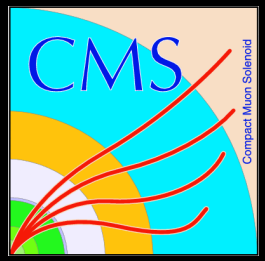
ECAL Clustering

Andrew Askew



Clustering:

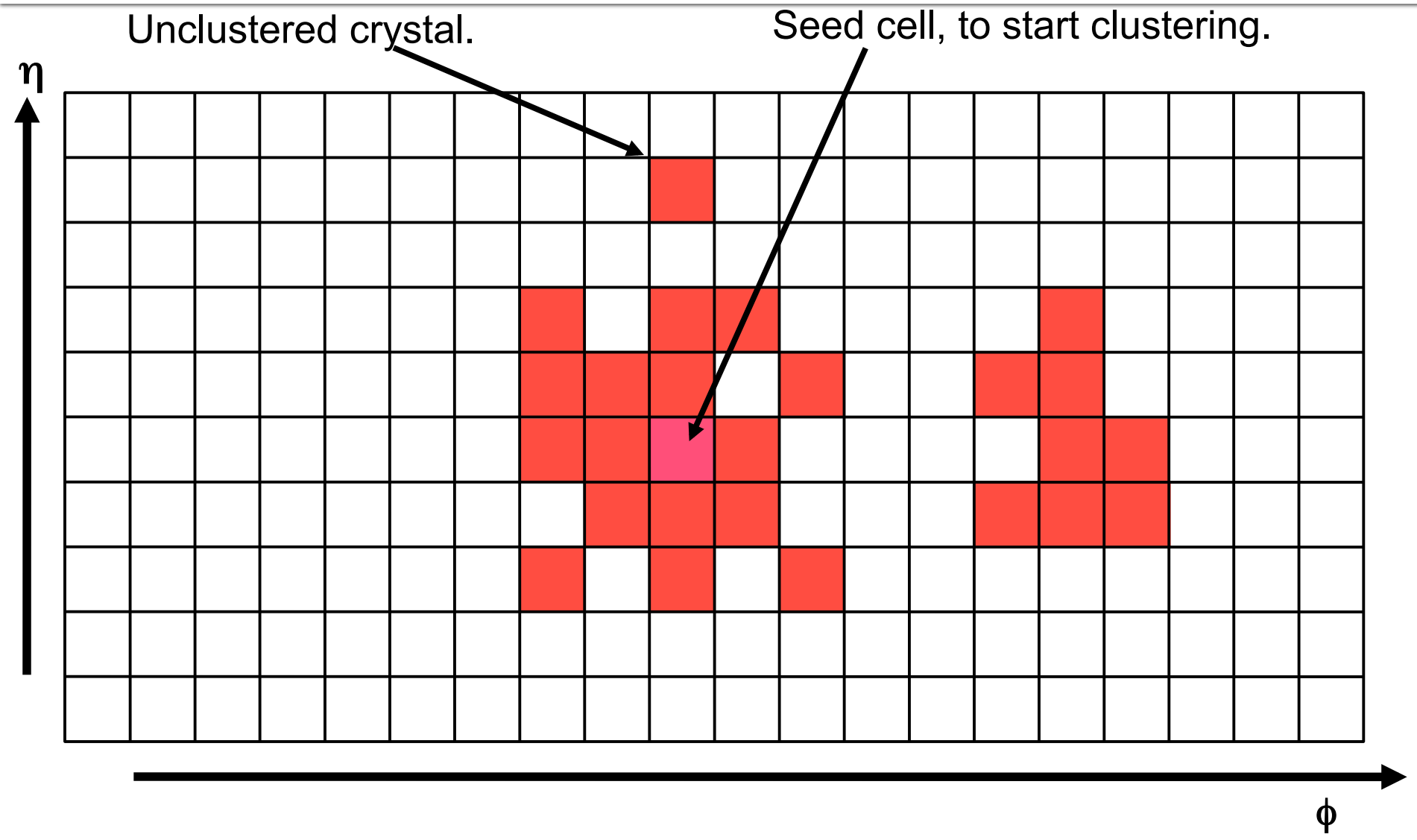
- What is the goal of clustering?
  - Logically group local depositions of energy under the hypothesis that the channels together represent an electromagnetic shower.
- What is Superclustering?
  - Logically grouping clusters under the hypothesis that they originated from the same electromagnetic particle produced in the collision.
- Hopefully this sounds consistent. What I have found in studying this is that typically complications arise in telling the algorithm when to terminate.

- We presently have have three different ways of clustering the crystals of the ECAL.
  - Barrel: Hybrid superclustering
  - Endcap: multi5x5 clustering/superclustering
  - Particle Flow clustering. Doesn't natively supercluster, but new developments put this at the block/particle level.
- All of these are "seeded" algorithms, they start with a particular channel that satisfies certain criteria.
- Let's examine what each of them does.

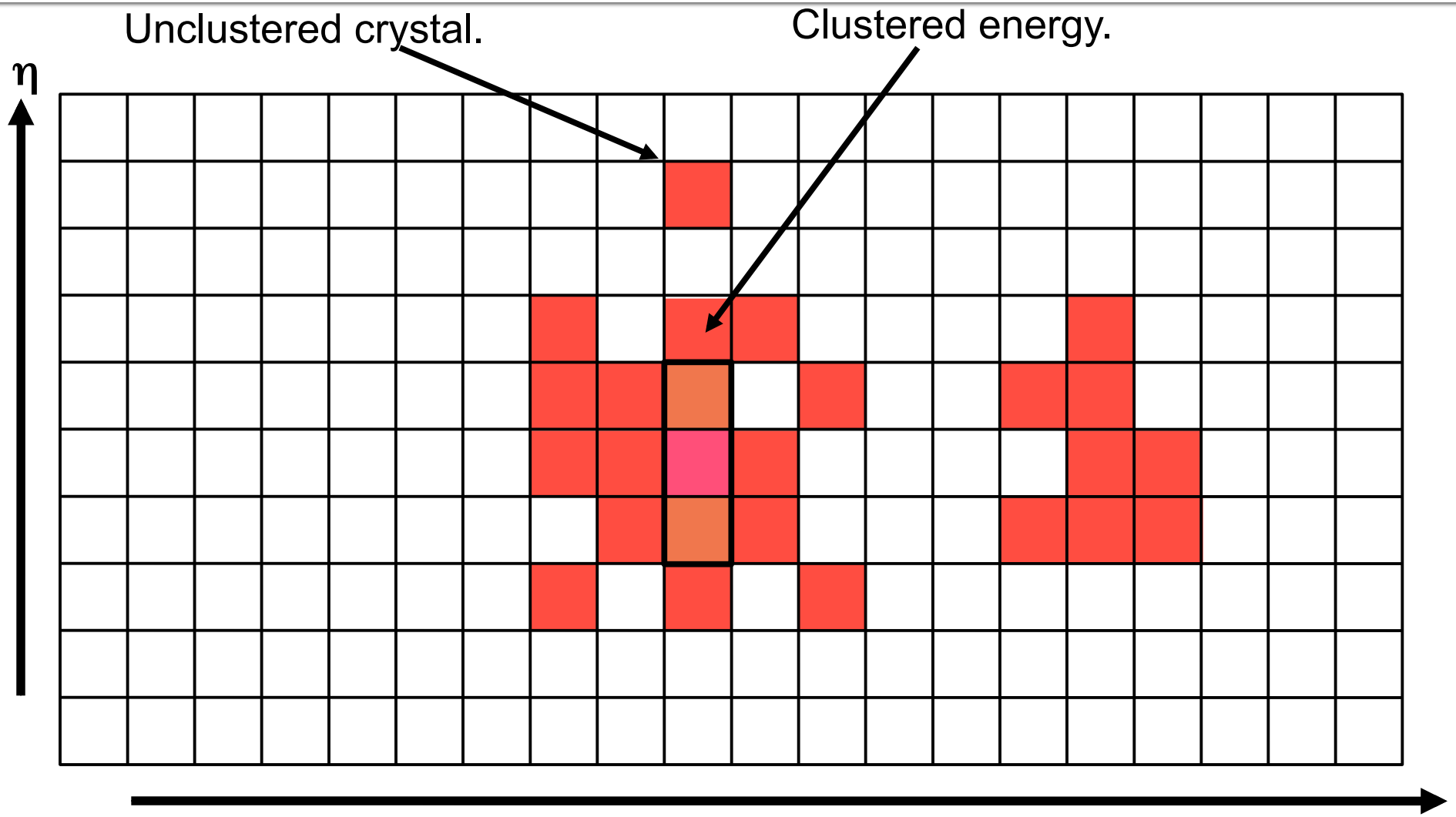
The "Hybrid" algorithm

- The "Hybrid" algorithm forms one logically large area about a seed, and then groups the local maxima within that area.
  - Thus it natively "superclusters"
  - Seeds here are just all crystals above a set threshold. Start with highest energy seeds first, and cannot reuse clustered crystals.
- The region it forms is a rectangle of fixed width in η, and a configurable length in ϕ. All crystals above a configurable thresholds are seeds, the seeds are sorted, and the algorithm starts from the highest energy crystal.

EB Clustering (same for e/γ):

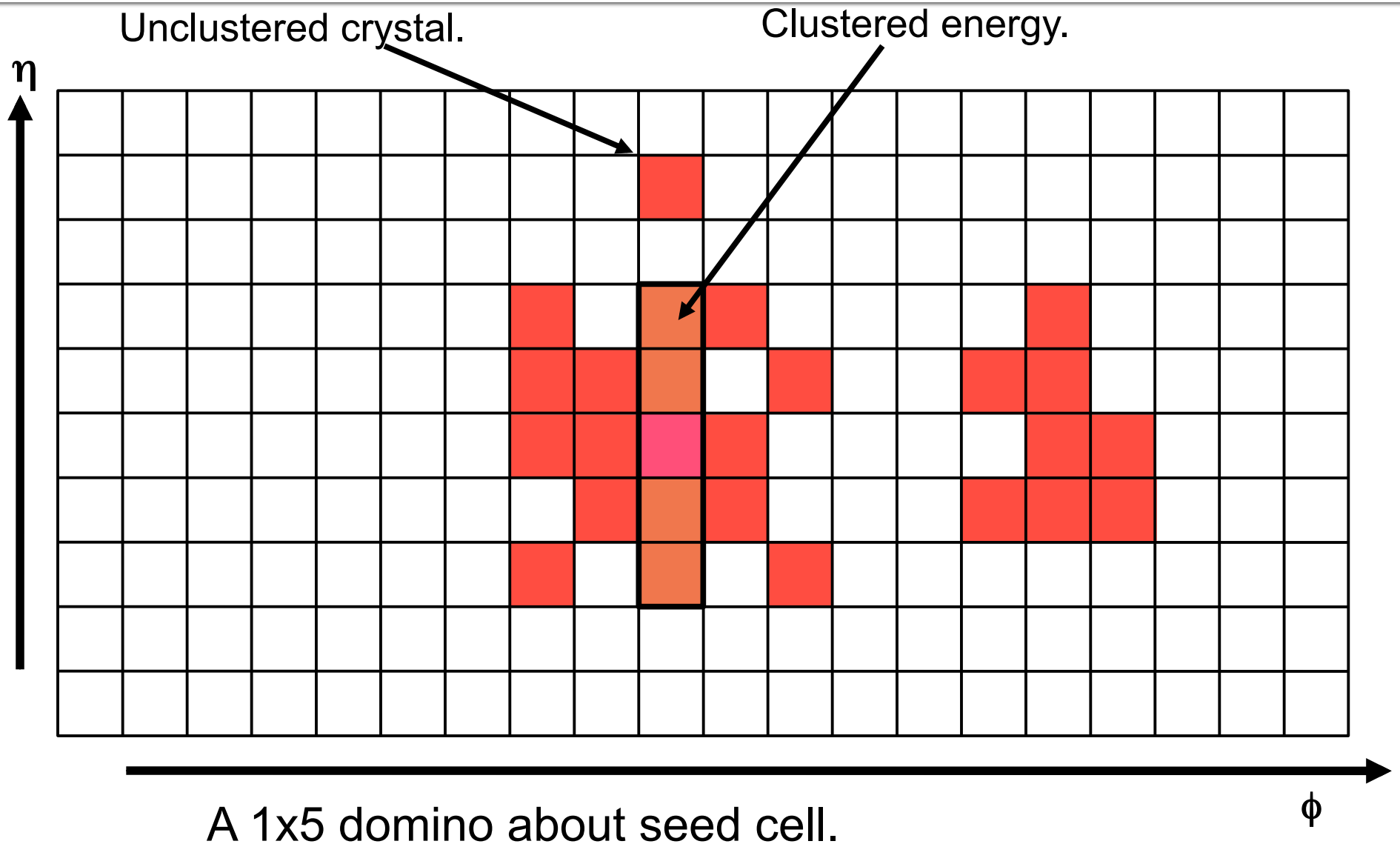


EB Clustering (same for e/γ):



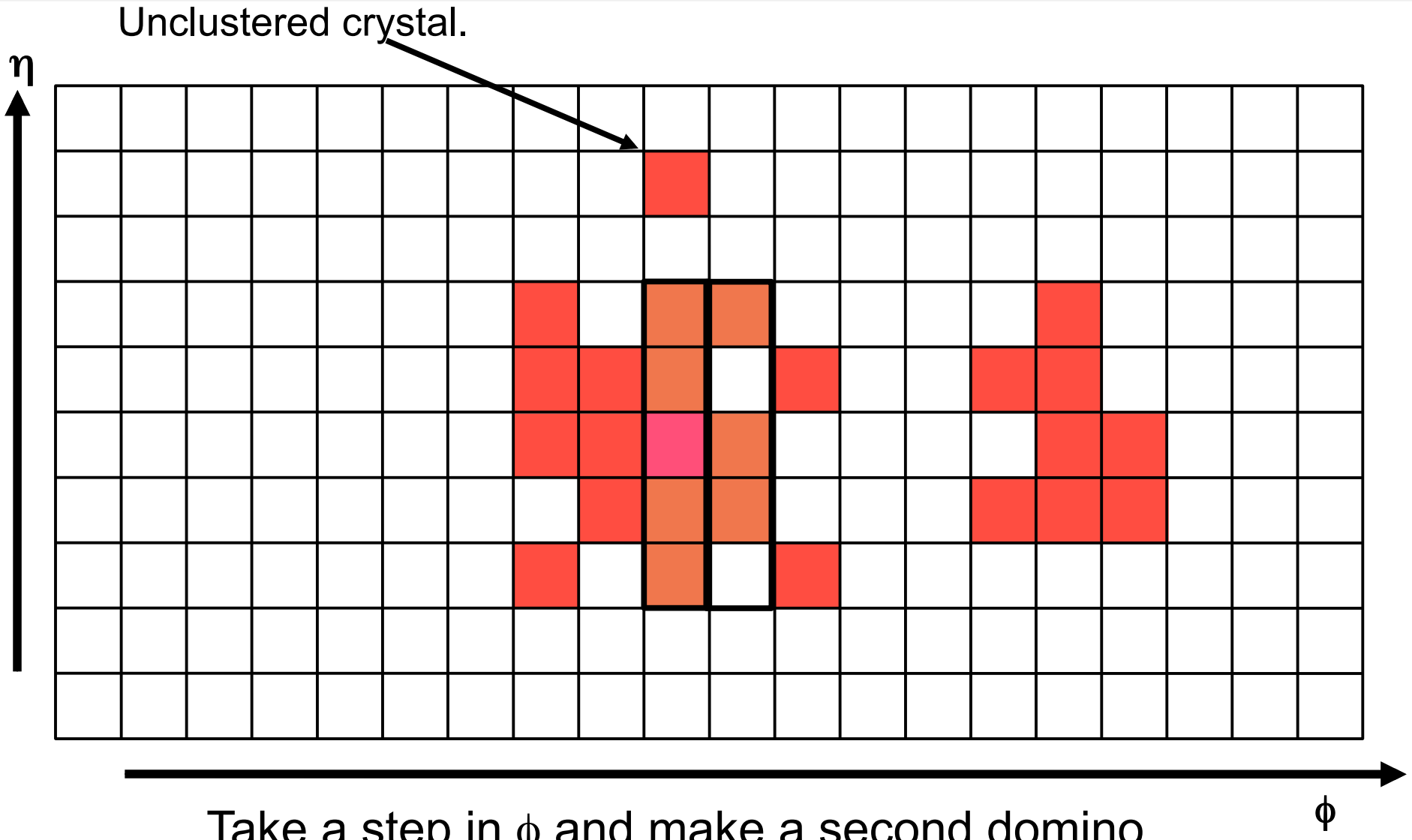
A 1x3 domino about seed cell. If there is sufficient energy, the “wing” cells will be added to make a 1x5

EB Clustering (same for e/γ):



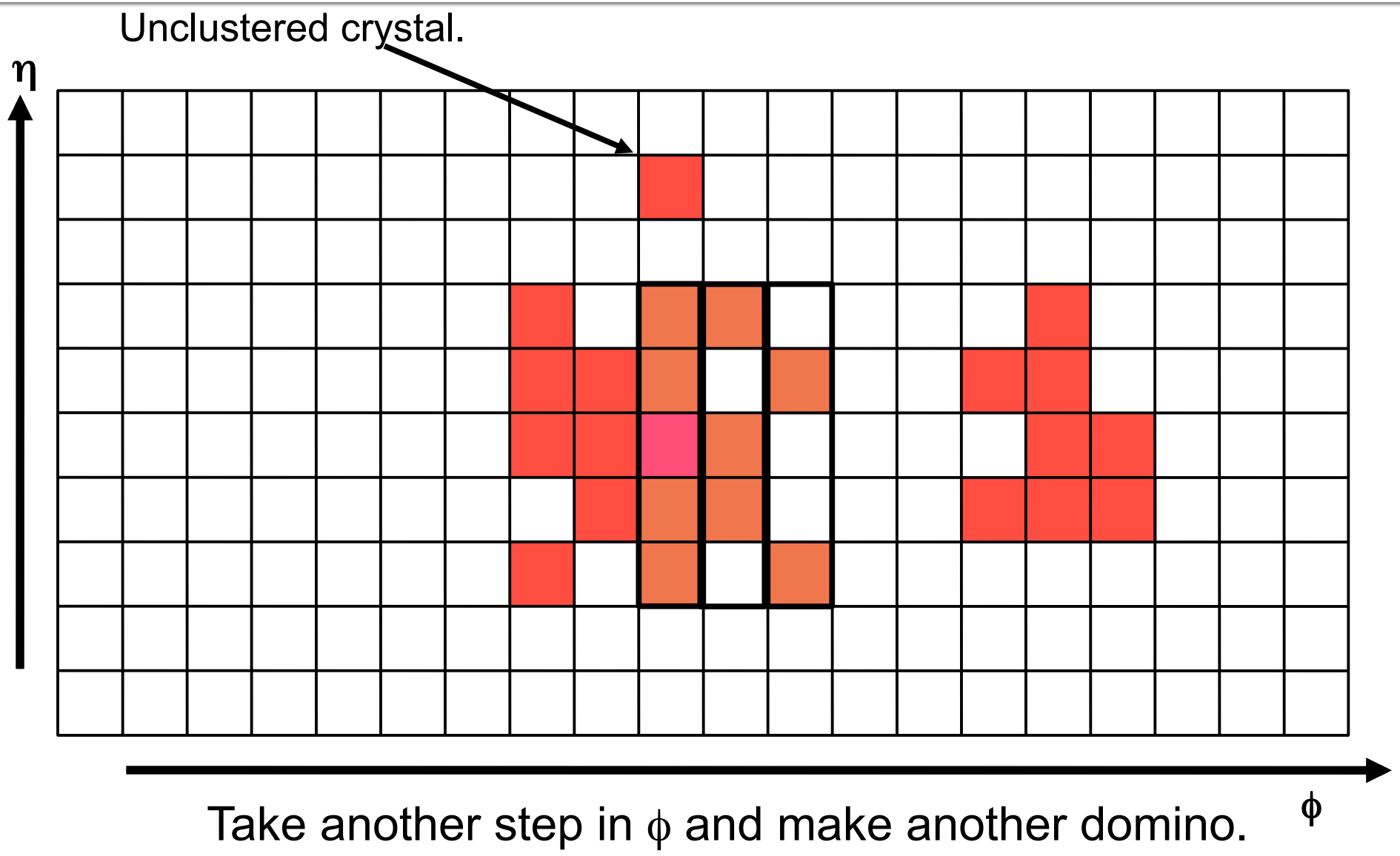
A 1x5 domino about seed cell.

