shrink, the rigid-motion family is inadequate for the correspondences, which is what one expects when the underlying surface is curved or the coordinates are distorted, because the exact transitions then lie outside G. The remedy is a richer family or an atlas that records its transitions and does not force them into G. If the overlap fits are exact and the holonomy defect around a ring is nonzero, the data describe a flat space with nontrivial holonomy, as a strip closed into a cylinder does, and no single chart in G exists. Demanding one would be the wrong target, and the defect is information about the space. If both defects shrink as observations improve, the cause is estimation error accumulating as in the proposition. For that case the constructive remedy is to estimate all transitions jointly, as in rotation averaging, which chooses the g_i that

minimize the total discrepancy over all overlaps rather than fitting overlaps one at a time [9]. The remedy requires the regions to send one another more than their overlaps: the estimated transition maps and the loops they close.

3 Reaction–Diffusion and Denoising

Two quite different procedures produce spatial structure from a random initial field by repeated local updates. The first is a reaction–diffusion system. The slides display the Gray–Scott model started from random initial values, and organize the discussion around the general form

$$\frac{\partial \mathbf{u}}{\partial t} = \mathbf{D}\,\nabla^2 \mathbf{u} + \mathbf{F}(\mathbf{u}),$$

the setting of Turing's analysis of morphogenesis [18, 6]. For the Gray–Scott system the reaction terms are

$$\frac{\partial u}{\partial t} = D_u \nabla^2 u - u v^2 + F(1-u), \qquad \frac{\partial v}{\partial t} = D_v \nabla^2 v + u v^2 - (F+k)v,$$

and for suitable parameters a uniform state loses stability to spatial patterns [8]. On a graph of processors the update is local by construction: each processor replaces its state with a combination of its own state and the differences to its neighbors, plus a reaction term that depends on its own state alone. After n updates a processor's value can depend on processors at most n hops away, and no further.

The second procedure is a generative diffusion model. A forward process corrupts a sample x_0 with Gaussian noise, $x_t = \sqrt{\bar{\alpha}_t}\, x_0 + \sqrt{1-\bar{\alpha}_t}\, \varepsilon$, and a learned reverse process turns noise into a sample through updates of the form

$$x_{t-1} = \frac{1}{\sqrt{\alpha_t}}\Big(x_t - \frac{\beta_t}{\sqrt{1-\bar{\alpha}_t}}\, \varepsilon_\theta(x_t, t)\Big) + \sigma_t z,$$

where z is fresh noise [15, 10]. In the continuous-time view the same process discretizes a reverse stochastic differential equation whose drift contains the score $\nabla \log p_t$ [16]. Both procedures start from noise and develop structure over successive updates, and in both the structure is not present in the initial field.

The comparison becomes useful when one asks where the information that organizes the result resides. In the reaction–diffusion system the update rule is the same at every site and contains no description of the final pattern. It fixes the scale at which patterns form, since the spatial wavelength is determined by the diffusion constants and reaction parameters, and it leaves the positions of the individual features to the initial perturbation. Information about where things go comes from the evolving state. Information about how things are shaped comes from the rule. In the denoising model most of the organizing information lies in the network weights, which encode a learned function of the whole current field. The function ε_θ is not local. A convolutional network has a receptive field of fixed finite radius, and global attention layers extend the dependence across the full image, whereas windowed attention does not. A single

reverse update at one location therefore may depend on the state at locations far outside any neighborhood of it, and its dependence radius can be as large as the image.

This difference between a bounded dependence radius per update and an unbounded one is the structural fact on which the later sections turn. Relaxation toward a harmonic function, reaction–diffusion, and denoising are all iterated local updates, though the first runs to a fixed point and the others evolve a state. The first two keep their dependence local, so that the spread of influence is limited by the number of updates. The third does not, and a distributed implementation of it has to decide how much of the nonlocal dependence to carry through communication.

4 Local Drift and Averaging

The mathematical core of the comparison concerns what is lost when parts evolve separately and are then merged by averaging. The setting is the familiar one of workers who start from the same parameters, take several steps on different local objectives, and average their endpoints [13, 17]. The layered schedule of such averaging is treated in the essay Hierarchical Synchronization [7], and the discussion here isolates the smallest instance of the discrepancy.

