Supplementary Material

Appendix 1: justification for gradient-free weight update rule

The optimization of the spot penalties uses a cost function that is entirely determined by the total probability \( P(s) \) of mapping a locus to a given spot \( s \) in the image:

\[
P(s) = \sum_L p_{L \rightarrow s} = \frac{1}{Z} \sum_L Z_{L \rightarrow s}.
\]

Both \( Z_{L \rightarrow s} \) and \( Z \) sum terms that are products of the spot penalties \( q_s \), owing to the fact that a direct influence on any one mapping probability indirectly influences all other mapping probabilities. However, when the penalty on some spot \( s \) is far away from its proper value, any mapping probability \( p_{L \rightarrow s} \) to that spot tends to be saturated very close to either 0 or 1. In this case we make the approximation that a given penalty factor \( q_s \) only affects any given \( p_{L \rightarrow s} \) directly as a multiplying factor, and does not affect the mapping probabilities to other spots, since the indirect influences are small until the mapping probability is out of saturation. Under this approximation \( Z_{L \rightarrow s} \approx f_s \cdot a_s \), and \( NZ = \sum_L \sum_s Z_{L \rightarrow s} \approx f_s \cdot a_s + b_s \), so

\[
P(s) \approx \sum_L \frac{f_s N}{f_s + (b/a_s)}.
\]

Since our objective is to find an updated \( q'_s \) causing the sum of mapping probabilities to be some target \( P'(s) \), we also write:

\[
P'(s) \approx \sum_L \frac{f'_s N}{f'_s + (b/a_s)}.
\]

Solving Eqs 5 and 6 together to eliminate the unknown \( b_s/a_s \) gives us the update rule:

\[
f'_s \approx \frac{f_s N - 1}{f_s N} \cdot f_s.
\]

The cost function contains two terms: 1) a penalty for the overall difference between the expected rate of missing spots \( p_f n \) versus that inferred from the summed \( P(s) \); and 2) a penalty on \( P(s) \) if it exceeds 1, which would indicate a > 100% likelihood of a locus mapping to that spot. Our current implementation simply makes two updates, one pushing \( P(s) \rightarrow p_f n \) and the second pushing \( P(s) \rightarrow 1 \) for spots violating normalization. Doing so leads to the update rule given by Eq. 1.

Appendix 2: series expansion derivations

Here we prove a) each series converges to the true partition function when all terms are counted, and that b) a given truncation of each series using our recommended selection of series terms counts every possible legal or overlapping conformation zero or more times (despite many series terms having negative coefficients), and therefore produces positive mapping probabilities. Notice that both expansions count all legal (non-overlapping) conformations once from \( Z_0 \), and the goal of considering higher-order series terms is thus to minimize the weight of the illegal overlapping conformations without their weights ever becoming negative. We note that conformations overlapping at adjacent loci are automatically eliminated from all terms in the calculation, so the individual conformations we consider here are assumed to overlap themselves only between non-adjacent loci.

For our proofs we will define an illegal overlap as a set of loci in an illegal conformation mapping to a given spot in the image. Illegal overlaps are to illegal conformations as illegal constraints are to higher-order series terms, and we use the same set notation for overlaps as we do for constraints.

Our derivations make repeated use of the following identity, taken from Eq. 3.1.7 in Ref. [15].

\[
\sum_{b=0}^a \binom{a}{b} (-1)^b = 0 \text{ for } a > 0
\]

This relation follows from the fact that the left-hand side is a series expansion for \((1 - 1)^a\).

Series expansion 1 A given conformation bearing a set of illegal overlaps \( \theta \) is counted by each constrained partition function whose set of illegal constraints (indices) are a subset of \( \theta \). Thus the weight \( W_\theta \) of this conformation in the full partition function \( Z \) is:

\[
W_\theta = \sum_{\phi \subseteq \theta} w_\phi = \sum_{n_\phi \geq 0} \binom{n_\phi}{n_\theta} (-1)^{n_\phi}
\]

where we have used Eq. 7 in the final step. This result shows that series expansion 1 counts only conformations having no overlaps, i.e. for which \( n_\theta = 0 \).

Suppose that one follows our prescription for evaluating a subset of terms, namely all terms that are a subset of overlaps \( \psi \). Then using the same formula, we reason that a conformation with overlaps \( \theta \) will be given the following weight:
\[ W_{\theta}^{(\psi)} = \sum_{\phi \subseteq (\theta \cap \psi)} w_\phi \]
\[ = \begin{cases} 
1 & \text{for } \theta \cap \phi = \emptyset \\
0 & \text{otherwise.} 
\end{cases} \]

Thus our selection of series terms for expansion 1 eliminates all conformations from the partition function having any overlaps contained in our set \( \psi \).