EB Clustering (same for e/γ):

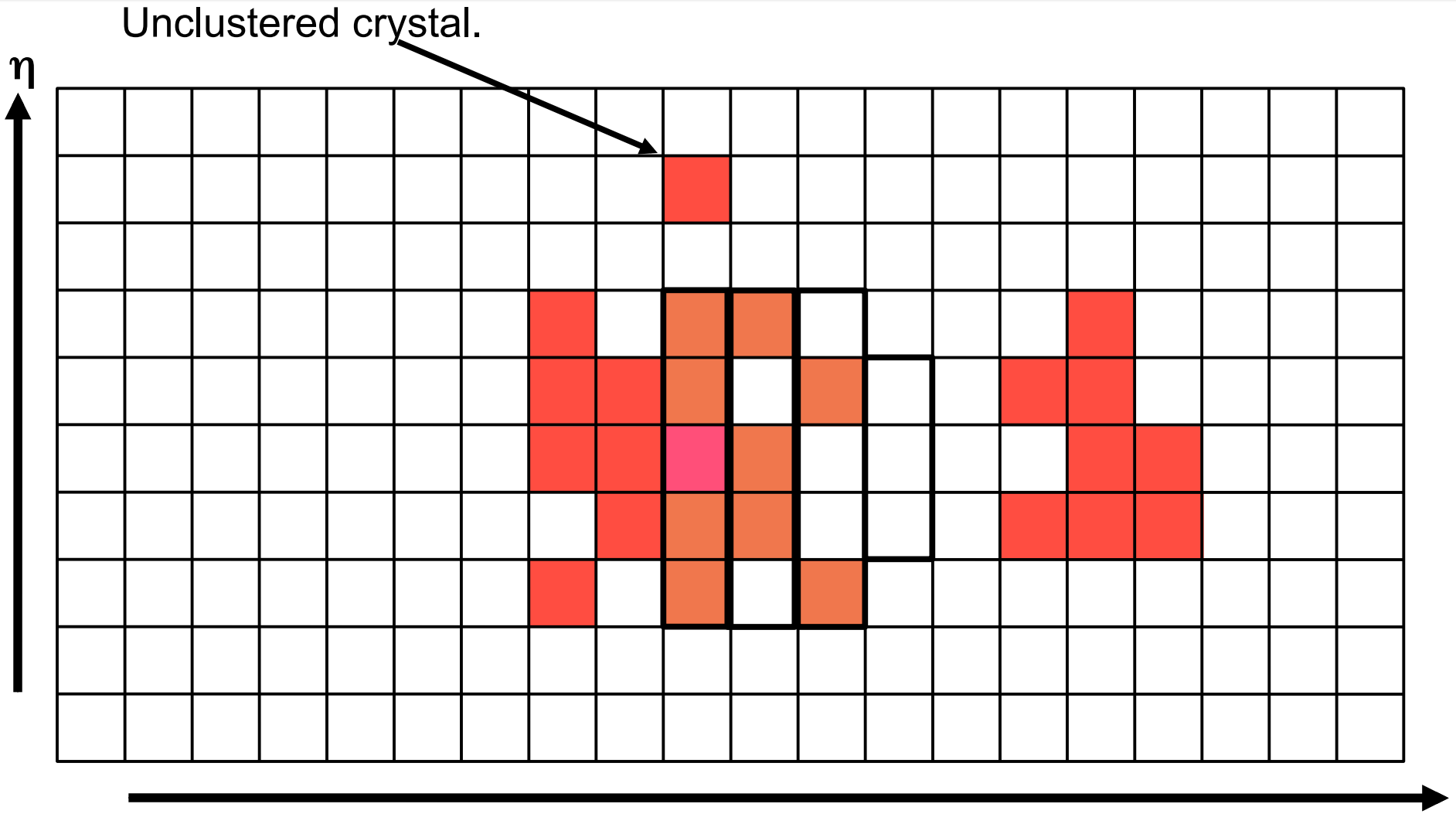


Take a step in ϕ and make a second domino.

EB Clustering (same for e/γ):

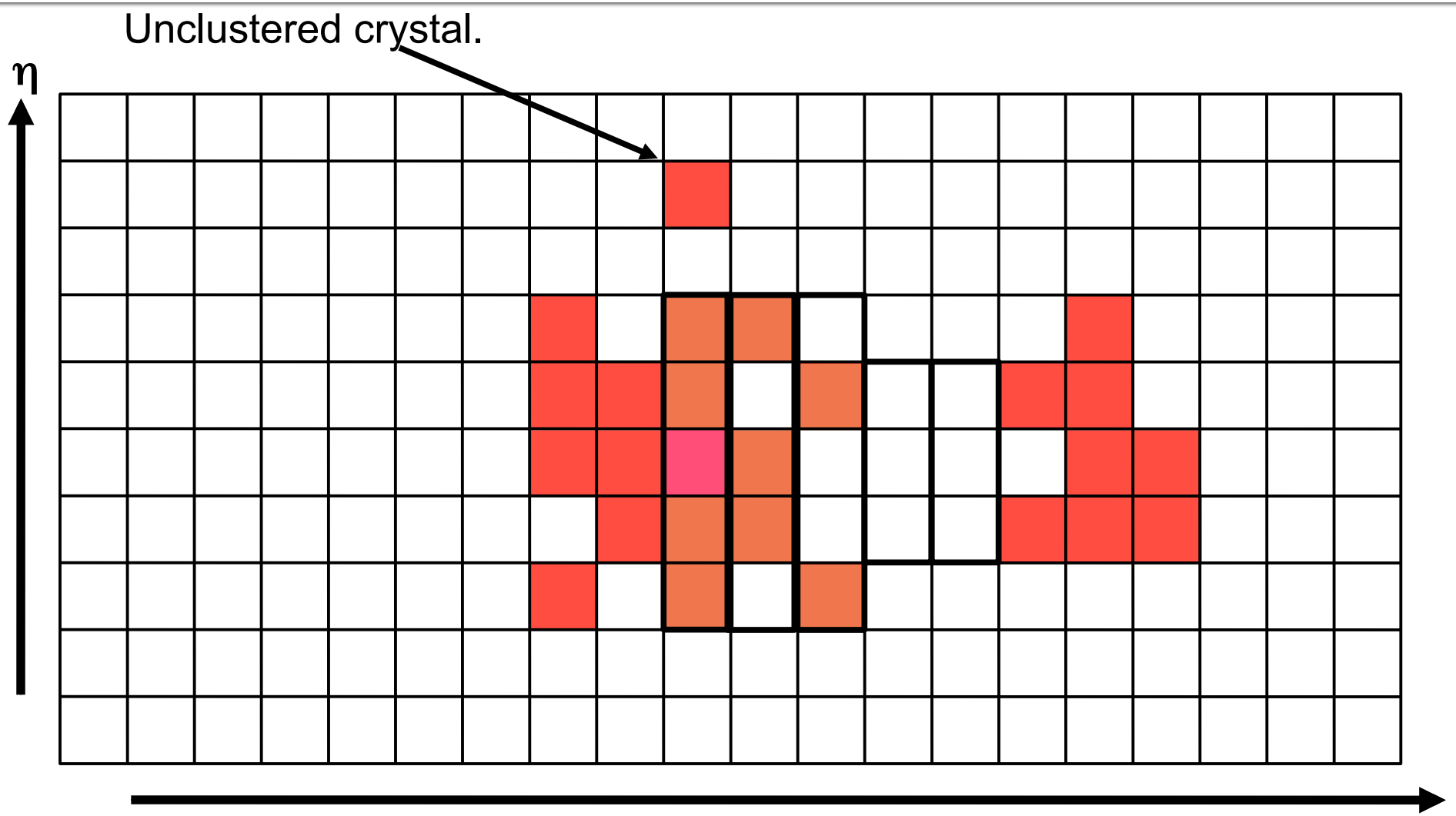


EB Clustering (same for e/γ):



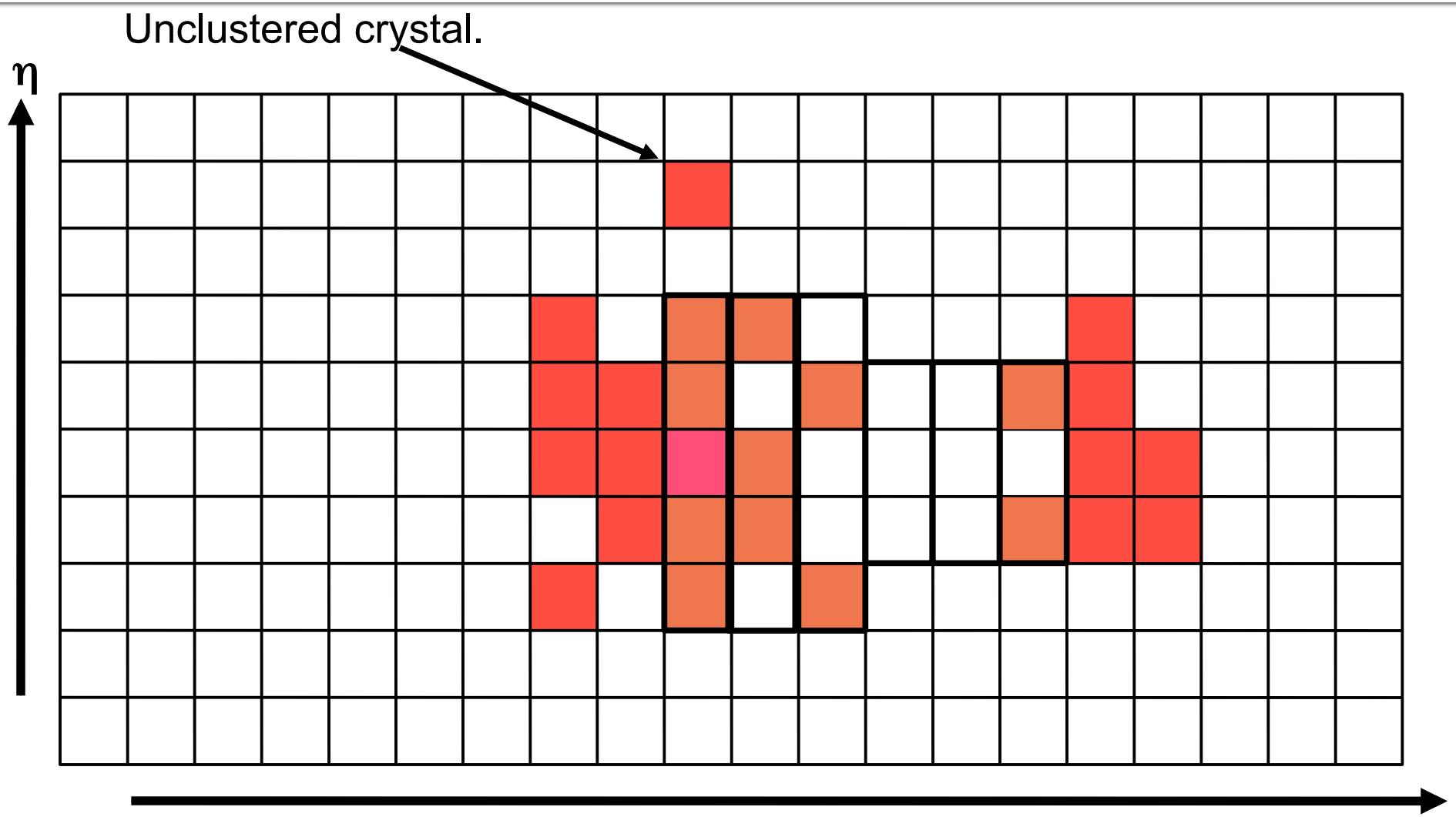
Take another step in ϕ. This intervening space is fine, the algorithm simply won't cluster it and move on in ϕ.

EB Clustering (same for e/γ):



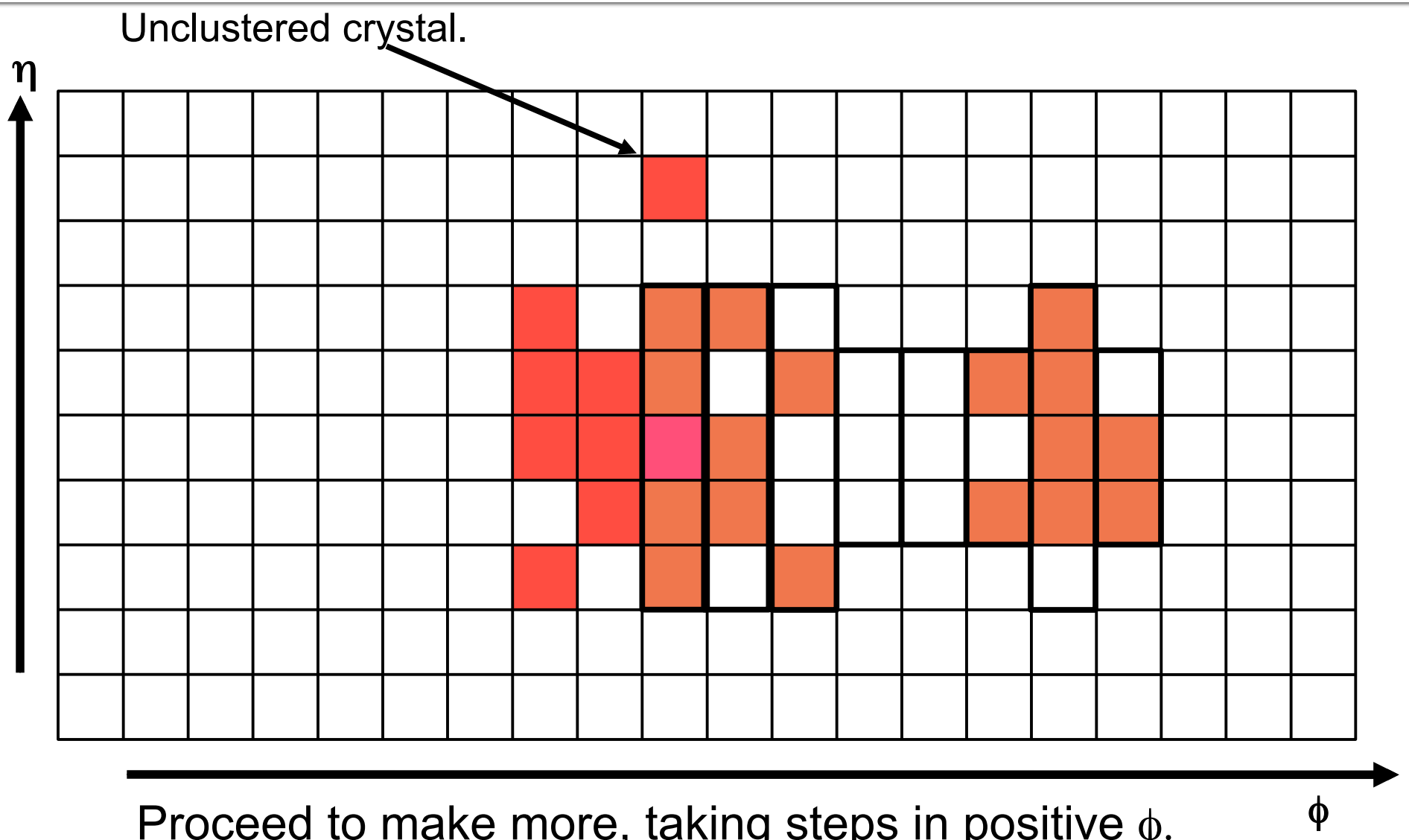
Take another step in ϕ. This intervening space is fine, the algorithm simply won't cluster it and move on in ϕ.

EB Clustering (same for e/γ):

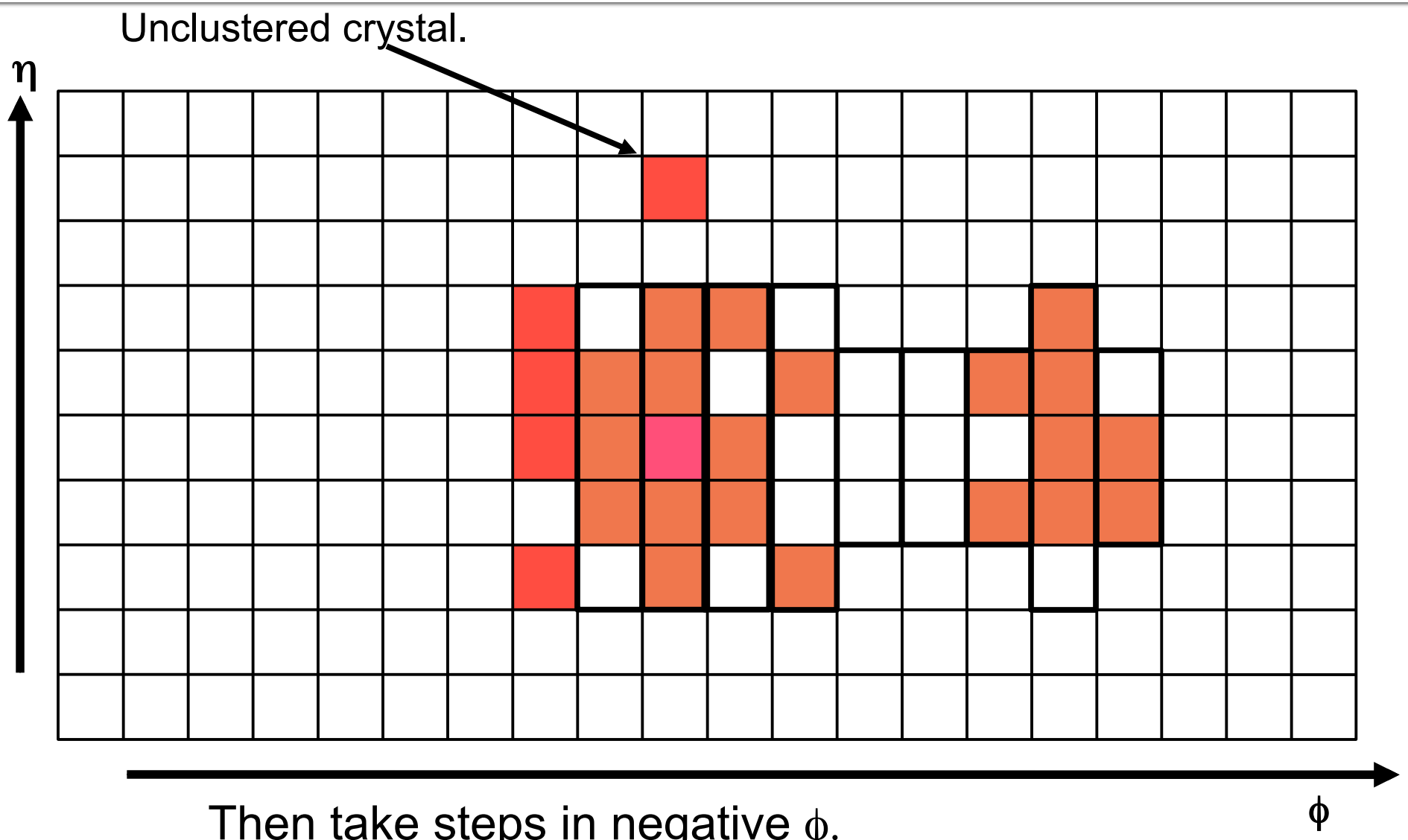


Take another step in ϕ. Found a new energy deposit, will be clustered, but will likely be a separate basic cluster.