Take two workers with scalar quadratic objectives $f_i(w) = \frac{1}{2}a_i(w - b_i)^2$, $i = 1, 2$, and the pooled objective $f = \frac{1}{2}(f_1 + f_2)$. All workers start at w_0 and use step size η. One gradient step on f_i maps $w - b_i$ to $(1 - \eta a_i)(w - b_i)$. After K local steps worker i is at

$$w_i^{(K)} = b_i + q_i^K(w_0 - b_i), \qquad q_i = 1 - \eta a_i,$$

and the averaged endpoint is $\hat{w}^{(K)} = \frac{1}{2}(w_1^{(K)} + w_2^{(K)})$. The pooled gradient is $\bar{a}\,(w - c)$ with $\bar{a} = \frac{1}{2}(a_1 + a_2)$ and $c = (a_1b_1 + a_2b_2)/(a_1 + a_2)$, so K pooled steps lead to

$$w_{\text{pool}}^{(K)} = c + (1 - \eta\bar{a})^K(w_0 - c).$$

For $K = 1$ the two agree exactly. The averaged endpoint is $w_0 - \eta \cdot \frac{1}{2}\sum_i a_i(w_0 - b_i)$, which is one pooled gradient step, because the average of the gradients at a common point is the pooled gradient. This is the sense in which a single communication per gradient step is lossless. For $K > 1$ each worker evaluates its gradient at its own displaced point, and the average of gradients evaluated at different points is no longer the gradient of the pooled objective at any point the procedure visits.

A concrete case exposes the discrepancy. Let $a_1 = 1$, $b_1 = 0$, $a_2 = 2$, $b_2 = 1$, $\eta = \frac{1}{2}$, and $w_0 = 0$. Then $q_1 = \frac{1}{2}$ and $q_2 = 0$. Worker 1 starts at its own minimizer and does not move. Worker 2 reaches its minimizer in one step and stays there. The averaged endpoint is $\frac{1}{2}$ for both $K = 1$ and $K = 2$. The pooled objective has $\bar{a} = \frac{3}{2}$ and $c = \frac{2}{3}$, and

$$w_{\text{pool}}^{(1)} = \tfrac{1}{2}, \qquad w_{\text{pool}}^{(2)} = \tfrac{2}{3} + \tfrac{1}{16}(-\tfrac{2}{3}) = \tfrac{5}{8}.$$

The two procedures agree at $K = 1$ and differ by $\frac{1}{8}$ at $K = 2$. The difference does not vanish over rounds. If averaging is repeated, the fixed point w^* of the round map satisfies

$\sum_i (1 - q_i^K)(w^* - b_i) = 0$, so

$$w^* = \frac{\sum_i (1 - q_i^K)\, b_i}{\sum_i (1 - q_i^K)}.$$

For $K = 1$ the weights $1 - q_i = \eta a_i$ reproduce the pooled minimizer c. For $K = 2$ the weights are $\frac{3}{4}$ and 1, and $w^* = \frac{4}{7}$, not $c = \frac{2}{3}$. Increasing the number of local steps shifts the weights from curvature toward saturation, so that a worker whose objective is already nearly minimized counts for less than it should. The procedure converges, but to a point that is not the minimizer of the objective it was meant to serve. This persistent offset is often called client drift in the federated setting [12, 11].

A control variate changes what is communicated. Each worker corrects its local gradient by the difference between a global and a local gradient estimate, so that the local direction becomes $\nabla f_i(w) - \nabla f_i(w_0) + \nabla f(w_0)$ with all estimates taken at the common starting point of the round [11]. The message now contains the pooled gradient $g_0 = \nabla f(w_0)$ in addition to w_0. For the quadratic example the local displacement $d = w - w_0$ obeys $d \leftarrow (1 - \eta a_i)d - \eta g_0$, which gives

$$d_i^{(K)} = -\eta g_0 \sum_{j=0}^{K-1} (1 - \eta a_i)^j.$$

At a pooled minimizer $g_0 = 0$, so every worker stays where it is and the pooled minimizer is a fixed point of the round for every K. The round endpoint is a polynomial of degree $K - 1$ in a_i, and the average of a polynomial equals the polynomial of the average only when the degree is at most one. The corrected average therefore matches the pooled trajectory for $K \leq 2$. In the numerical example above $g_0 = -1$ and the corrected average at $K = 2$ is $\frac{5}{8}$, the pooled value. At $K = 3$ the two differ by exactly $-\eta^3 g_0 \operatorname{Var}(a)$, where $\operatorname{Var}(a) = \left(\frac{a_1 - a_2}{2}\right)^2$ is the variance of the curvatures.