**Series expansion 2** Expansion 2 allows unconstrained loci to revisit spots that already have constrained loci. In this case, a conformation whose set of nonadjacent overlaps is denoted \( \theta \) will be counted by a partition function having overlaps \( \phi \) if each individual overlap \( \phi_i \) is contained in an individual overlap \( \theta_i \), in the sense that \( \phi_i \) maps any subset of \( n_\phi^i > 1 \) loci in \( \theta_i \) to the same spot as \( \theta_i \). Then the weight of this conformation in the expansion is

\[ W_\theta = \sum_{\phi = (\theta \subseteq \theta_i)} w_\phi \]
\[ = \prod_{k=1}^{n_{s}} \left[ 1 + \sum_{n_{k}^\phi = 1}^{n_k} \binom{n_k^\phi}{n_k^\phi} (-1)^{n_k^\phi-1}(n_k^\phi - 1) \right] \]
\[ = \prod_{k=1}^{n_{s}} \sum_{n_{k}^\phi = 0}^{n_k^\phi} \binom{n_k^\phi}{n_k^\phi} (-1)^{n_k^\phi-1}(n_k^\phi - 1) \]
\[ = \prod_{k=1}^{n_{s}} n_k^\phi \binom{n_k^\phi - 1}{n_k^\phi - 1} (-1)^{n_k^\phi - 1} \]

where the fact that \( n_k^\phi > 0 \) allowed Eq. 7 to eliminate a term in the last line. Defining \( m_k^\phi = n_k^\phi - 1 \) gives us:

\[ W_\theta = \prod_{k=1}^{n_{s}} n_k^\phi \sum_{m_k^\phi = 0}^{n_k^\phi - 1} \binom{n_k^\phi - 1}{m_k^\phi} (-1)^{m_k^\phi} \]
\[ = \begin{cases} 
1 & \text{for } n_s = 0 \\
0 & \text{otherwise.} 
\end{cases} \]

Again using Eq. 7. Therefore the complete series counts only non-overlapping conformations.

Our prescription for choosing series terms is to select a set of single-locus-to-spot mappings \( \Psi \), and generate all series terms whose illegal overlaps are built only from mappings in \( \Psi \). To be formal, we will compile all locus-to-spot mappings in \( \Psi \) for each given spot \( s \) (assuming there are \( 2 \) or more) into a single overlap, and use \( \psi \) to denote the set of these overlaps: then the series order \( N_\psi \) is the sum of the number of loci over all elements \( \psi_i \). Our rule is to include all series terms \( \phi \) whose illegal overlaps are built entirely from subsets of \( \psi \), in the sense that a given overlap to spot \( s \) contains a subset of the loci in the element \( \psi_i \) that maps to spot \( s \). For example if \( \psi \) contains the element \( (ACE)_s \) then \( \phi \supseteq \{(AC)_s,(CE)_s,(AE)_s,(ACE)_s\} \). Using these rules, a given conformation having illegal overlaps \( \theta \) will be given the following weight in the expansion:

\[ W_\theta^{(\psi)} = \sum_{\phi = (\theta \subseteq \theta_i)} w_\phi \]
\[ = \prod_{k=1}^{n_{s}} \sum_{n_{k}^\theta = 0}^{n_k^\theta} \binom{n_k^\theta}{n_k^\theta} (-1)^{n_k^\theta-1}(n_k^\theta - 1) \]
\[ = \begin{cases} 
1 & \text{if all } n_k^\theta < 2 \\
0 & \text{otherwise.} 
\end{cases} \]

Thus series 2 eliminates all conformations having any overlap \( (L_1,L_2,\ldots) \to s \) where \( L_1 \to s \) and \( L_2 \to s \) are contained in \( \Psi \).
Appendix 3: Supplementary figures

Figure S1. Information recovery from individual simulations. Conformation #1 of each series is shown at far left. Panels to the right show unrecovered information \( I \), entropy \( S \) and \( \log Z \) as a function of the number of series terms included. Dot-dashed lines show the unrecovered information of \( \tilde{Z}_{opt} \) using optimized spot penalties on top of the given number of the series terms.
Figure S2. Accuracy of entropy as a proxy for unrecovered information. Distributions showing the difference between entropy $S$, which is a blind estimate of unrecovered information, and actual unrecovered information $I$ in each of the 3 simulation scenarios considered, as a function of number of series terms. No spot penalties were used for these results. Each distribution shown encompasses the $S - I$ curves of all 100 simulated reconstructions in one scenario.
Figure S3. Accuracy of mapping probabilities. Binned mapping probabilities (x axis) versus the fraction of true mapping probabilities in each bin (y axis), averaged over the various simulated experiments in each experimental scenario. Grey shaded regions show the 3σ range of uncertainty due to counting error.
Figure S4. Partition function $Z$ versus number of series terms. A. Relationship between log $Z$ and the number of series terms for each simulated scenario, calculated as a median average of the relationships found in each of the 100 individual simulations in each scenario. B. Distributions of the difference between log $Z$ calculated using series 2 and series 1. The fact that this quantity is generally negative when some but not all series terms are included shows that series 2 recovers information (i.e. removes unrecovered information $I$) faster than series 1.