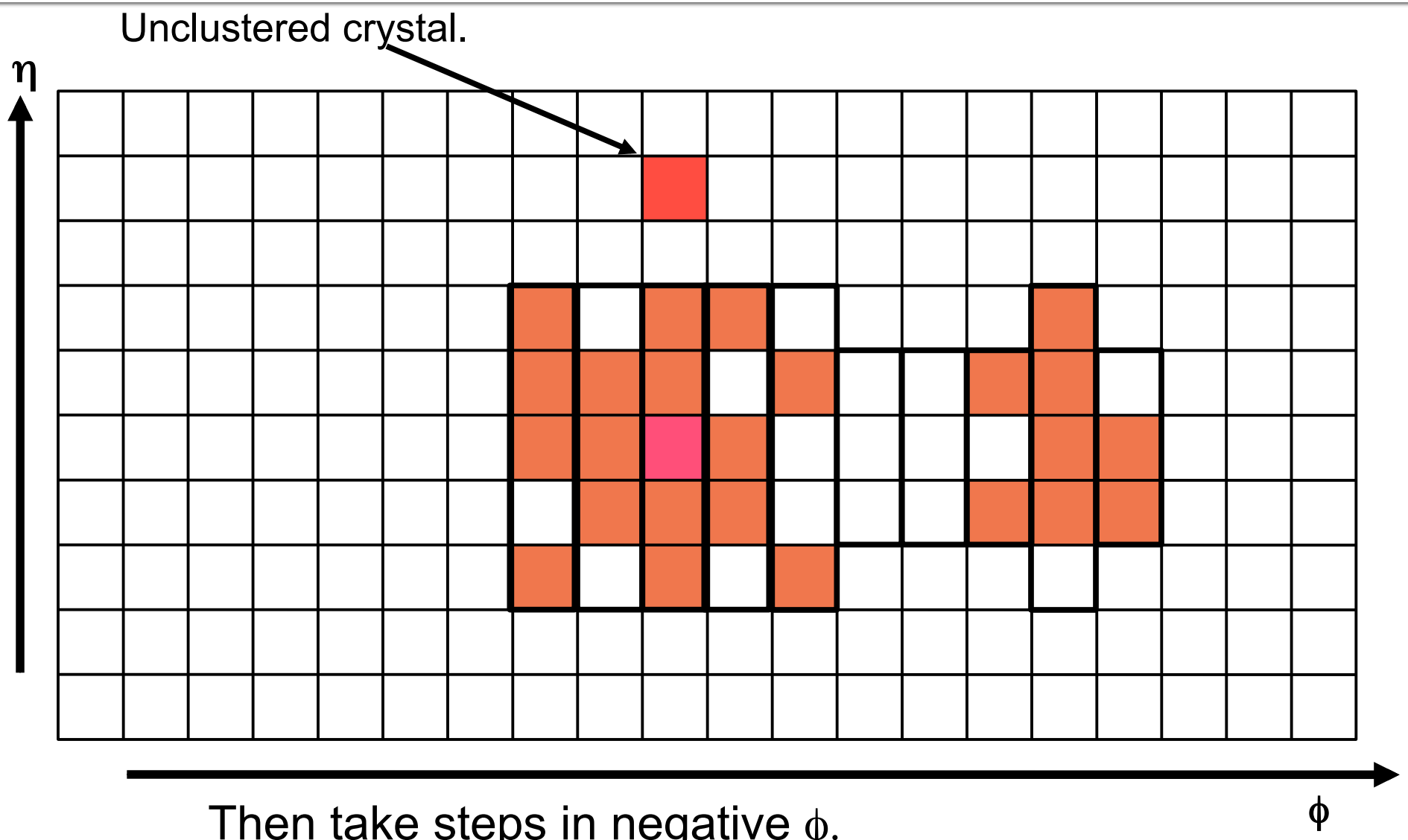
EB Clustering (same for e/γ):



EB Clustering (same for e/γ):



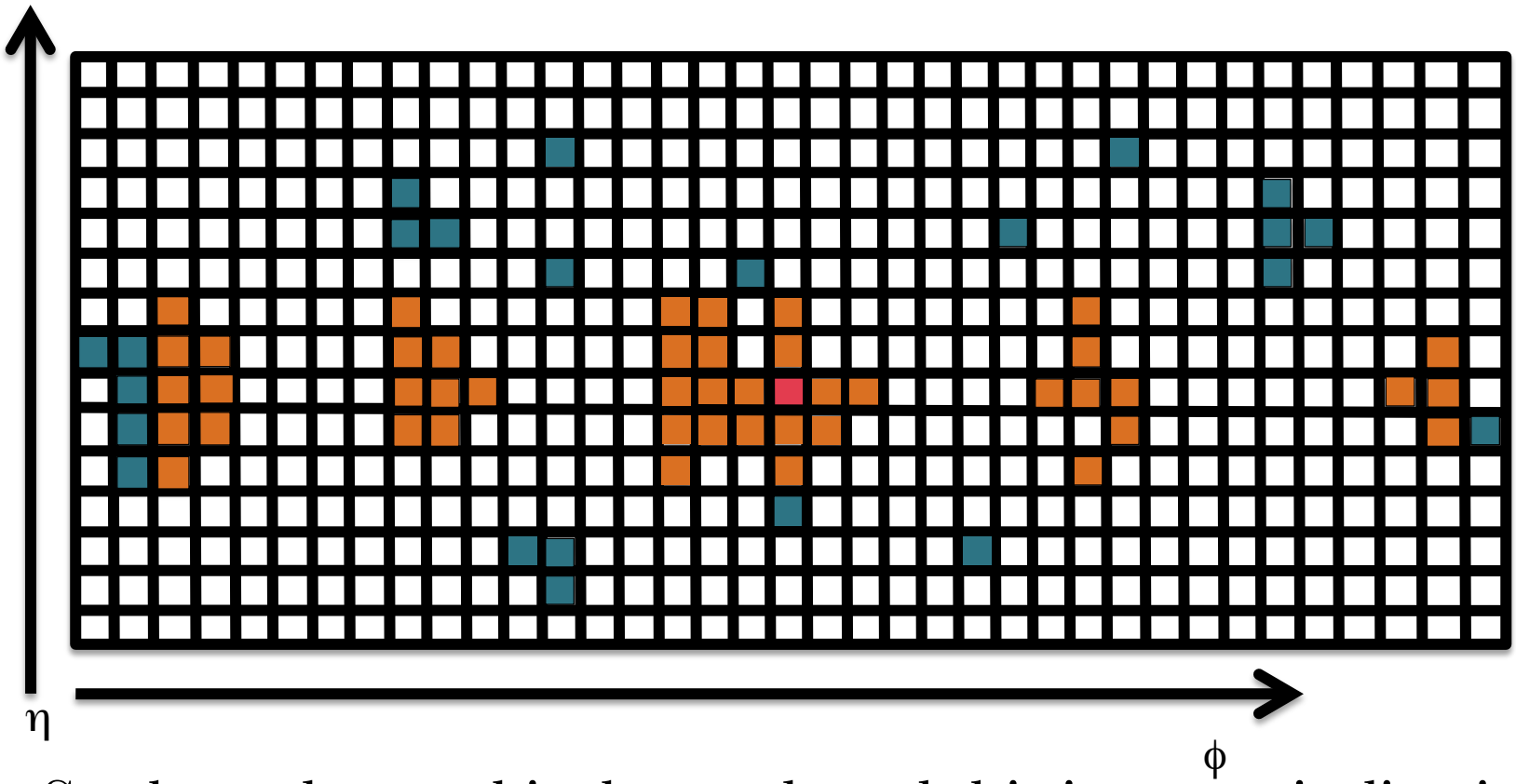
EB Clustering (same for e/γ):



Hybrid Clustering Parameters:

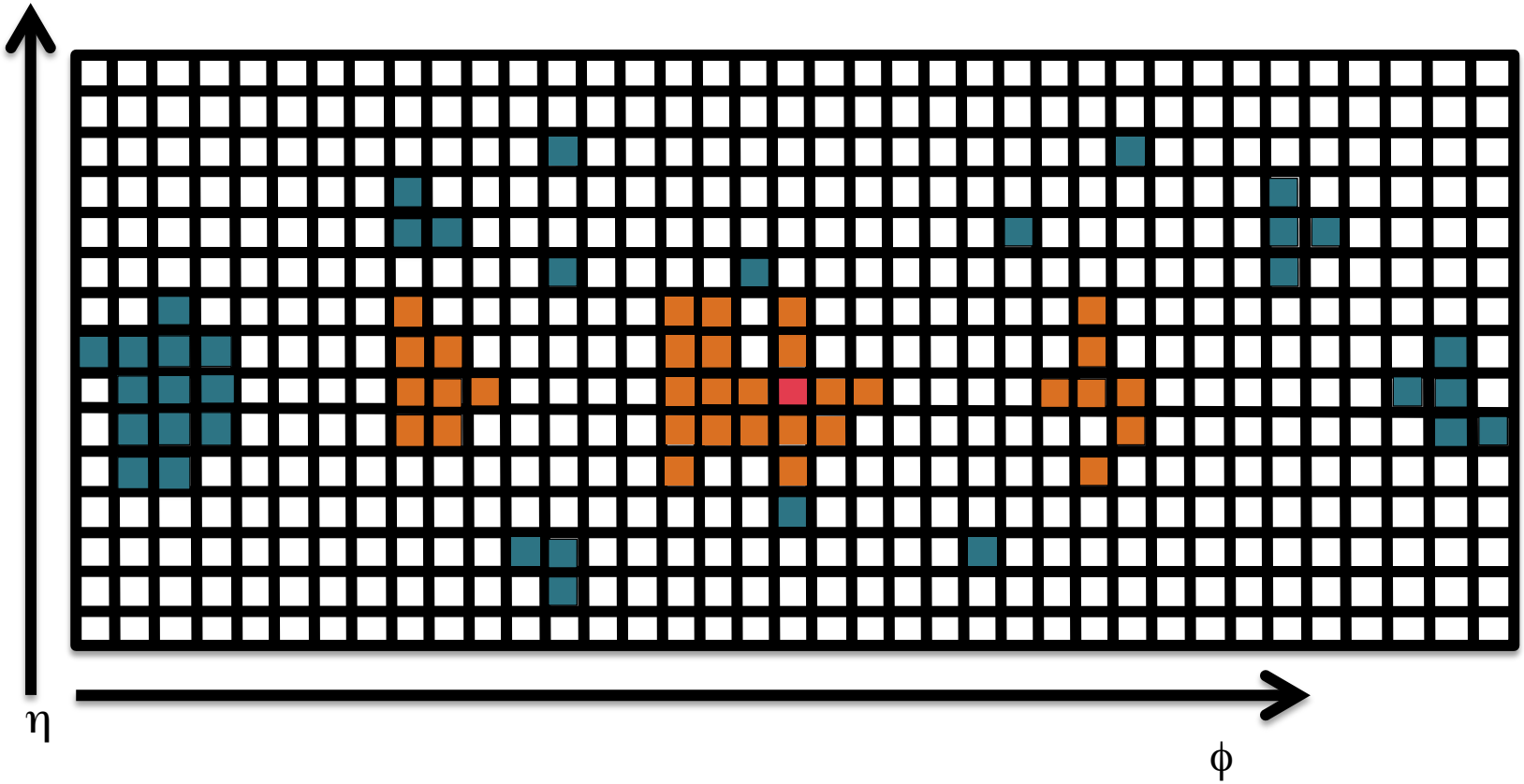
- HybridBarrelSeedThr: crystal transverse energy needed to start a supercluster (1 GeV in E_T)
- Step: Number of crystals in positive and negative phi to consider (17).
- Ethres: Threshold for including a particular domino (0.100).
- Eseed: Threshold for turning a local maximum into a basic cluster (0.350)
- Ewing: Threshold for forming a 1x5 domino instead of a 1x3 domino (0).
- That's a lot of parameters. I made some graphics to try and illustrate what they do.

Hybrid Clustering Graphics



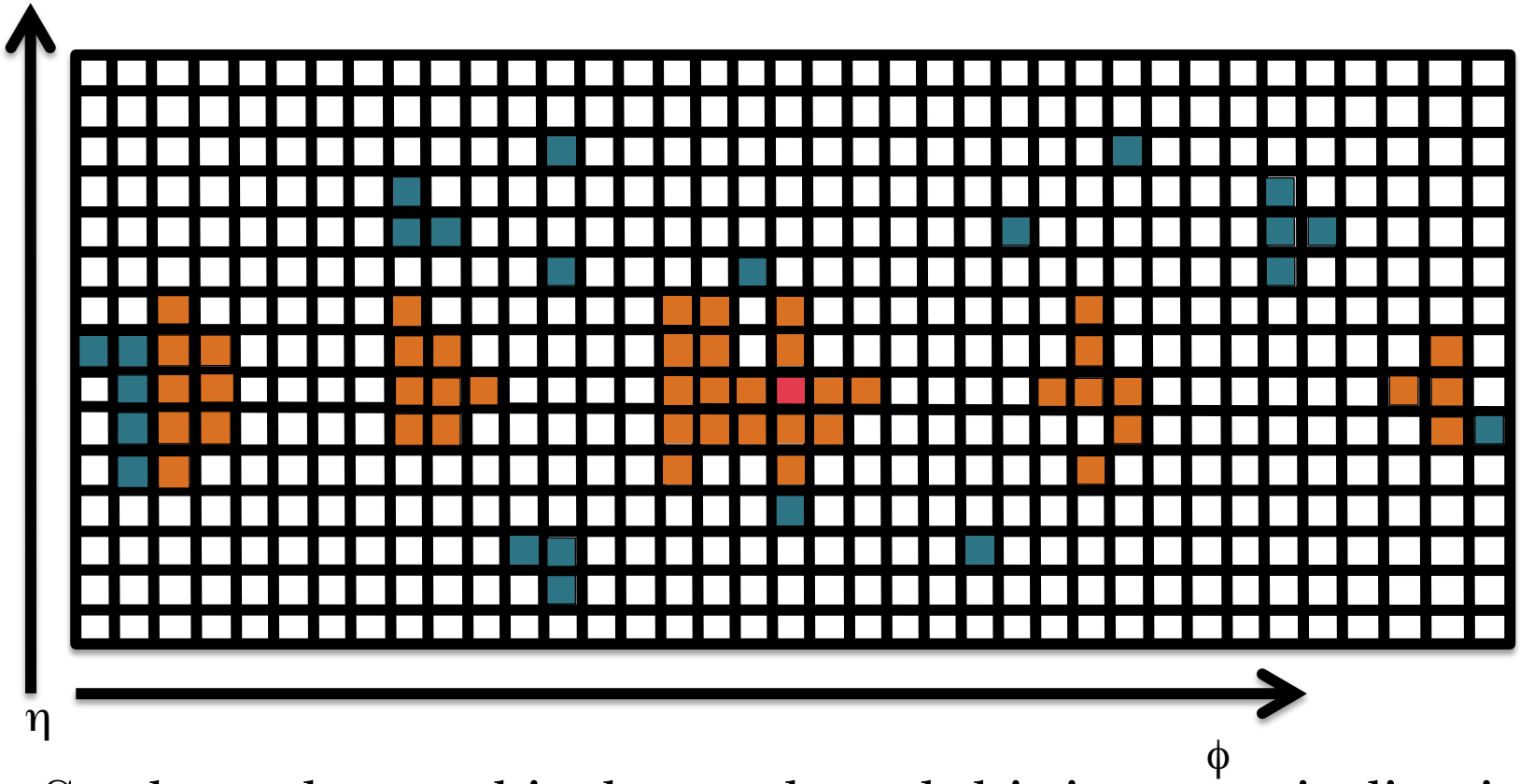
- So the red crystal is the seed, and this is pretty indicative of our current clustering settings, step=17, ethres=0.100, eseed=0.350, ewing=0.

Hybrid Clustering Graphics



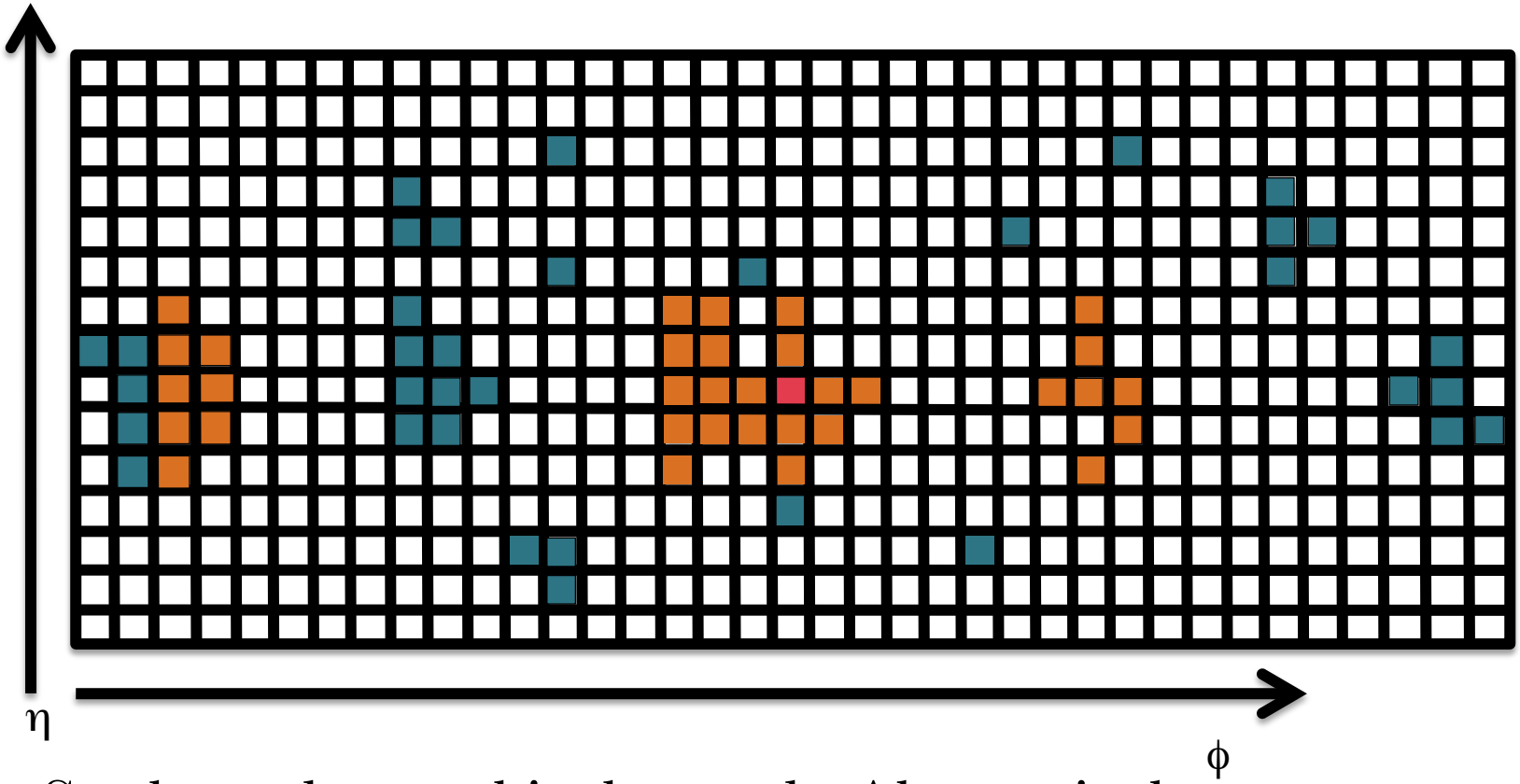
- So the red crystal is the seed. If you change step, you alter what crystals are considered. This is step=14.

Hybrid Clustering Graphics



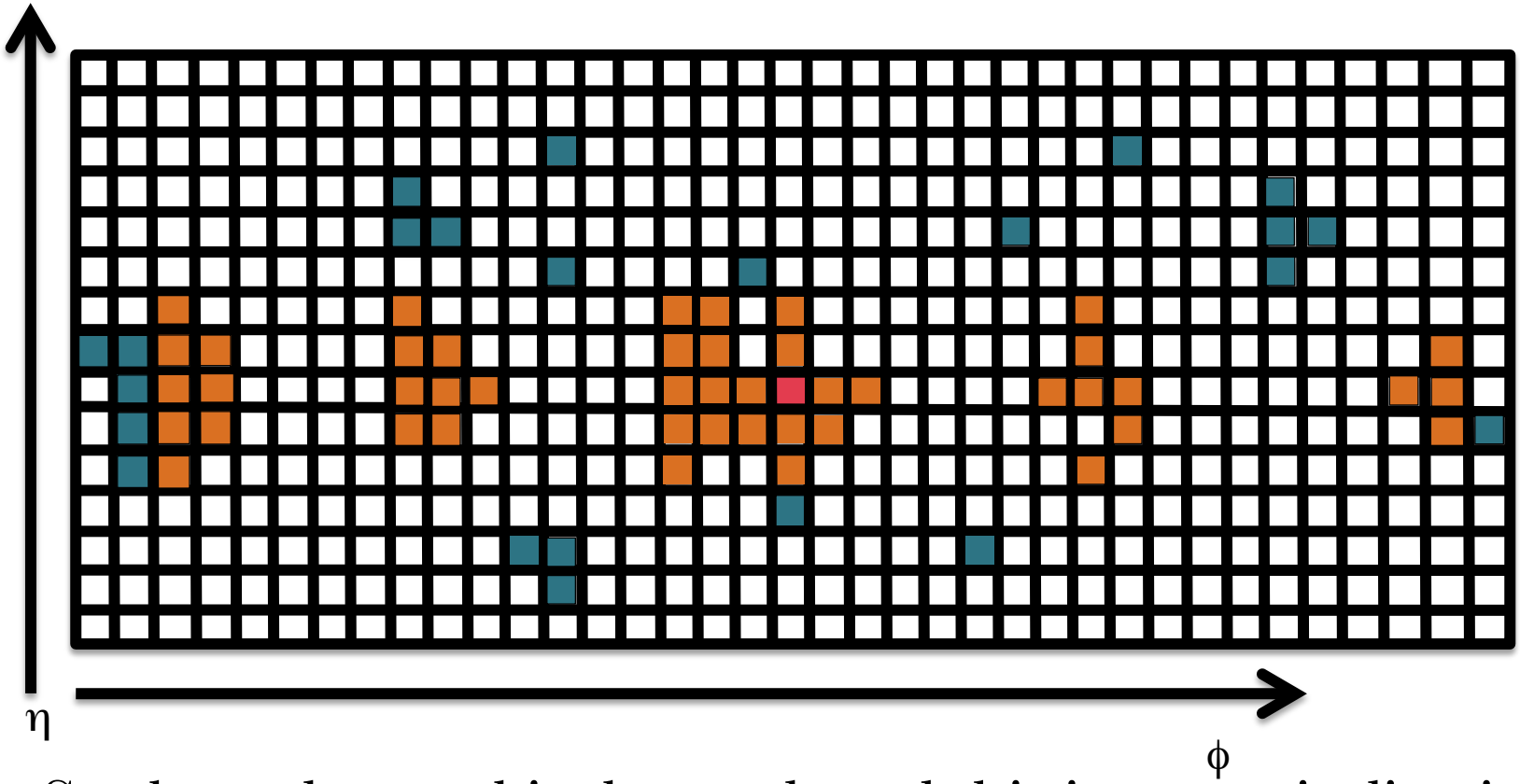
- So the red crystal is the seed, and this is pretty indicative of our current clustering settings, step=17, ethres=0.100, eseed=0.350, ewing=0.