Two targets are now in play, and the example shows exchanges that are sufficient for one and insufficient for the other. The first target is that the pooled minimizer be a fixed point of the round. The second is that the round endpoint equal the pooled endpoint after K steps. Three exchanges can be compared. The first sends w_0 alone. It is sufficient for both targets at $K = 1$. For $K \geq 2$ it fails the first target within the class of K-step gradient rules, as the offset $\frac{4}{7}$ versus $\frac{2}{3}$ shows, and it fails the second target for every local rule: at $K = 2$ the pooled endpoint is $w_0 - 2\eta g_0 + \eta^2 \bar{a} g_0$, and the product $\bar{a} g_0$ contains the cross term $a_1 a_2 (u_1 + u_2)/4$ with $u_i = w_0 - b_i$, which couples the two workers and which no sum of terms, each computed by one worker from its own objective and w_0, can produce. The second exchange sends (w_0, g_0). It is sufficient for the first target for every K and for the second target through $K = 2$, since the corrected endpoint is then a sum of terms each depending on one worker's curvature and on the shared g_0, with no product of the two workers' curvatures. For $K = 3$ it is insufficient for the second target under any local rule, because the pooled endpoint contains a term in $\bar{a}^2 g_0$ and hence a cross product $a_1 a_2 g_0$ that no term computed by one worker from its own curvature, w_0, and g_0 can supply. The third exchange sends $(w_0, g_0, \bar{a})$, from which $c = w_0 - g_0/\bar{a}$ is determined, and it is sufficient for the second target for every K, since $\bar{a}$ and c fix the pooled endpoint. At that point, however, each worker could compute the pooled trajectory without using its own objective at all. The quadratic family admits this because its pooled problem

has a two-number summary, and for general objectives no fixed-size summary of this kind is available.

The example supports the central claim in a precise form. Frequency of communication, which is K, and content of communication, which is whether workers exchange only parameters or also gradient corrections and curvature, are both parts of the algorithm. The exchange of (w_0, g_0) recovers the pooled fixed point for every K in the example, while exact agreement with the pooled trajectory is guaranteed only through two steps. Which of the two is the right target depends on what the computation is for, and the sufficiency of the message cannot be judged without naming it.

5 Tiled Denoising

The same issue appears in image generation when a large image is split into overlapping tiles and each tile is denoised on its own for several steps between exchanges. A scheme that fuses the tiles at every step by weighted least squares gives each pixel an average of the predictions of the tiles that cover it [2]. Exchanging less often reduces communication and invites the failure studied here. Take as the running example a single contour that crosses a tile boundary. Each tile sees one side of the contour inside its own window and must infer how it continues. Two tiles can hold nearly identical values on their common overlap and still move the contour in different directions at the next step, because each update depends on context that lies outside the overlap and differs between the tiles.

A guarantee can be stated under a locality assumption. Assume that the two tiles apply identical update rules with identical conditioning and identical noise samples, and that one denoising step has dependence radius R, meaning that the new value at pixel p depends on the current tile values within distance R of p and on nothing else. Identical noise is the idealized case and is essential to the argument: if the tiles draw independent noise, their stochastic terms differ at the first step and the guarantee does not apply, and what remains is an empirical question about the size of the resulting seam error.

Proposition 2. *Let x^i and x^j be the states of tiles Ω_i and Ω_j immediately after an exchange, and suppose they agree on the overlap $\Omega_i \cap \Omega_j$. After s further steps without communication, the two tiles are equal at every pixel p whose ball $B(p, sR)$ lies inside $\Omega_i \cap \Omega_j$. In an idealized overlap that is a band of width w between two tile edges, such pixels form a core of width $w - 2sR$ when $w > 2sR$.*

Proof. By induction on the number of steps. After one step the two tiles agree at every pixel whose ball of radius R lies in the overlap, since the update uses only values on which they agree, with identical rules, conditioning, and noise. After m steps they agree on pixels whose ball of radius mR lies in the overlap, because the ball of radius R around such a pixel lies in a region where they agreed after $m - 1$ steps. □

The bound is a sufficient geometric condition for a guarantee, and it is no more. Width greater than $2sR$ yields a core of positive width in the idealized band, and the proposition

asserts pointwise equality on that core and nothing about the pixels outside it. A particular update may agree on a narrower overlap, for instance when its actual dependence is shorter than the nominal radius or when the discrepancy outside the overlap is small, and the proposition does not claim disagreement there. It does show what each of the evident remedies acts on. A wider overlap enlarges the guaranteed core. More frequent exchange reduces s and so reduces the radius by which the core erodes. Neither helps when R is unbounded, which is the case for a network with global attention, since then no overlap is wide enough and the guaranteed core is empty. This is the practical form of the distinction drawn in Section 3.