Hybrid Clustering Graphics



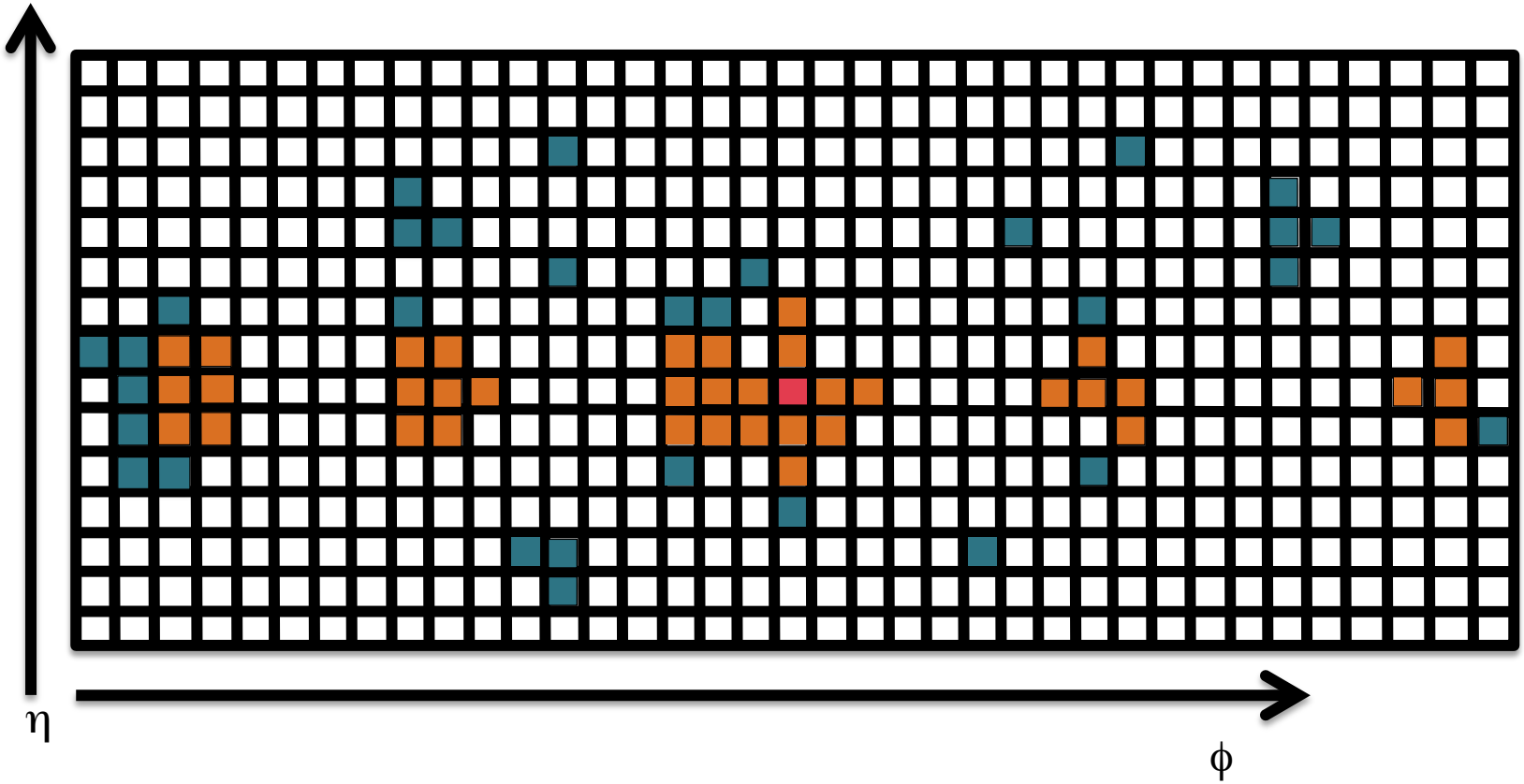
- So the red crystal is the seed. Alternatively, you can increase Eseed, and this can kill whole subclusters. This is a threshold on the local maxima.

Hybrid Clustering Graphics



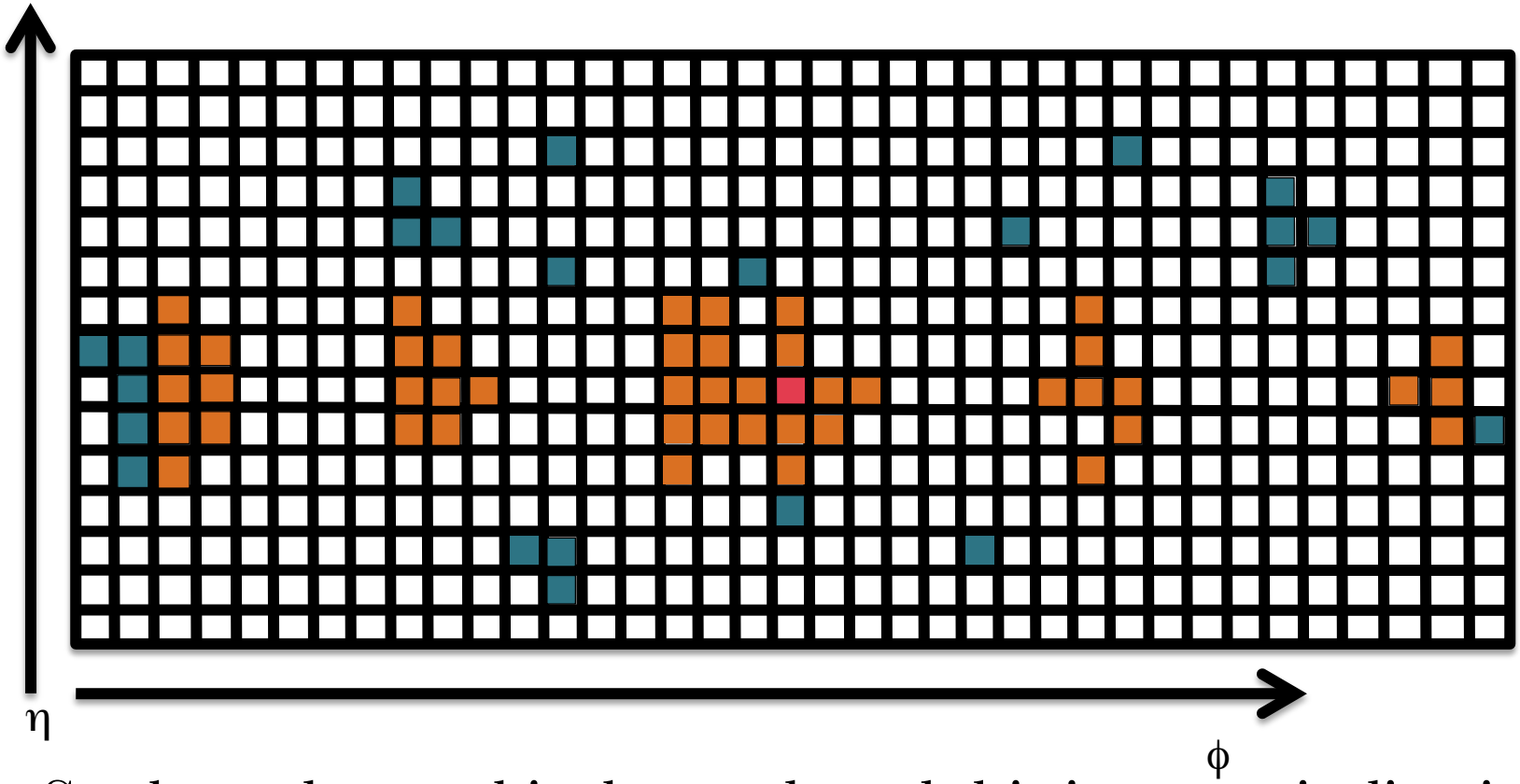
- So the red crystal is the seed, and this is pretty indicative of our current clustering settings, step=17, ethres=0.100, eseed=0.350, ewing=0.

Hybrid Clustering Graphics



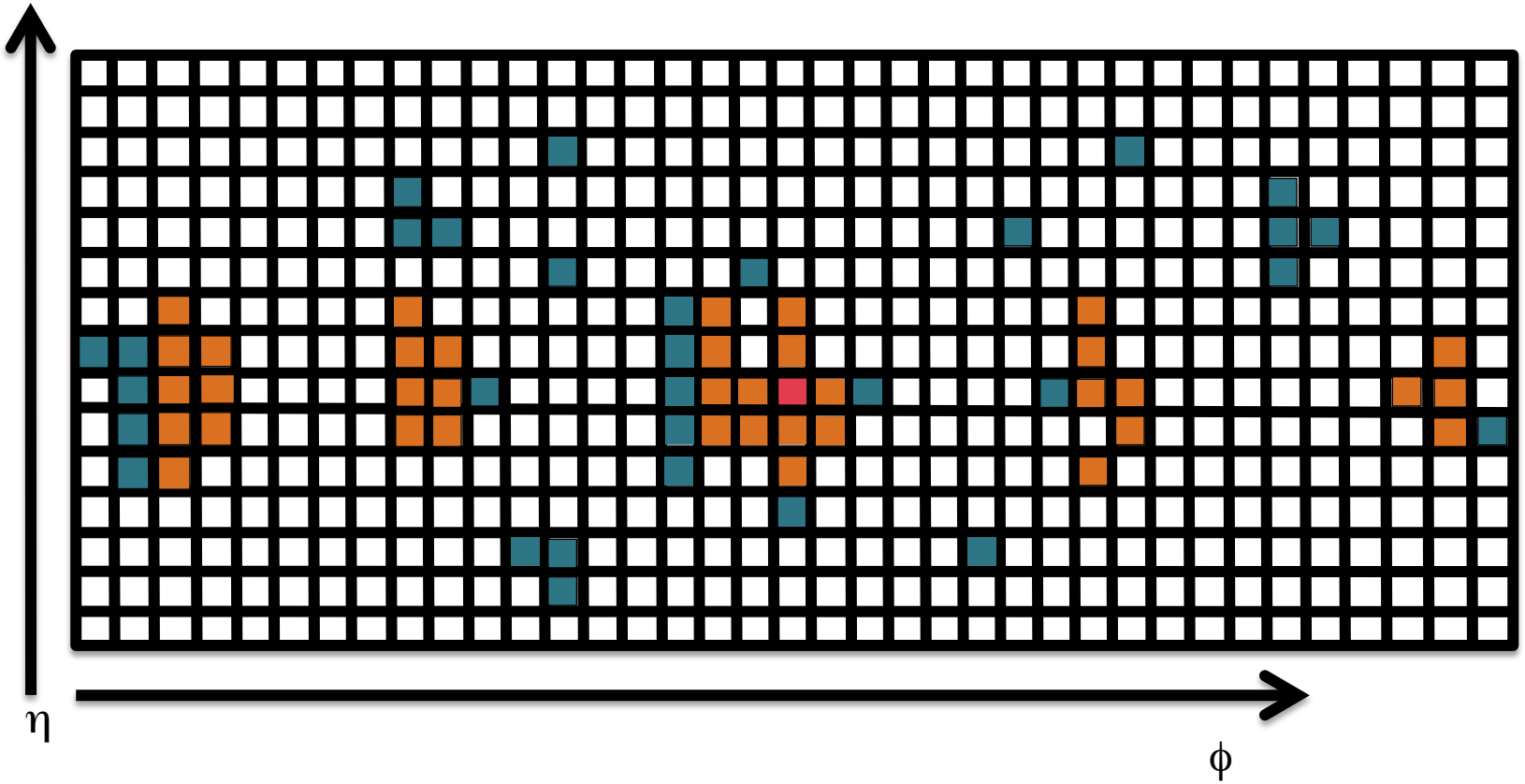
- So the red crystal is the seed. You could increase Ewing, which can round things off a bit.

Hybrid Clustering Graphics



- So the red crystal is the seed, and this is pretty indicative of our current clustering settings, step=17, ethres=0.100, eseed=0.350, ewing=0.

Hybrid Clustering Graphics

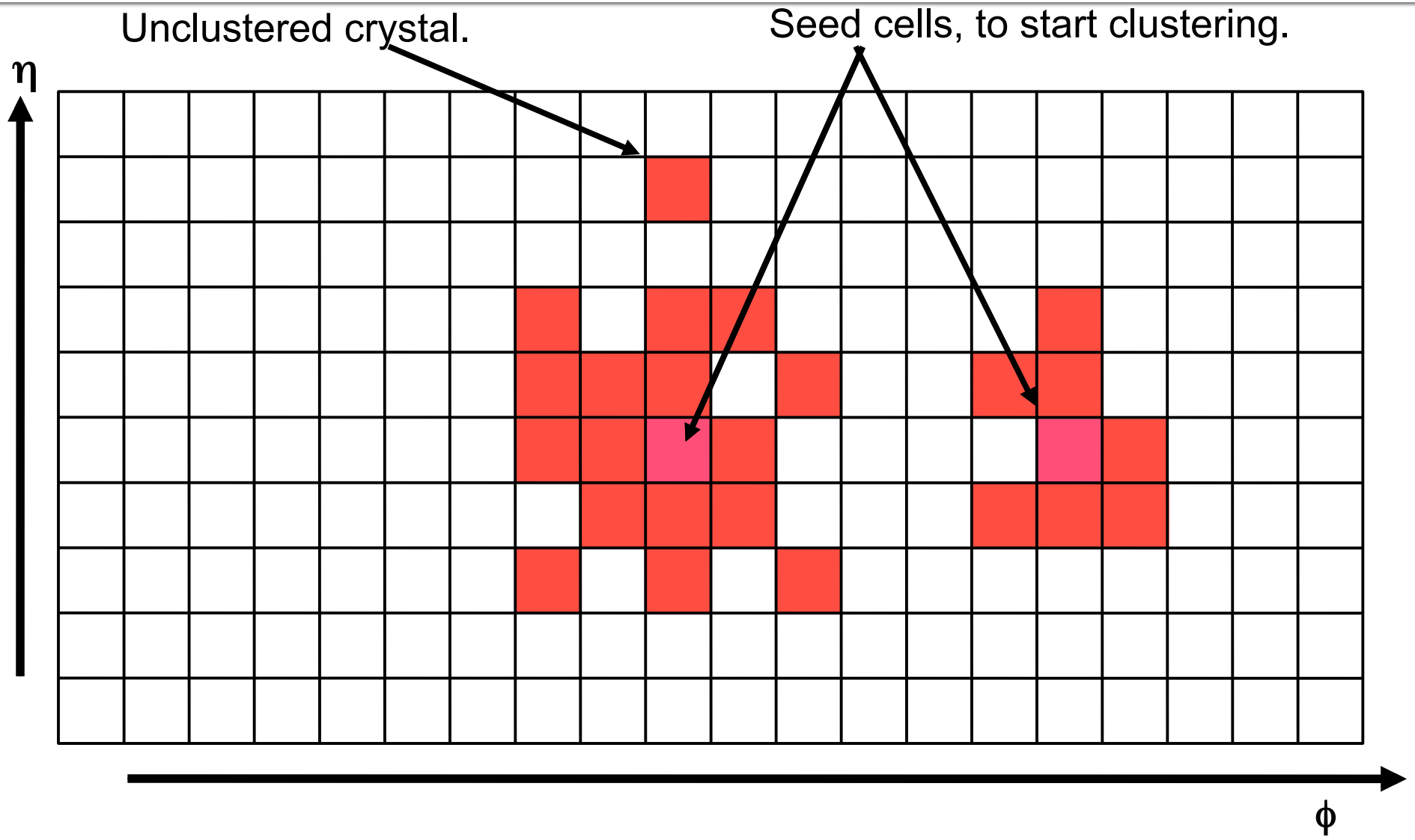


- So the red crystal is the seed. You could also change the domino threshold itself (Ethres).

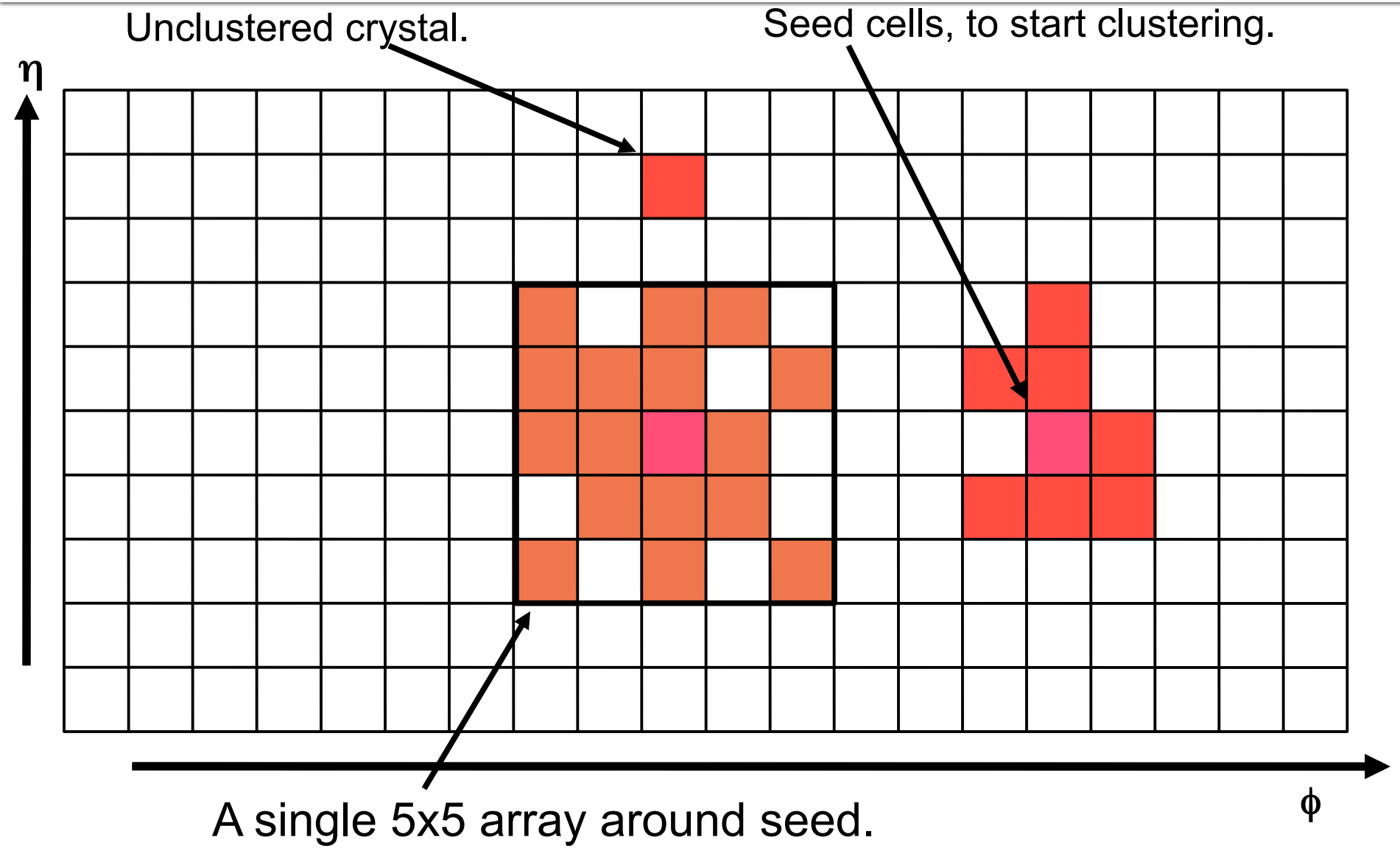
The "multi5x5" algorithm

- The "multi5x5" algorithm natively creates smaller clusters first, and then groups them together into superclusters based on their proximity in η/ϕ.
- Besides the range considered for grouping clusters together, and the thresholds, there are really very few parameters to speak of, besides the threshold to create a cluster.
- Seeds in this context are required to be local maxima of the crystals that share faces, and the 8 crystals that surround a seed are forbidden from further seeding. The outer right may, but only if they are local maxima.
- The seed threshold is 0.18, and is applied in E and then E_T. This is for BasicClusters

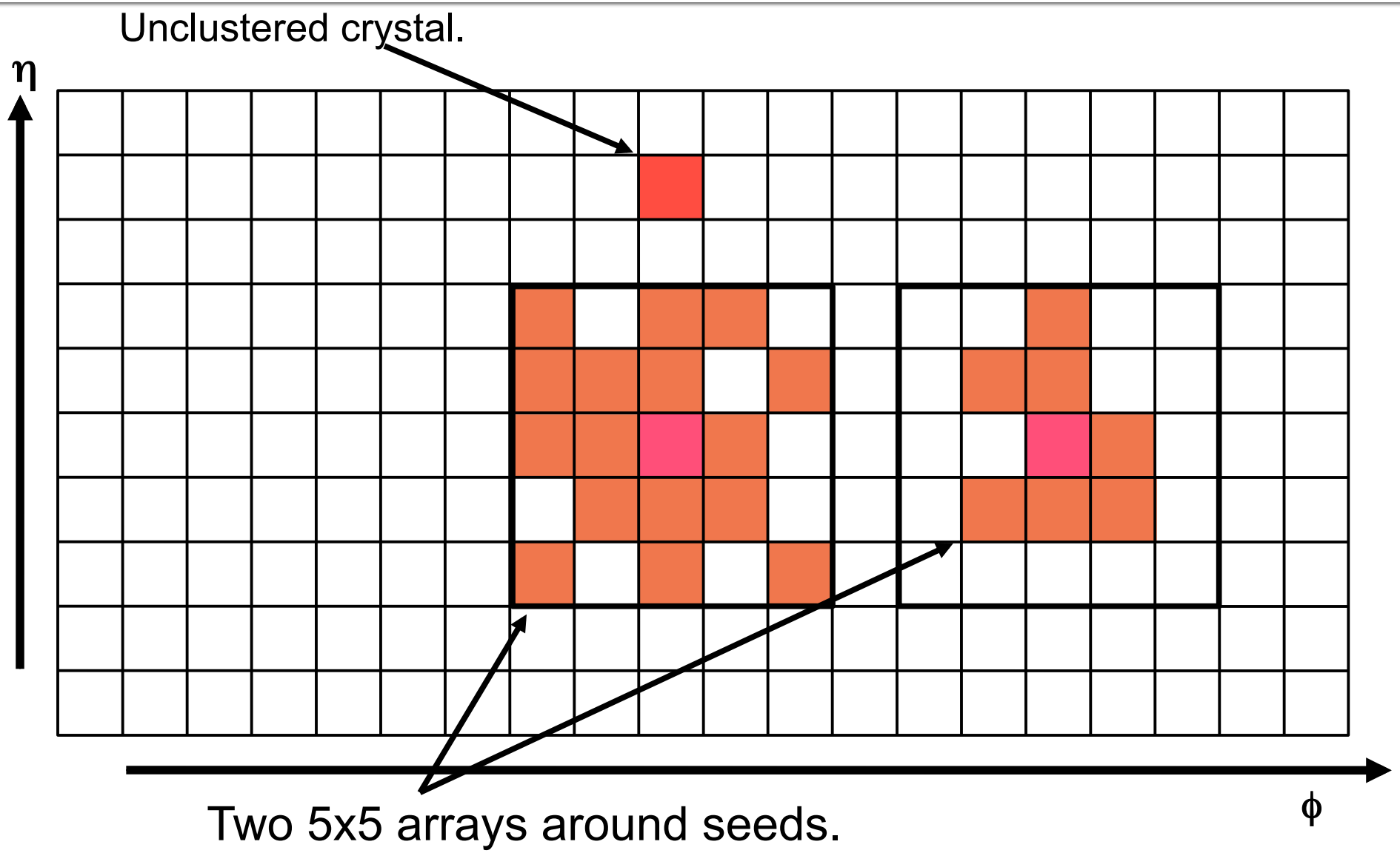
EE Clustering (same for e/γ):



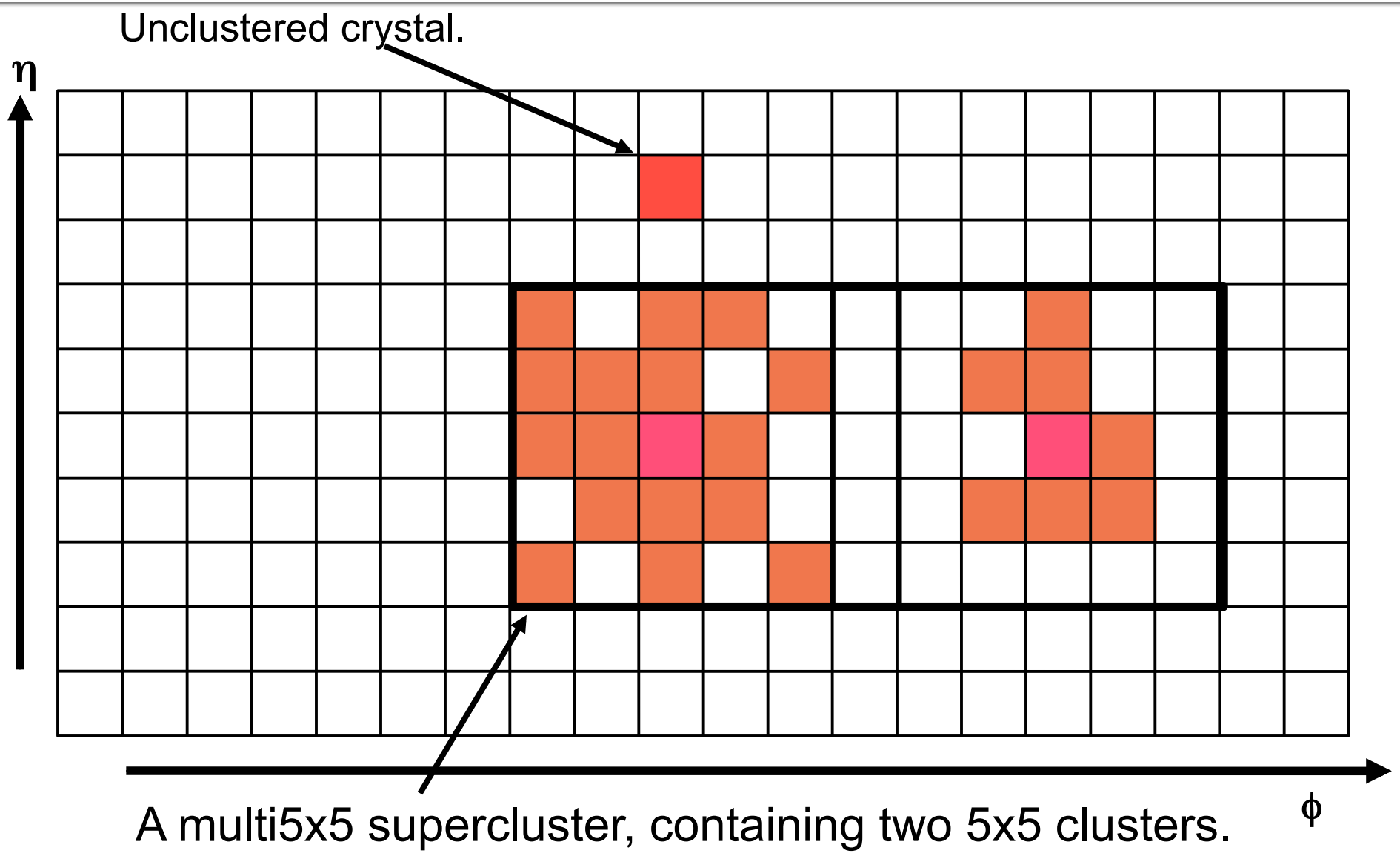
EE Clustering (same for e/γ):



EE Clustering (same for e/γ):



EE Clustering (same for e/γ):



Multi5x5 Superclustering

- There really aren't that many parameters to the "multi5x5" algorithm, so there's not much tuning that can be done. However, how you choose to group the clusters into superclusters is configurable.
  - seedTransverseEnergyThreshold = 1
  - endcapPhiSearchRoad=0.6
  - endcapEtaSearchRoad=0.14
- To Supercluster then, we start from an energy ordered list of clusters, and if they fall within the window of η/ϕ defined by the search roads, they get grouped together.

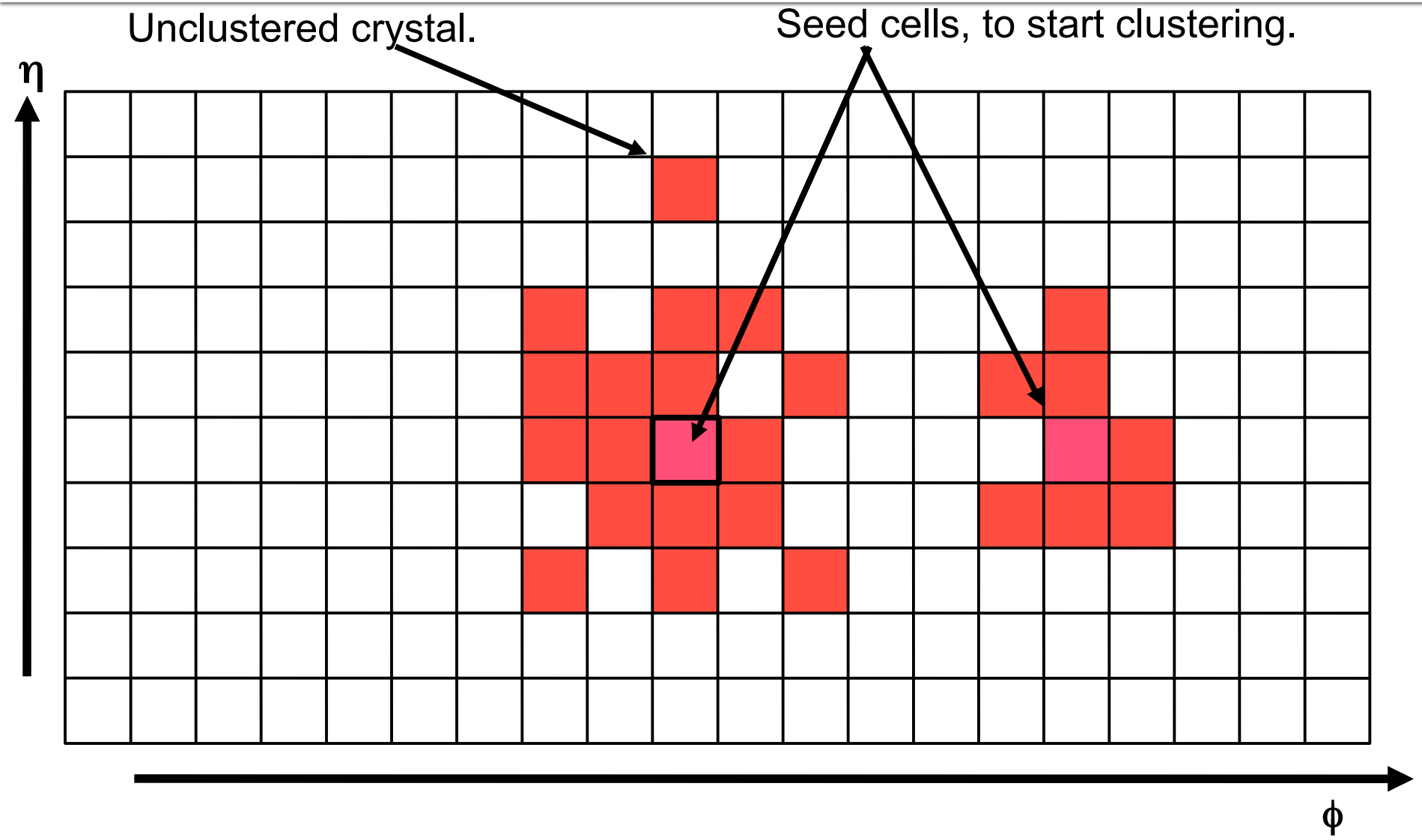
The Particle Flow algorithm

- Recognize that this is a very different beast.
  - The Pflow clustering algorithm runs on each layer of the calorimeter with a slightly different configuration, but mainly the same algo choices.
  - Only clusters are made here, and unlike the other algorithms, clusters can share energy. This is complicated to show simply.
  - In more detail, after each seed has been formed (and seeds must be the highest energy crystal in a 3x3), the local energy is clustered into a "topological" cluster.
  - This cluster is then split between the constituent seeds, unless there is only one, in which case, all the energy belongs to that seed.

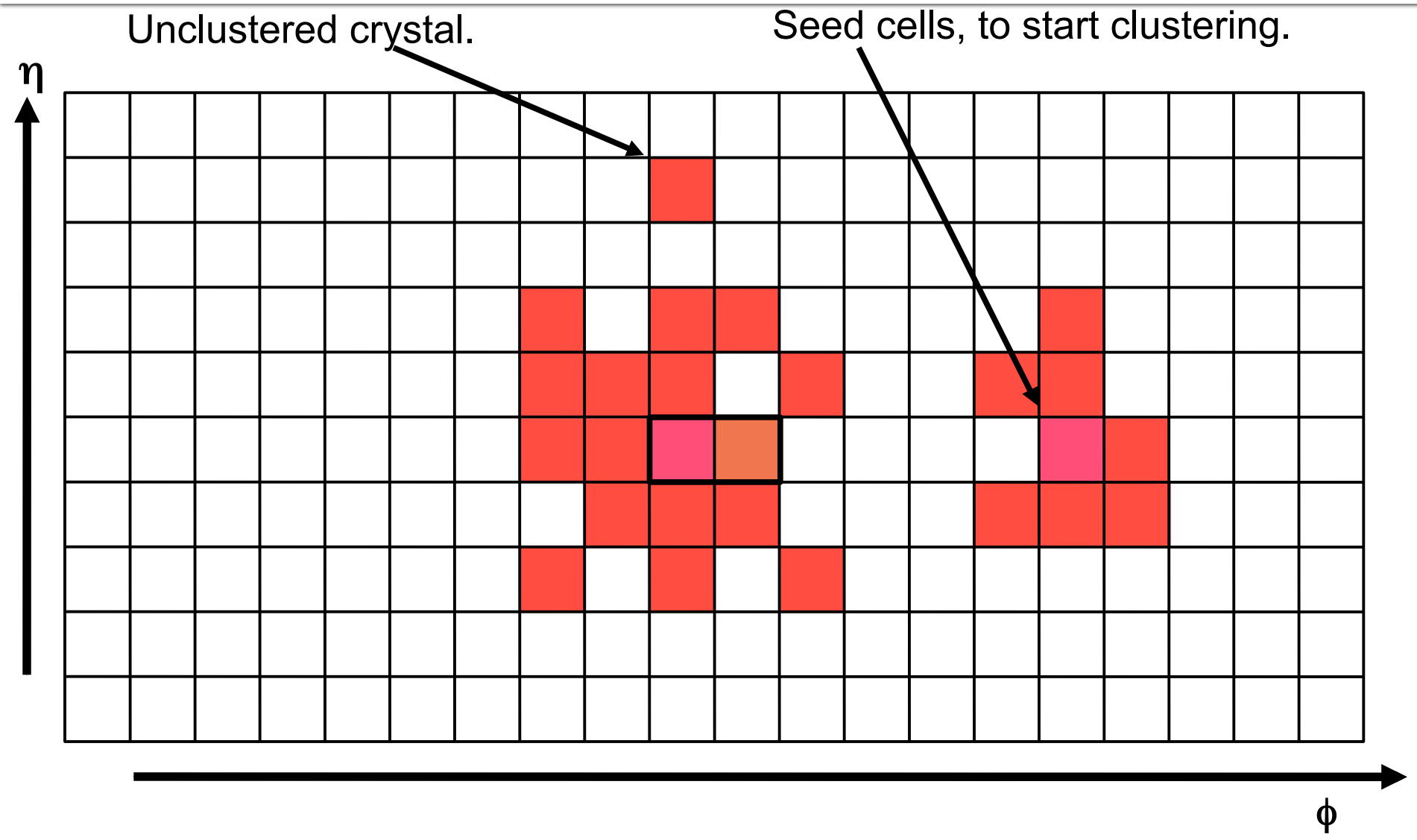
The Particle Flow algorithm

- If you recognize the jargon, the process by which crystals are added to a topological cluster is a recursive nearest-neighbor algorithm.
  - E.g. I take a step in a given direction and check my neighbor…
    - Then I take a step from the neighbor, to the next neighbor…
      - Then I take a step from the neighbor's neighbor, to the next neighbor…
        - And so on and so forth, until I run out of neighbors (and proceed to the next step, working back up the chain) or go insane…
- The illustration will make this slightly easier, but the entire contiguous area of energy will be added to the topological cluster, terminating where the crystals drop below threshold. In this case, the crystals that are white are either negative energy, or below the threshold for adding them.
  - The threshold for a seed is 0.23.
  - The threshold for adding crystals in the ECAL is 0.08.