A shared coarse representation addresses that case under a stronger condition. Suppose each tile receives the same coarse representation c of the whole image, held fixed over the interval, and suppose the update at a pixel factors through the pair formed by the local values within radius R and c, so that distant context affects the pixel only through c. Then the proposition applies with the same geometric condition on the local ball, together with equality of c across tiles: the coarse channel accounts for distant context, and the local dependence still needs adequate coverage by the overlap. If the update does not factor through c, the channel may still reduce disagreement in practice, but no guarantee follows from the argument. What has to be exchanged is then not more of the neighbor's pixels but a summary of everything the tile cannot see, and the quality of the summary is measured by how nearly the update factors through it.

The guarantee also fixes what an experiment can and cannot establish. A test of whether wider overlaps reduce the mean seam error is an empirical question, because the proposition gives pointwise equality only on the core and leaves the behavior elsewhere open. A protocol that separates the quantities is the following. Run a full-image denoiser from a fixed initial noise as the reference. Run the tiled scheme from the same noise with a chosen overlap width, exchange interval, and optional coarse channel. At the end of each interval, before the next fusion, record the seam error, the mean squared discrepancy between the tiles on their overlap,

$$E_{\text{seam}} = \frac{1}{|\Omega_{ij}|} \sum_{p \in \Omega_{ij}} \left\| x^i(p) - x^j(p) \right\|^2,$$

separately on the guaranteed core and outside it, and the structure error, the distance between the fused image and the reference after a low-pass operator L that retains object-scale content, for example a Gaussian blur whose width is set to the object scale,

$$E_{\text{struct}} = \left\| L \left(x_{\text{fused}} - x_{\text{ref}} \right) \right\|.$$

On the core, for tiles that agreed on the overlap at the preceding exchange, the seam error should vanish up to numerical precision if the assumptions of the proposition hold for the implementation, so a nonzero value there is a diagnosis of the implementation and not of the scheme. Outside the core, whether wider overlaps and shorter intervals reduce the seam error, and whether a coarse channel reduces the structure error, are the questions the experiment answers.

6 Growing Points and Repair

Coore's growing-point programs give the amorphous setting a mechanism for construction rather than only for coordination [4, 6]. A growing point is a locus of activity that moves through the medium, deposits material, responds to gradients, and may split, join, or terminate. The slides present the example of a chain of inverters built by a caricature of differentiation. A growing point seeded at an input contact lays down polysilicon for a fixed lifetime and then starts two further growing points, one heading up and one heading down. The upward point follows an increasing gradient toward a supply signal and, on sensing it, starts a transistor growing point and terminates. The transistor point runs for its own lifetime and starts diffusion growing points, one toward the supply and one toward a contact. The latter follows an attractor secreted by complementary diffusion, and when it reaches it, it drops a contact, starts a new growing point from an input contact with a fresh identifier, and terminates [6]. Each growing point secretes an inhibitor and carries an identifier that scopes its signals, which keeps one construction from interfering with another.

The slides emphasize that this very parallel process can be described through a serial metaphor, because the growing points provide a serial locus of control although the implementation is a uniform state machine in each computational particle [6]. This is the feature that makes repair a question about information. The program of a construction is spread across the particles, and at any moment its execution consists of the built material together with the state of each active growing point, including its type, its remaining lifetime, its tropism, and its identifier. Call the pair of built structure and active growing-point states the construction state σ.

Suppose a region fails halfway through building a connection. The failure erases both parts of σ inside the region. The surviving system may be able to rebuild coordinates for the damaged area by running the wave and relaxation steps of Section 1 again from the boundary values its neighbors still hold. Coordinates are a function of the field, and the field can be recomputed. The construction state is different. The survivors must determine which branch was active, where it was headed, and whether a connection had already been made, and the answers lie partly in what was built and partly in what was not recorded in the structure. Some of the missing state can be inferred from the surviving structure. If growth proceeds at a known rate from a known start, the length of a deposited segment records how many steps of the lifetime have elapsed, and a constant tropism toward a surviving gradient fixes the direction. Other state is not inscribed. Whether a growing point had already split, when the split children leave no material in the surviving region, is invisible to a survivor that holds only a partial segment, and an identifier that existed only in the failed particles cannot be reconstructed from material that does not carry it.

This is an instance of Definition 1. Let the target be the set of acceptable completed structures, so that the required output of a repair step is a continuation lying in that set. Let π be the map from the full construction state to what survives the failure, which comprises the surviving structure together with any records the neighbors retained. The map π plays the role of (x_i, m_i, r_i) in Definition 1: it is what the survivors can see, and the question is whether what

they see determines a continuation. A failure is repairable for a given program exactly when any two construction states with the same image under π call for continuations that can be served by a common continuation, which is to say that the fiber $\pi^{-1}(\pi(\sigma))$ lies within a single class of acceptable completions. A failure that destroys exactly the information distinguishing two classes is not repairable without additional retained state or external information. The target matters here as well. The surviving structure can suffice for the target that some acceptable structure be completed and fail for the target that the originally intended structure be completed, and a program that tolerates either of two functionally equivalent wirings is easier to repair than one that requires a specific layout.

The amorphous literature contains a second style of construction that avoids the problem by a different choice of where state lives. In the self-maintaining line, a gradient is created from one end point, and the opposite end point detects the gradient and tells its closer neighbor that it belongs to the line, which propagates up the gradient and marks the line as it goes [3]. The line is a function of a field that the processors keep refreshing, so a disruption of the line is followed by recomputation of the marking from the gradient. A growing point is a pointer, and a pointer destroyed in a failure takes its state with it. A field-based construction carries its state in a quantity that can be reconstituted from its surroundings. The tradeoff is that field-based constructions pay for continual refreshing and are limited to structures that can be expressed as functions of the field, while pointer-based constructions can express serial processes and need a record of how far the process had progressed. How much of the construction state must survive, or be retained elsewhere, is then a quantitative question: the smallest record such that every failure in a stated class leaves a fiber inside one class of acceptable completions.

7 Coherent Continuation

Return to the scattered processors. The first signals gave each processor enough information to locate itself approximately. The next exchanges had to relate overlapping coordinate systems, then to carry the state of a computation across communication intervals, then to keep tiles of a single object in step, and finally to recover a construction after an interruption. Each task puts a different demand on what the processors share, and the demands have a common form given by Definition 1: an exchange is judged against a target, and it is sufficient exactly when each part’s required output can be recovered from what that part holds.

The four settings each exhibit an exchange that is sufficient for one target and insufficient for another. Pairwise fits on overlaps suffice for the target that each overlap be matched and not for the target that the local systems assemble into one chart, because the loop defect is invisible to every pairwise fit, and the target determines which defect, if any, is the one to remove. Parameters together with a gradient correction suffice for the pooled fixed point at every number of local steps and for the pooled trajectory only through two, and a message that adds the pooled curvature suffices for the trajectory at every number of steps in the quadratic example. Overlap agreement suffices for pointwise equality on a core whose width is the overlap width minus twice the distance information travels between exchanges, and it

does not suffice for agreement on object-scale structure when the update depends on distant context, a dependence that a coarse channel carries only to the extent that the update factors through it. Surviving structure suffices for some acceptable completion where the fiber lies within one class and does not suffice for the originally intended one. In none of these cases does the agreement present at the moment of exchange decide the matter. The target does, together with what the exchange carries.

Two consequences follow. First, frequency and content trade against each other. A shorter interval requires less in the message, because less can diverge before the next exchange, and a longer interval requires a message that carries what the longer evolution will need. Wider overlap against a coarse channel, and fewer local steps against a control variate, are the same trade in different units. Second, admissibility is relative to a declared target. A single flat chart is a target for which the holonomy defect should vanish, an atlas that records its holonomy is a different target, a pooled optimum is a target for which the fixed point must be correct, and a reference image is a target for which both seam and structure errors are measured. Where the target is not stated, whether an exchange suffices has no answer.

The question that remains is one of cost, and it can be posed exactly. For a fixed target, a tolerance, and an interval between exchanges, what is the smallest message, or the smallest retained record, that still suffices? In the quadratic example the answer can be read off for each target. For the trajectory target, one number, w_0, suffices at $K = 1$, two numbers, w_0 and g_0, are needed through $K = 2$, and three, adding the pooled curvature, suffice at every K. For the fixed-point target the pair (w_0, g_0) suffices at every K. For denoising the corresponding comparison involves three quantities: the overlap width, the exchange frequency, and the size of the coarse representation, each bought against the seam error on and off the guaranteed core and against the object-scale structure error. For repair the question is how much construction state must survive a failure of a given extent, and in what form, so that the survivors determine an acceptable completion. For assembly it is how much about the loops a region must be told beyond its own overlaps. Only the quadratic case is answered here, and even there the cheapest message that is sufficient for the trajectory at every K is so rich that it makes the local computation redundant, so the question has to be asked of exchanges that leave the local evolution doing real work. The others are stated so that the answer in each can be sought against a named target, and so that the cost of a sufficient exchange is measured by what it carries and not by how much agreement it displays when it is sent.

References

[1] Harold Abelson, Don Allen, Daniel Coore, Chris Hanson, George Homsy, Thomas F. Knight Jr., Radhika Nagpal, Erik Rauch, Gerald Jay Sussman, and Ron Weiss. Amorphous Computing. *Communications of the ACM*, 43(5):74–82, 2000.

[2] Omer Bar-Tal, Lior Yariv, Yaron Lipman, and Tali Dekel. MultiDiffusion: Fusing Diffusion Paths for Controlled Image Generation. In *Proceedings of the 40th International Conference on Machine Learning*, 2023.

[3] Lauren Clement and Radhika Nagpal. Self-Assembly and Self-Repairing Topologies. In *Workshop on Adaptability in Multi-Agent Systems, RoboCup Australian Open*, 2003.

[4] Daniel Coore. Botanical Computing: A Developmental Approach to Generating Interconnect Topologies on an Amorphous Computer. PhD thesis, Massachusetts Institute of Technology, 1999.

[5] Daniel Coore, Radhika Nagpal, and Ron Weiss. Paradigms for Structure in an Amorphous Computer. A.I. Memo 1614, Massachusetts Institute of Technology, Artificial Intelligence Laboratory, October 1997.

[6] DARPA Ultrascale Computing Program. Amorphous Computing. Presentation slides, 1 April 1998. `http://groups.csail.mit.edu/mac/projects/amorphous/tucson-talk/`

[7] Flyxion. Hierarchical Synchronization. Essay, 2026.

[8] Peter Gray and Stephen K. Scott. Autocatalytic Reactions in the Isothermal, Continuous Stirred Tank Reactor: Oscillations and Instabilities in the System $A + 2B \to 3B$; $B \to C$. *Chemical Engineering Science*, 39(6):1087–1097, 1984.

[9] Richard Hartley, Jochen Trumpf, Yuchao Dai, and Hongdong Li. Rotation Averaging. *International Journal of Computer Vision*, 103(3):267–305, 2013.

[10] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising Diffusion Probabilistic Models. In *Advances in Neural Information Processing Systems 33*, 2020.

[11] Sai Praneeth Karimireddy, Satyen Kale, Mehryar Mohri, Sashank J. Reddi, Sebastian U. Stich, and Ananda Theertha Suresh. SCAFFOLD: Stochastic Controlled Averaging for Federated Learning. In *Proceedings of the 37th International Conference on Machine Learning*, 2020.

[12] Xiang Li, Kaixuan Huang, Wenhao Yang, Shusen Wang, and Zhihua Zhang. On the Convergence of FedAvg on Non-IID Data. In *International Conference on Learning Representations*, 2020.

[13] H. Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Agüera y Arcas. Communication-Efficient Learning of Deep Networks from Decentralized Data. In *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*, 2017.

[14] Radhika Nagpal. Organizing a Global Coordinate System from Local Information on an Amorphous Computer. A.I. Memo 1666, Massachusetts Institute of Technology, Artificial Intelligence Laboratory, August 1999.

[15] Jascha Sohl-Dickstein, Eric A. Weiss, Niru Maheswaranathan, and Surya Ganguli. Deep Unsupervised Learning Using Nonequilibrium Thermodynamics. In *Proceedings of the 32nd International Conference on Machine Learning*, 2015.

[16] Yang Song, Jascha Sohl-Dickstein, Diederik P. Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-Based Generative Modeling through Stochastic Differential Equations. In *International Conference on Learning Representations*, 2021.

[17] Sebastian U. Stich. Local SGD Converges Fast and Communicates Little. In *International Conference on Learning Representations*, 2019.

[18] Alan M. Turing. The Chemical Basis of Morphogenesis. *Philosophical Transactions of the Royal Society of London B*, 237(641):37–72, 1952.