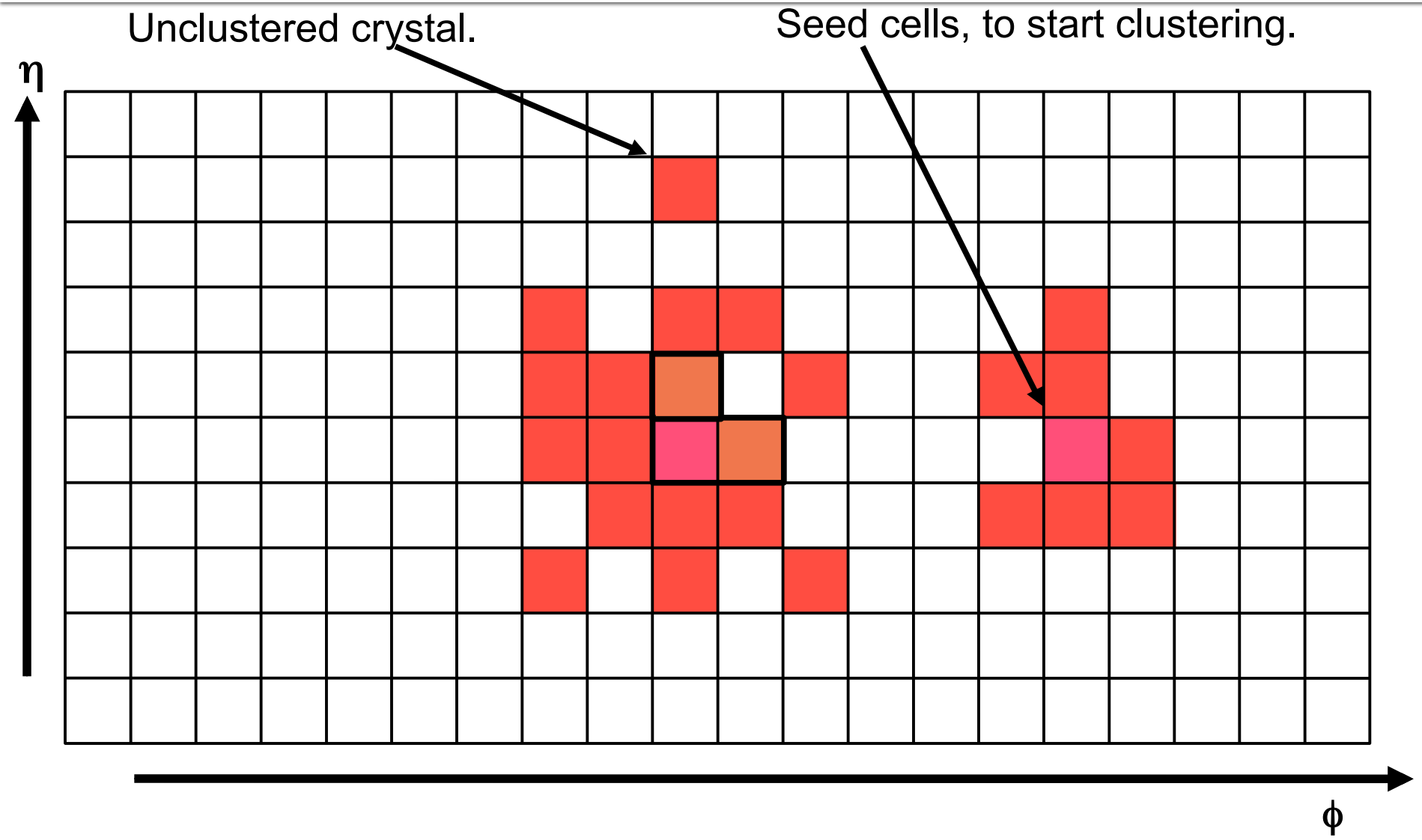
PFlow Clustering:



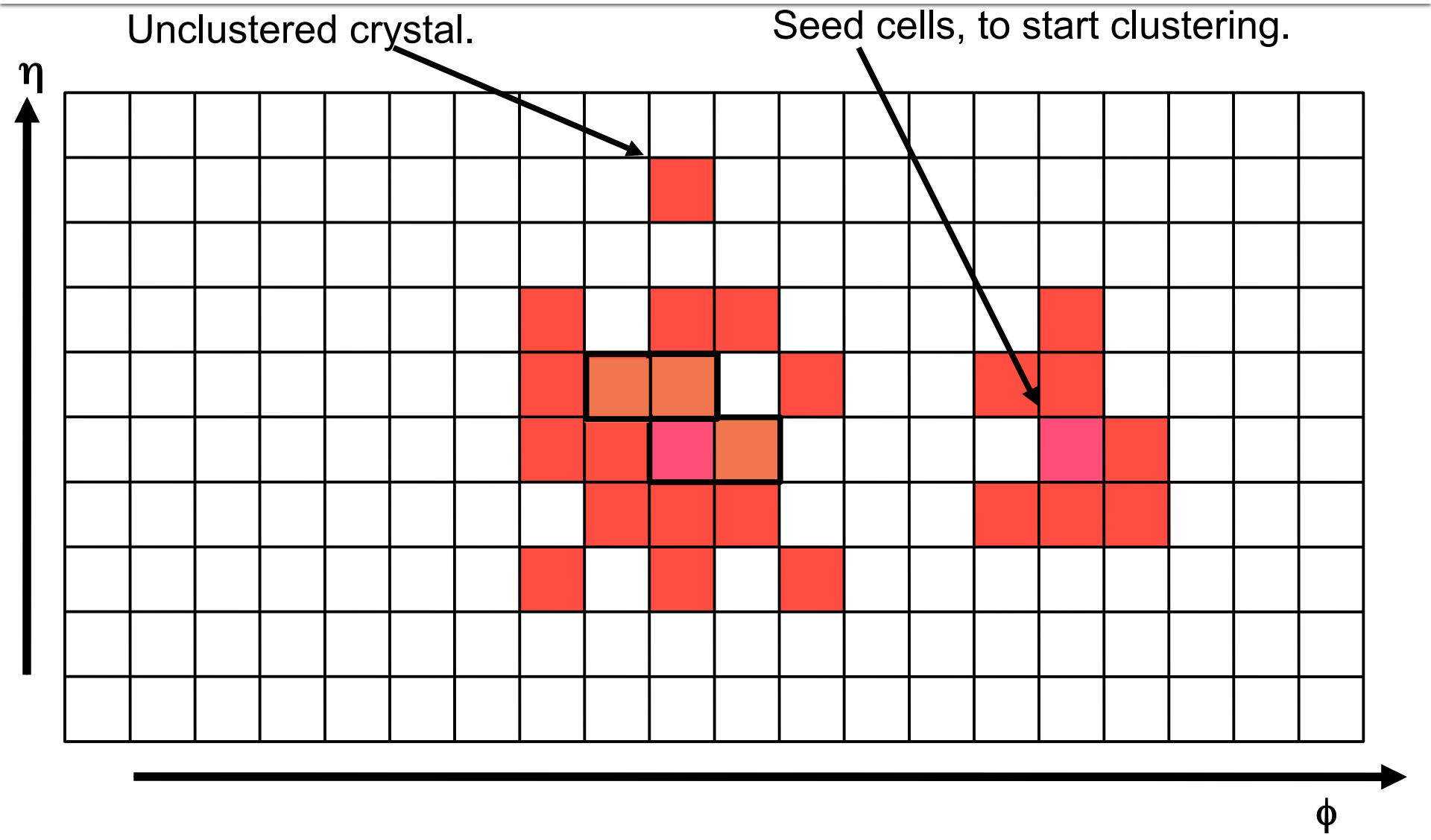
PFlow Clustering:



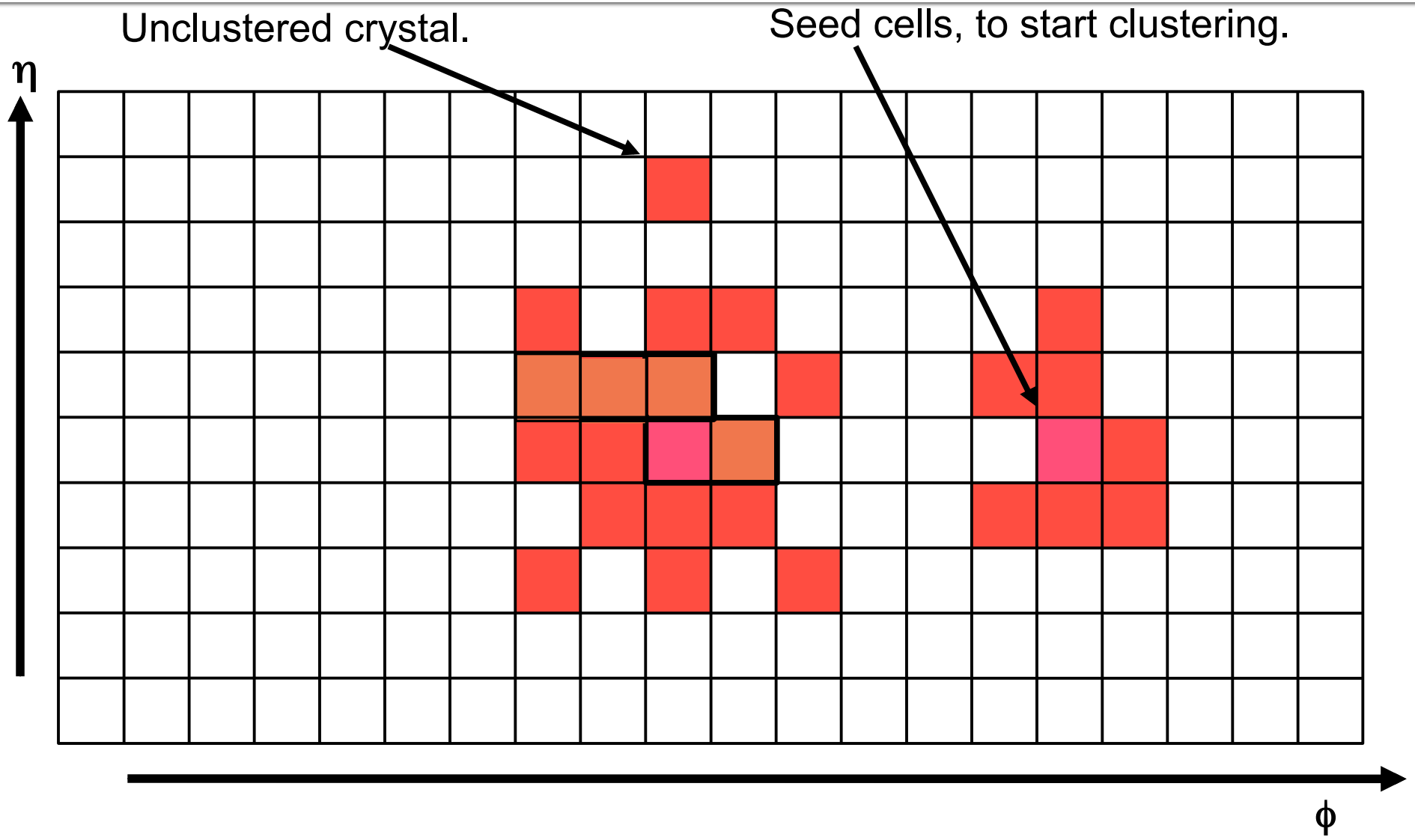
PFlow Clustering:



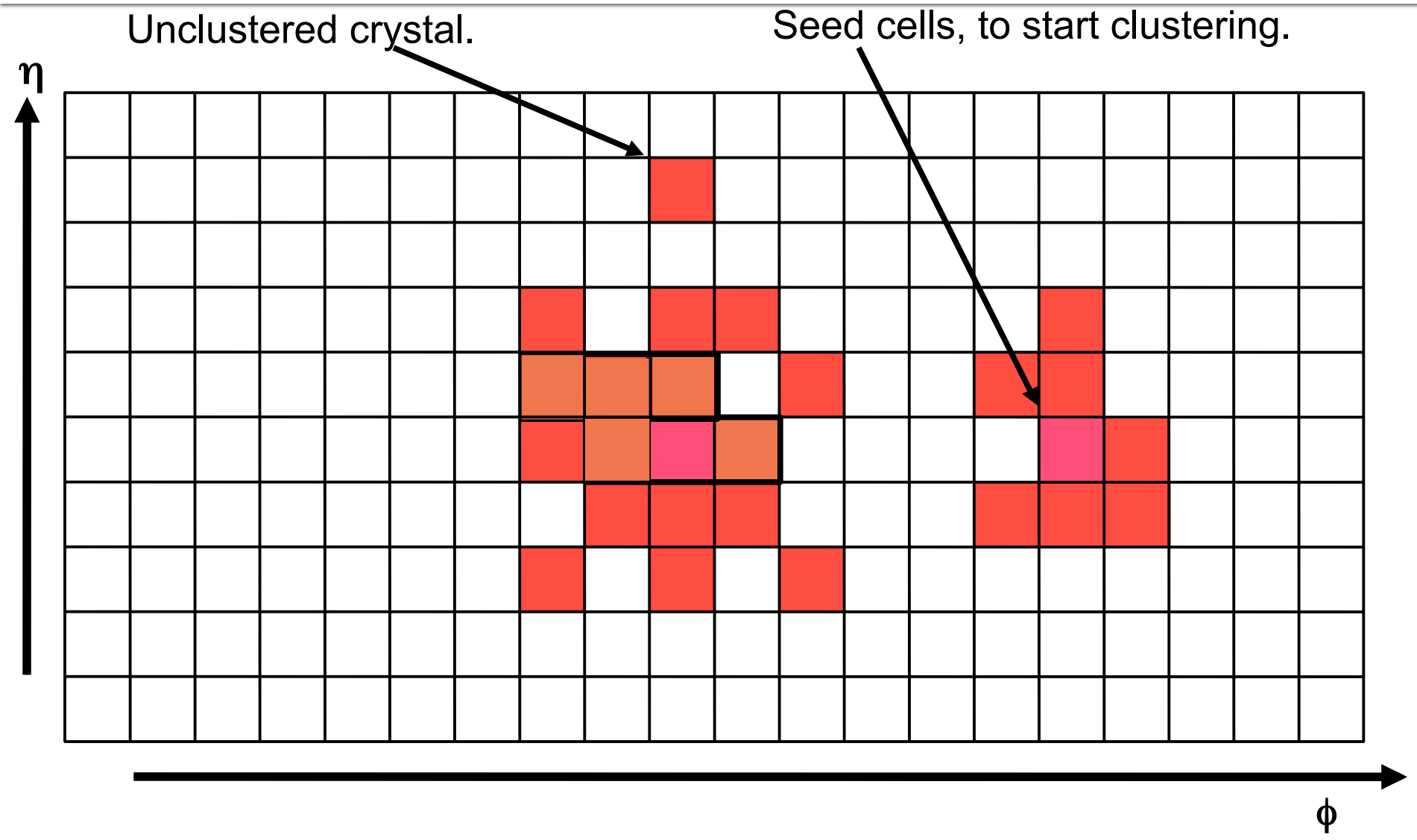
PFlow Clustering:



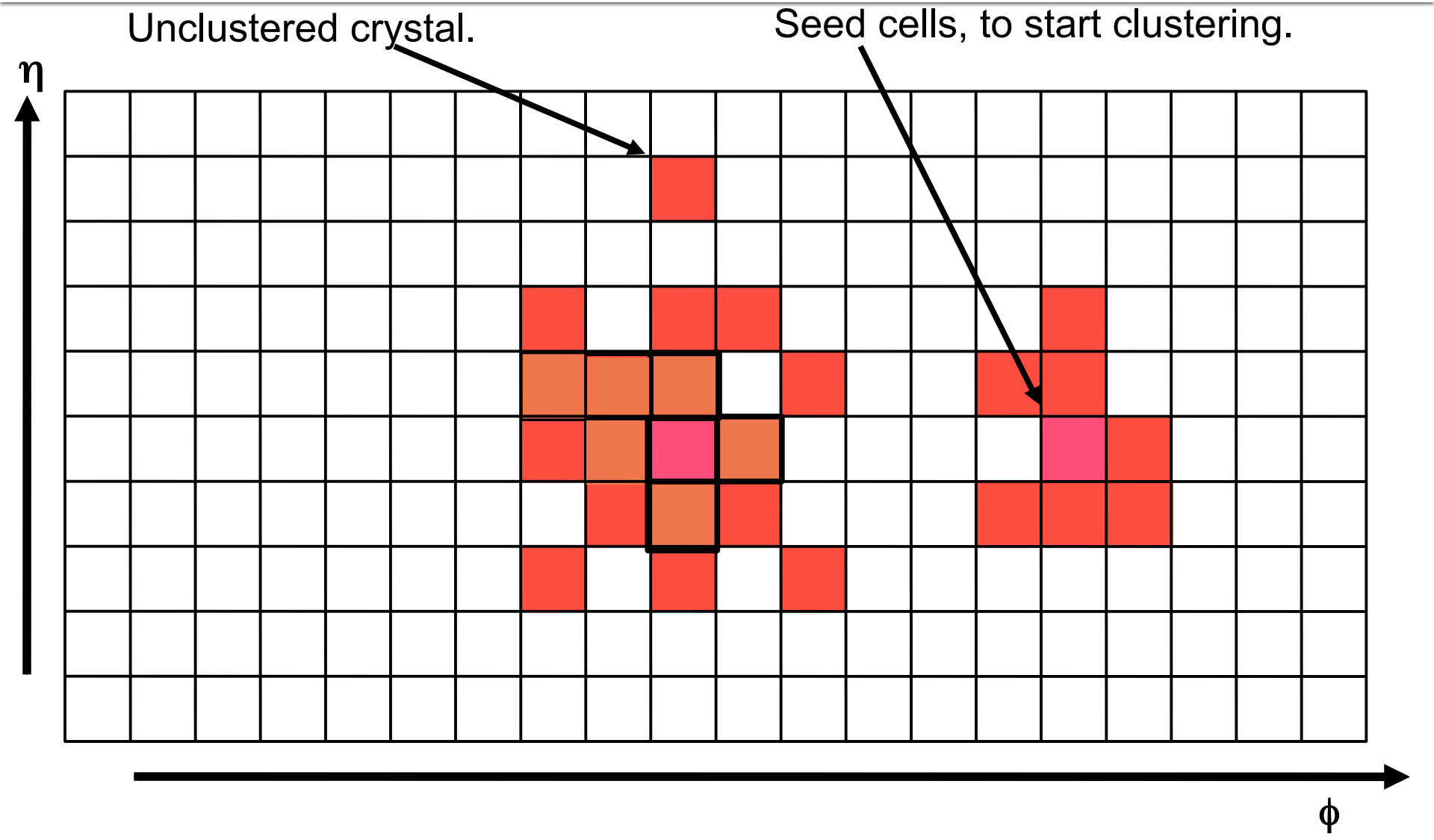
PFlow Clustering:



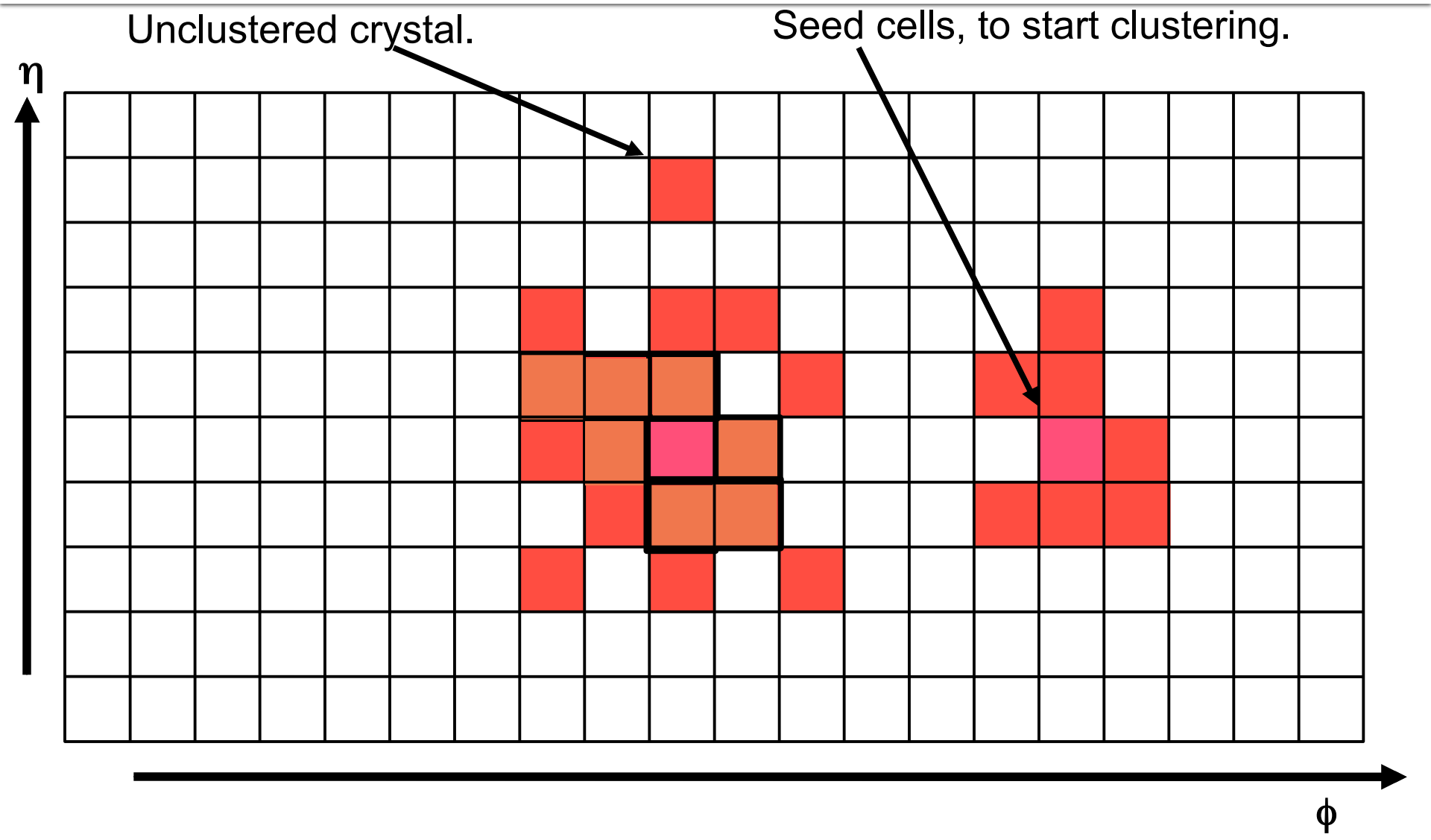
PFlow Clustering:



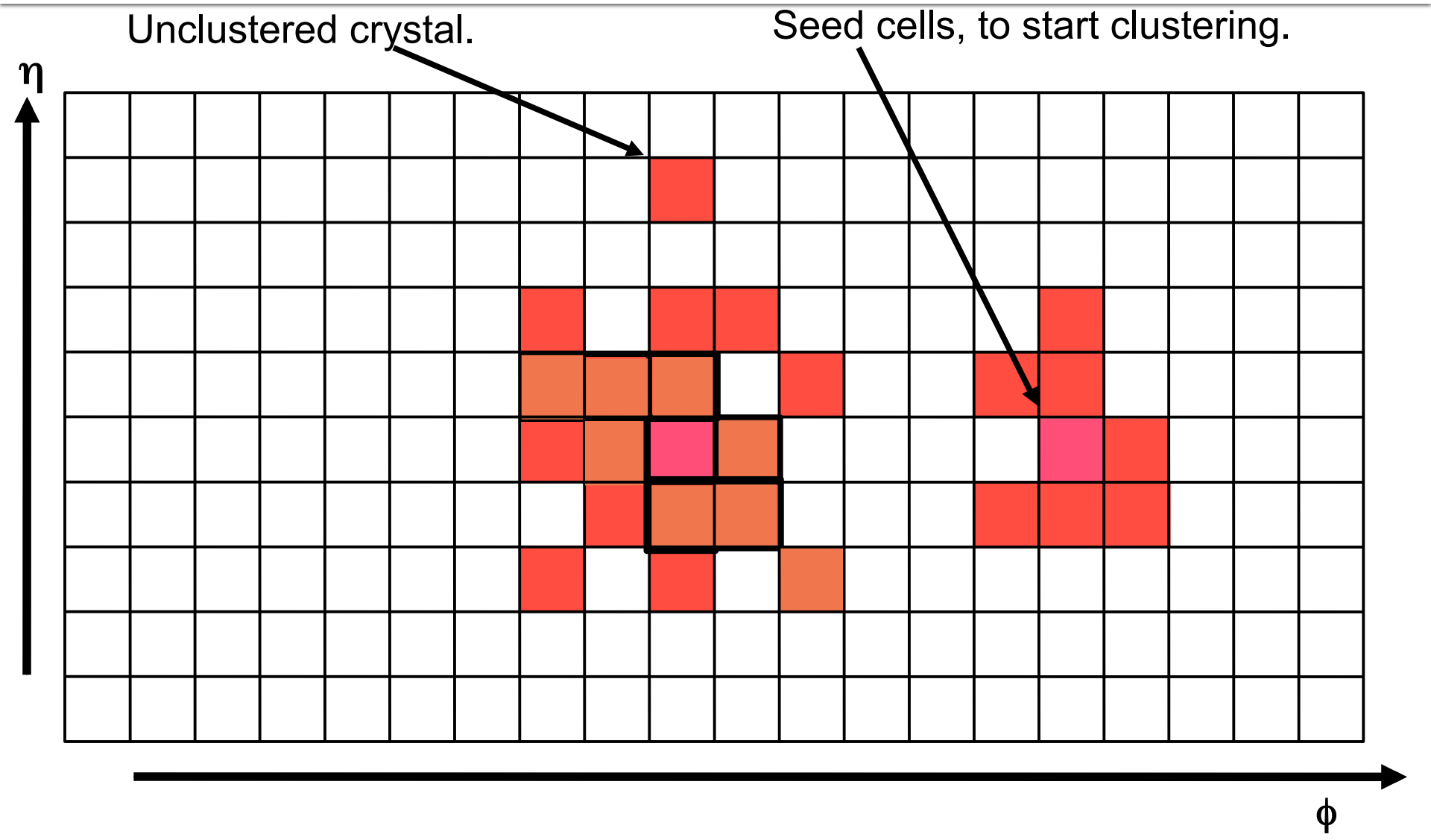
PFlow Clustering:



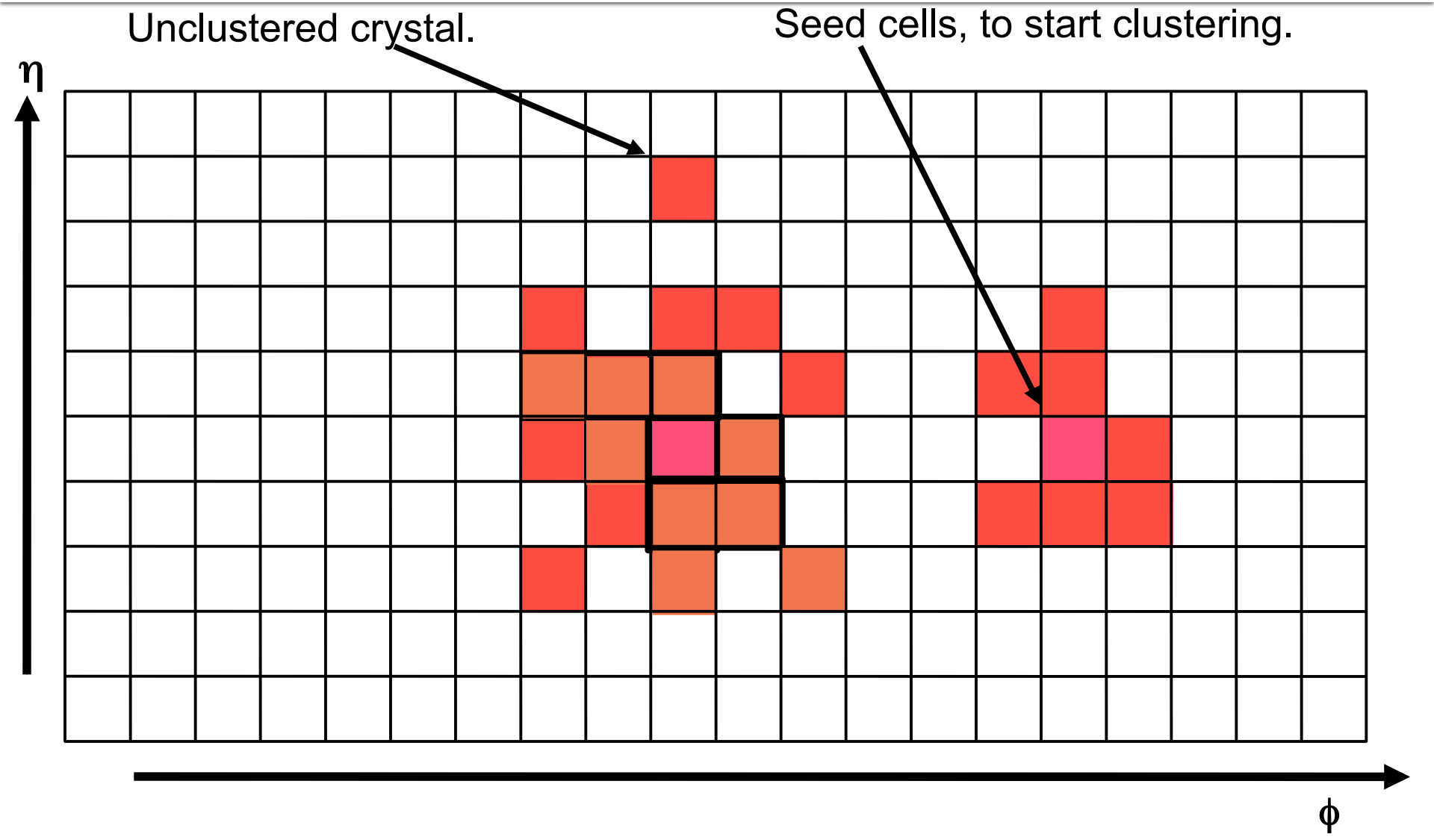
PFlow Clustering:



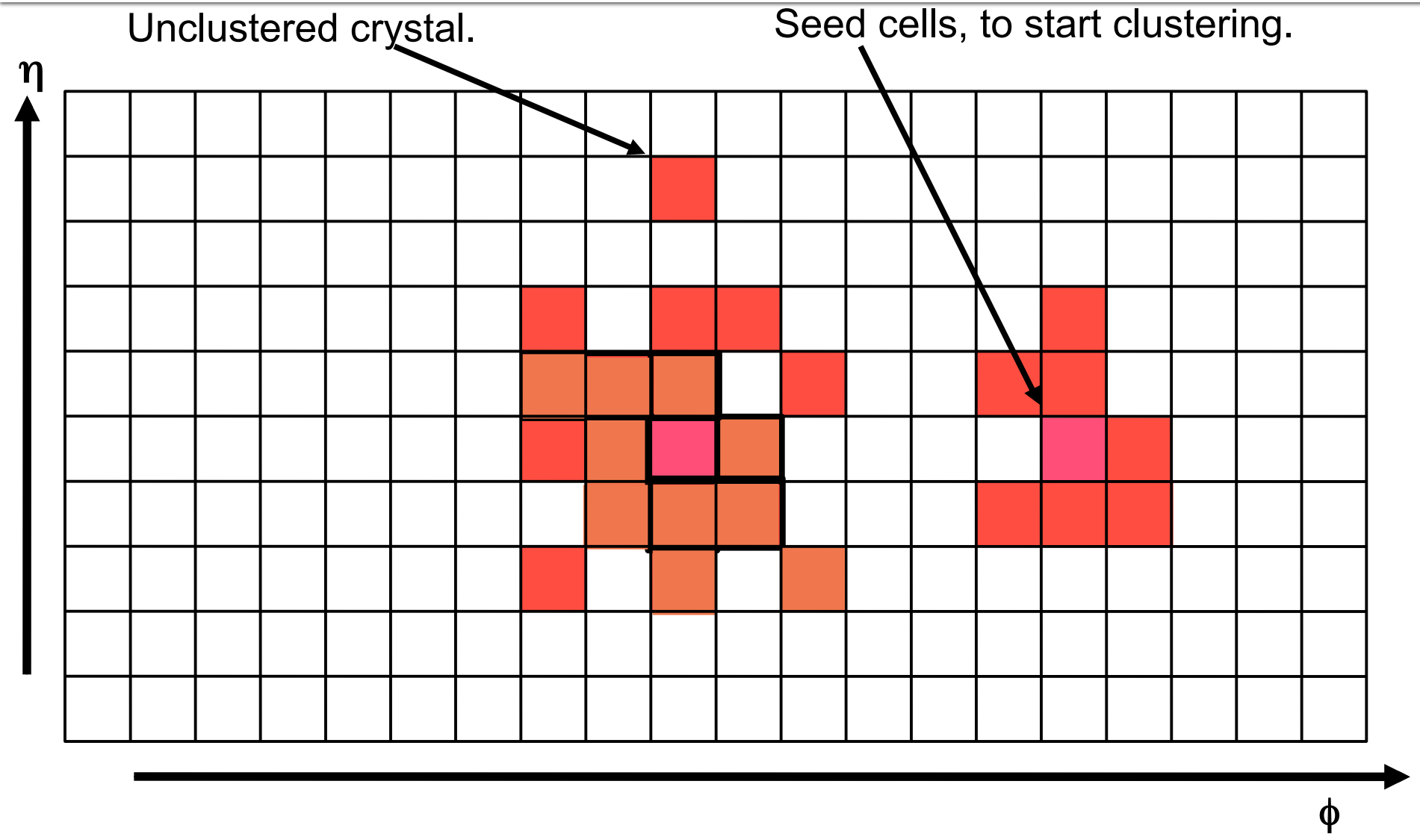
PFlow Clustering:



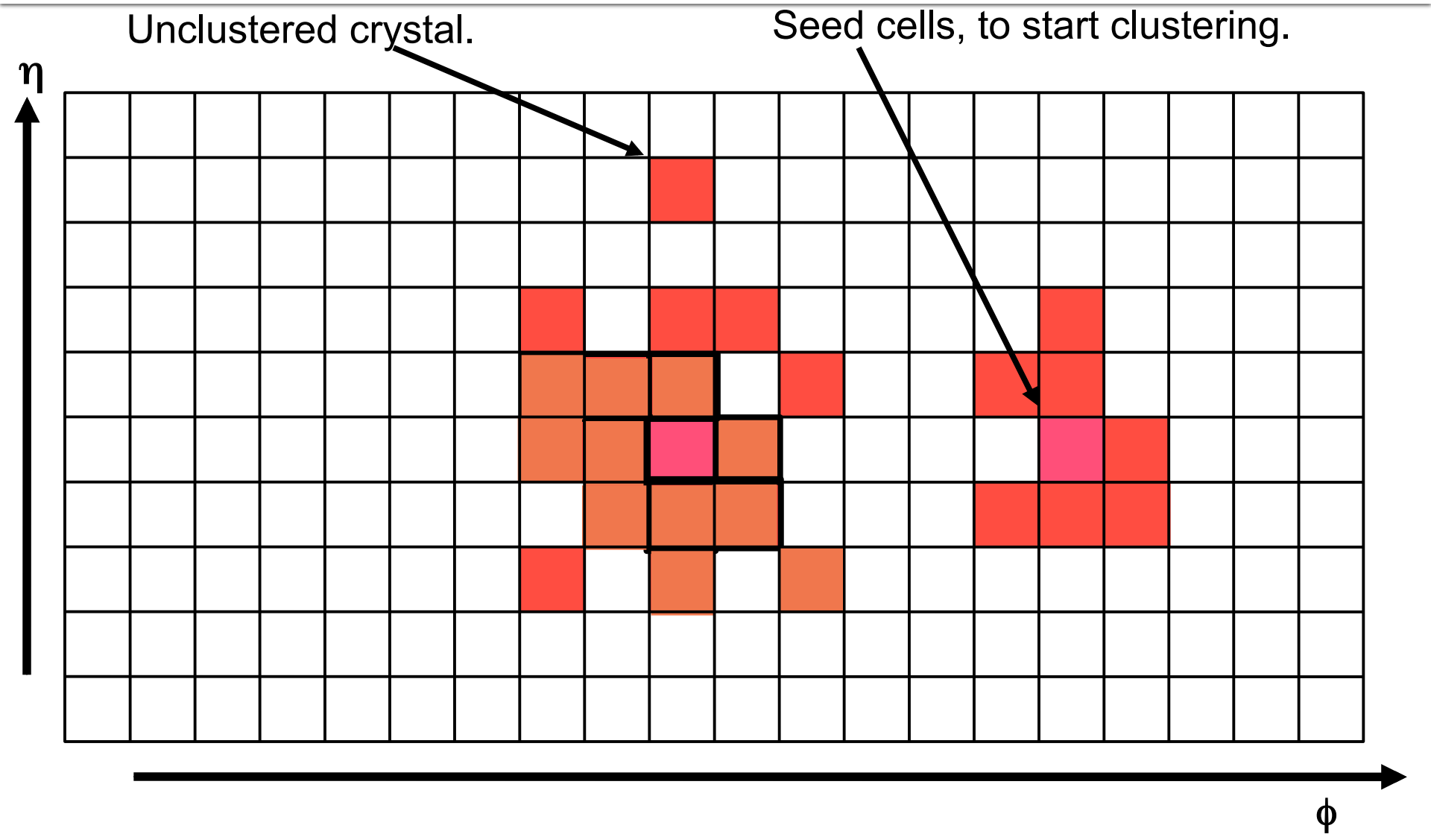
PFlow Clustering:



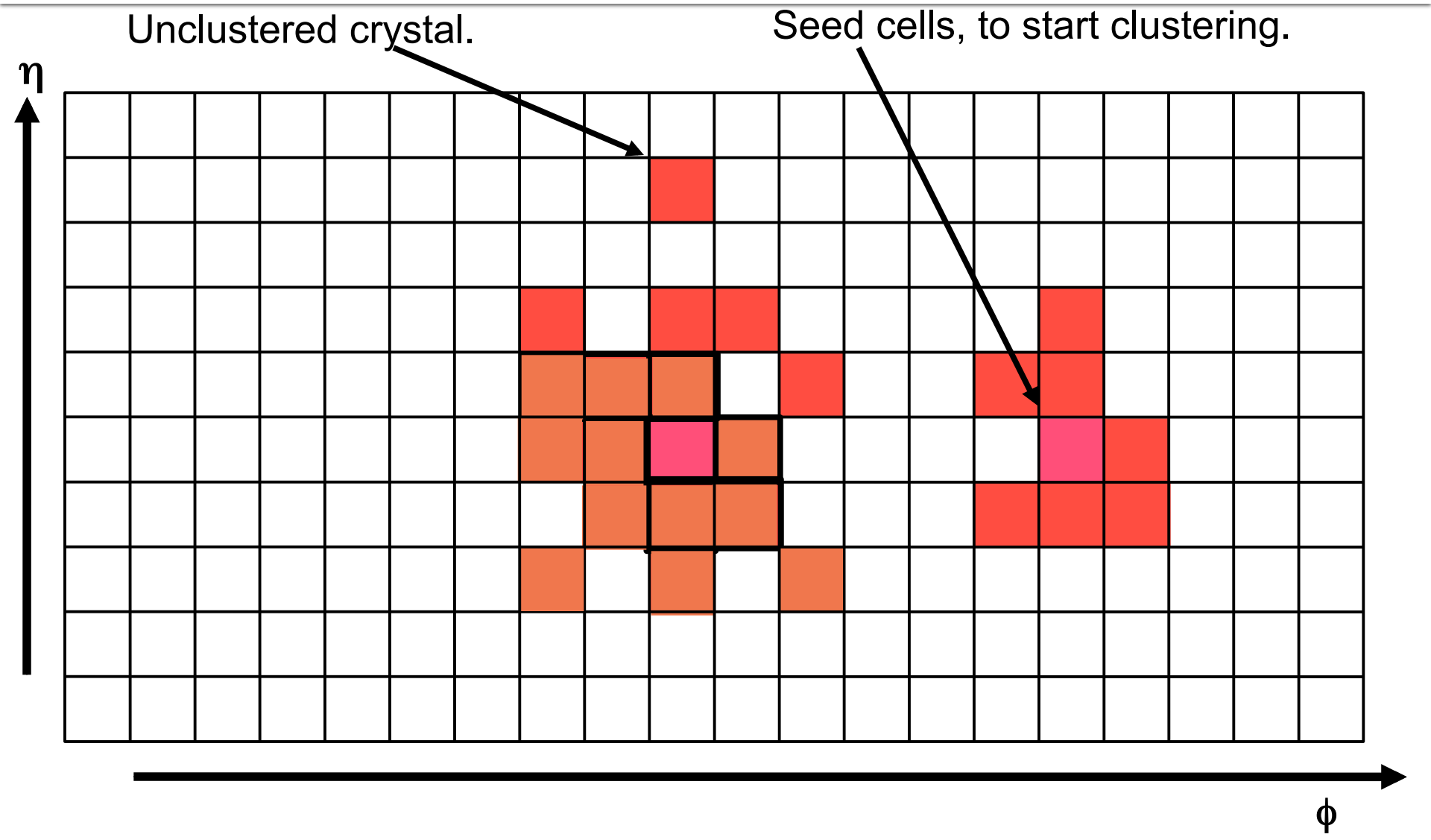
PFlow Clustering:



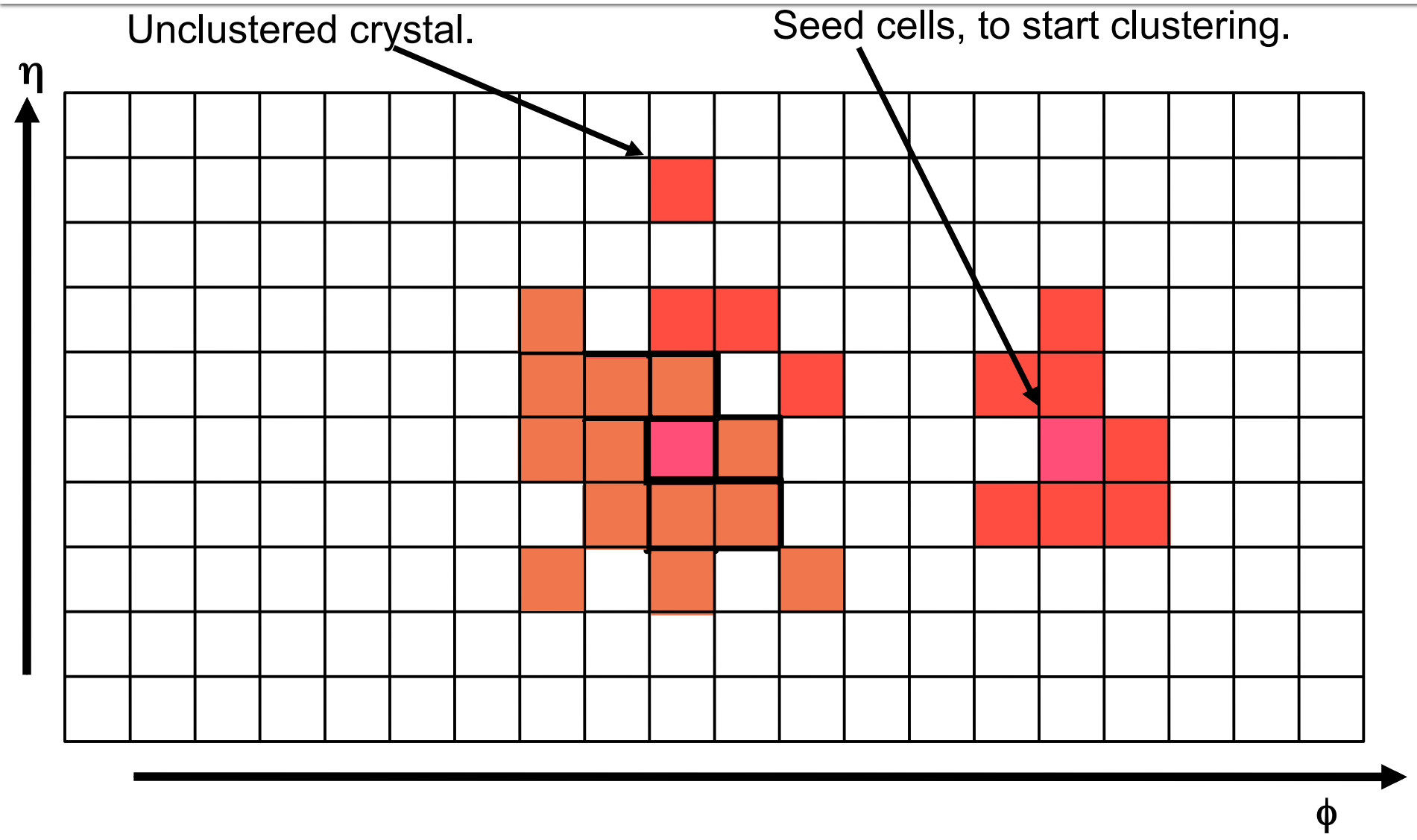
PFlow Clustering:



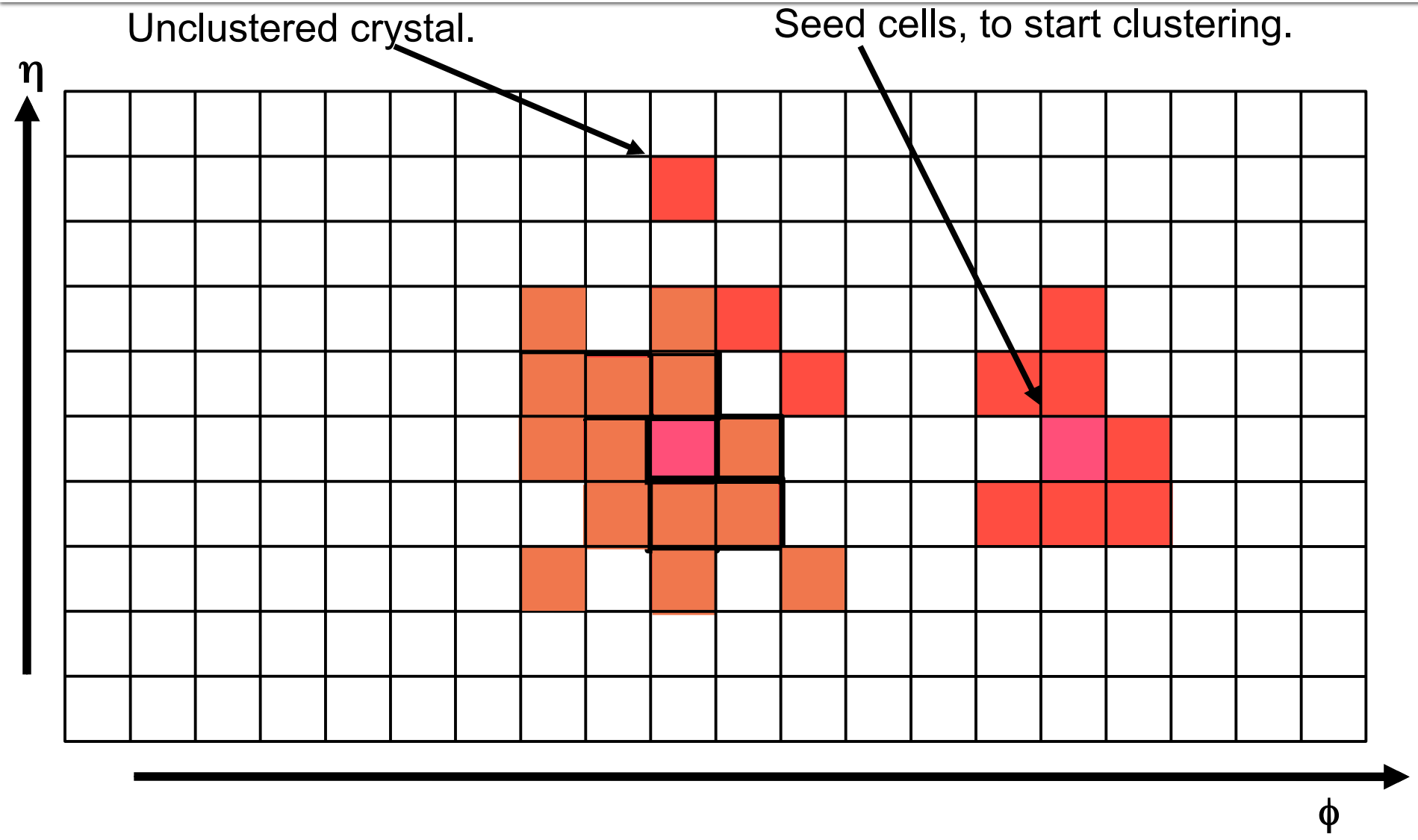
PFlow Clustering:



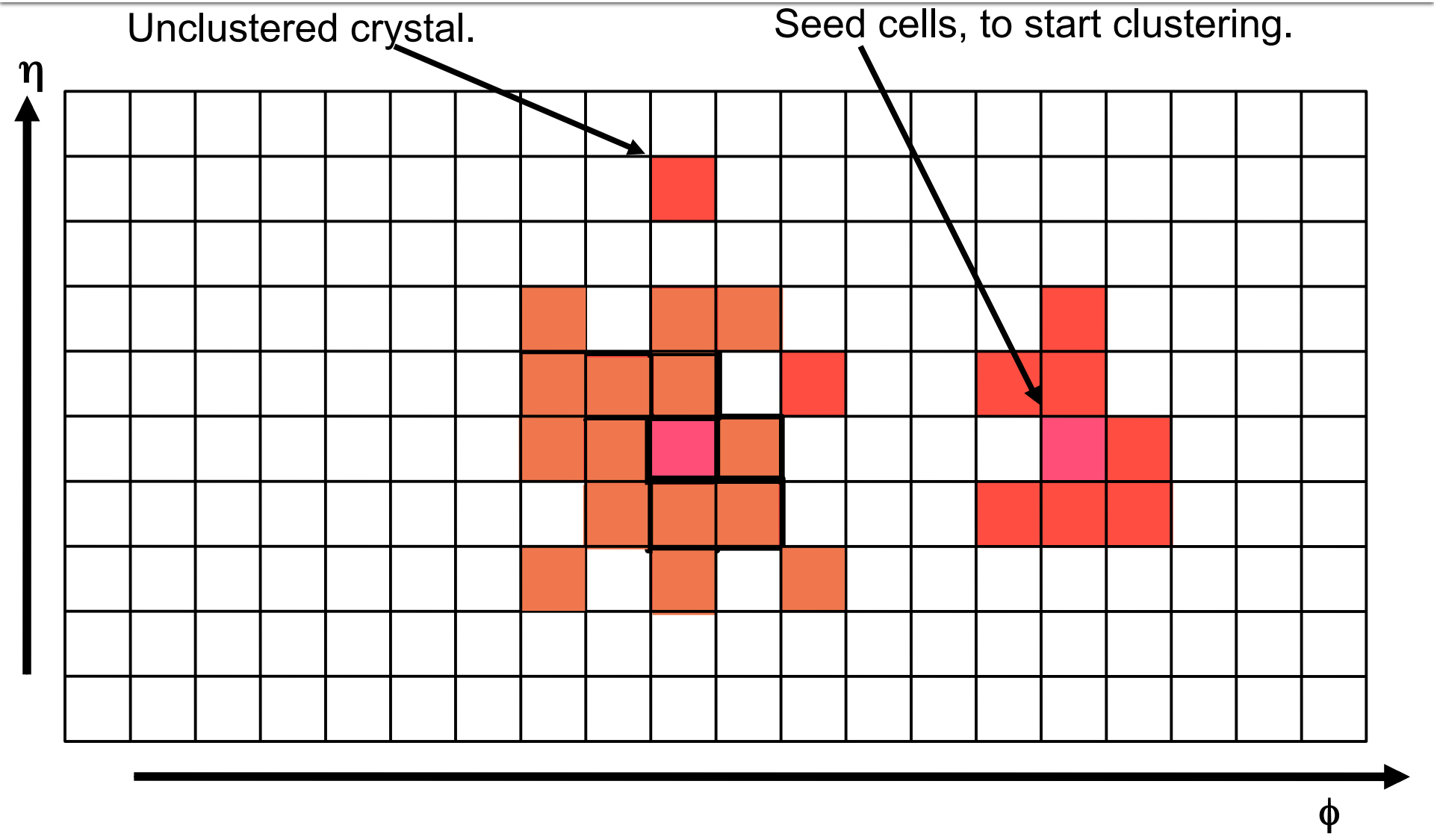
PFlow Clustering:



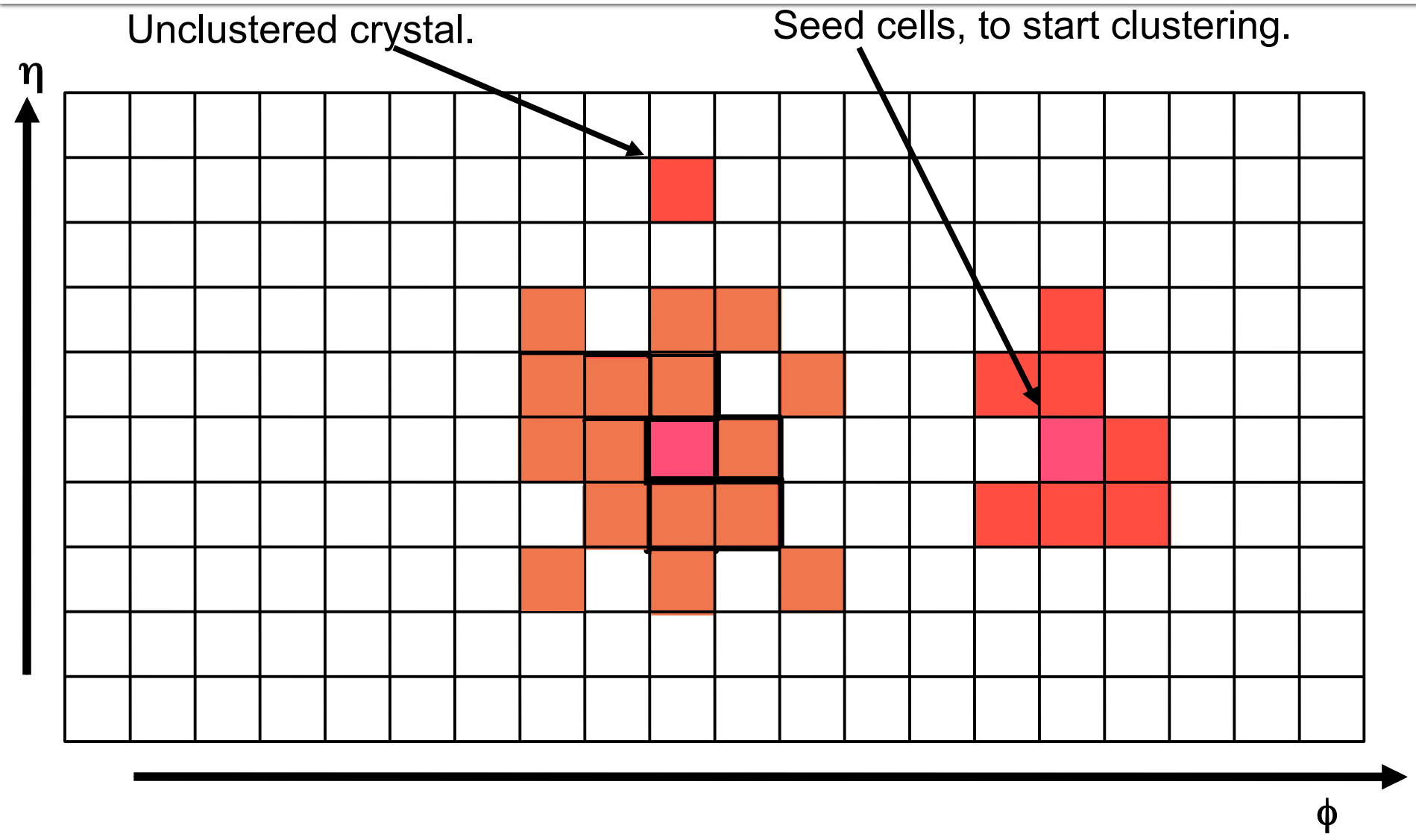
PFlow Clustering:



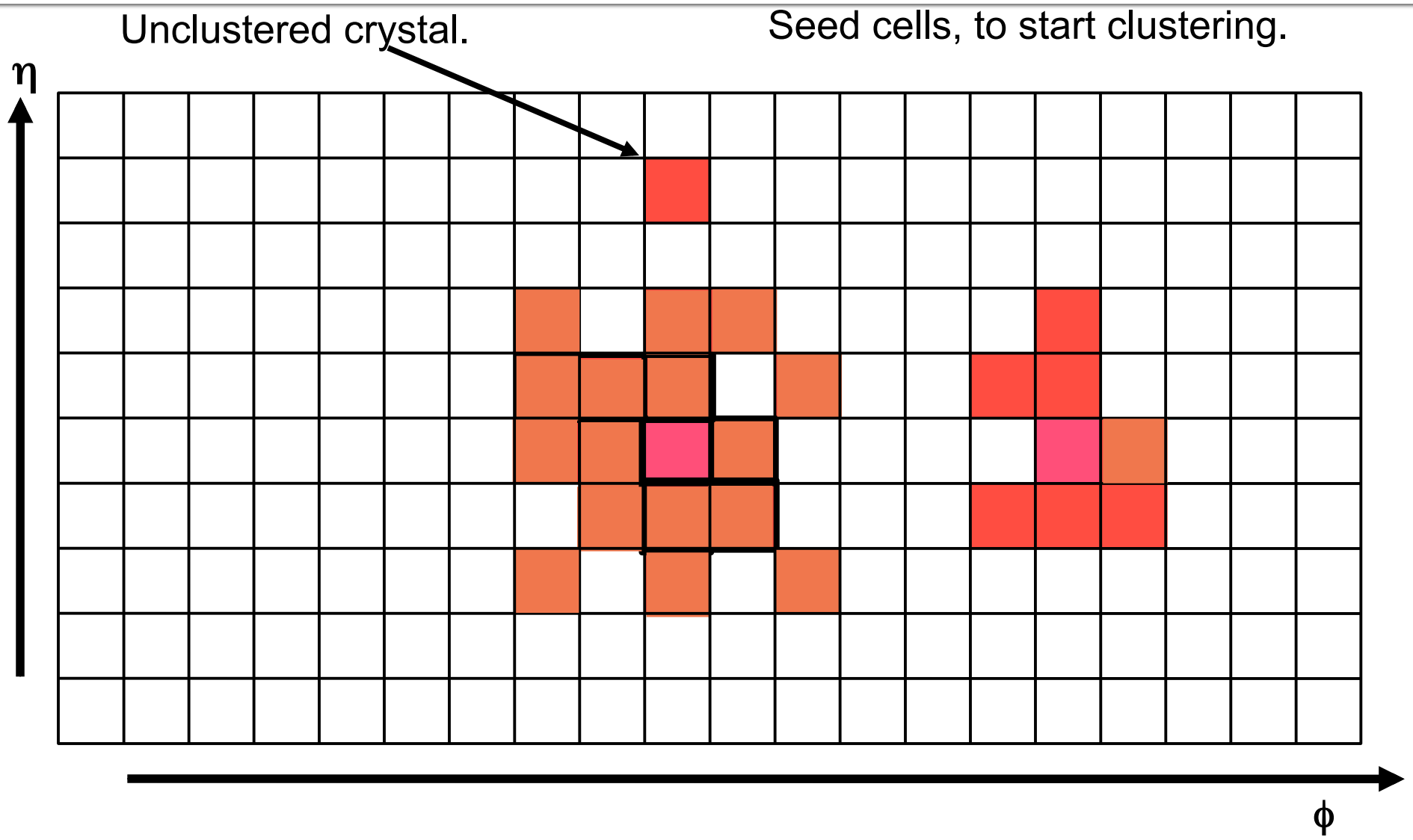
PFlow Clustering:



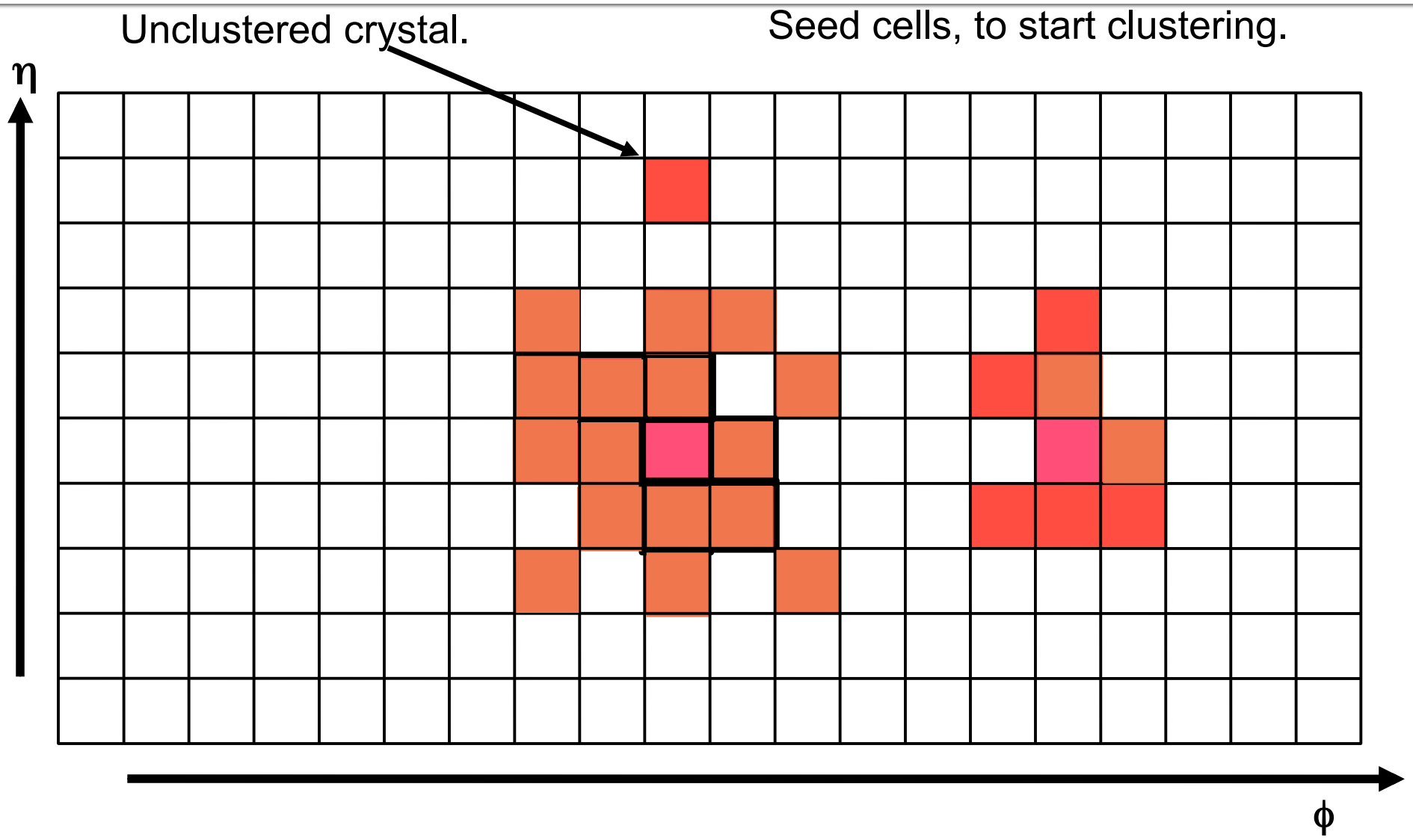
PFlow Clustering:



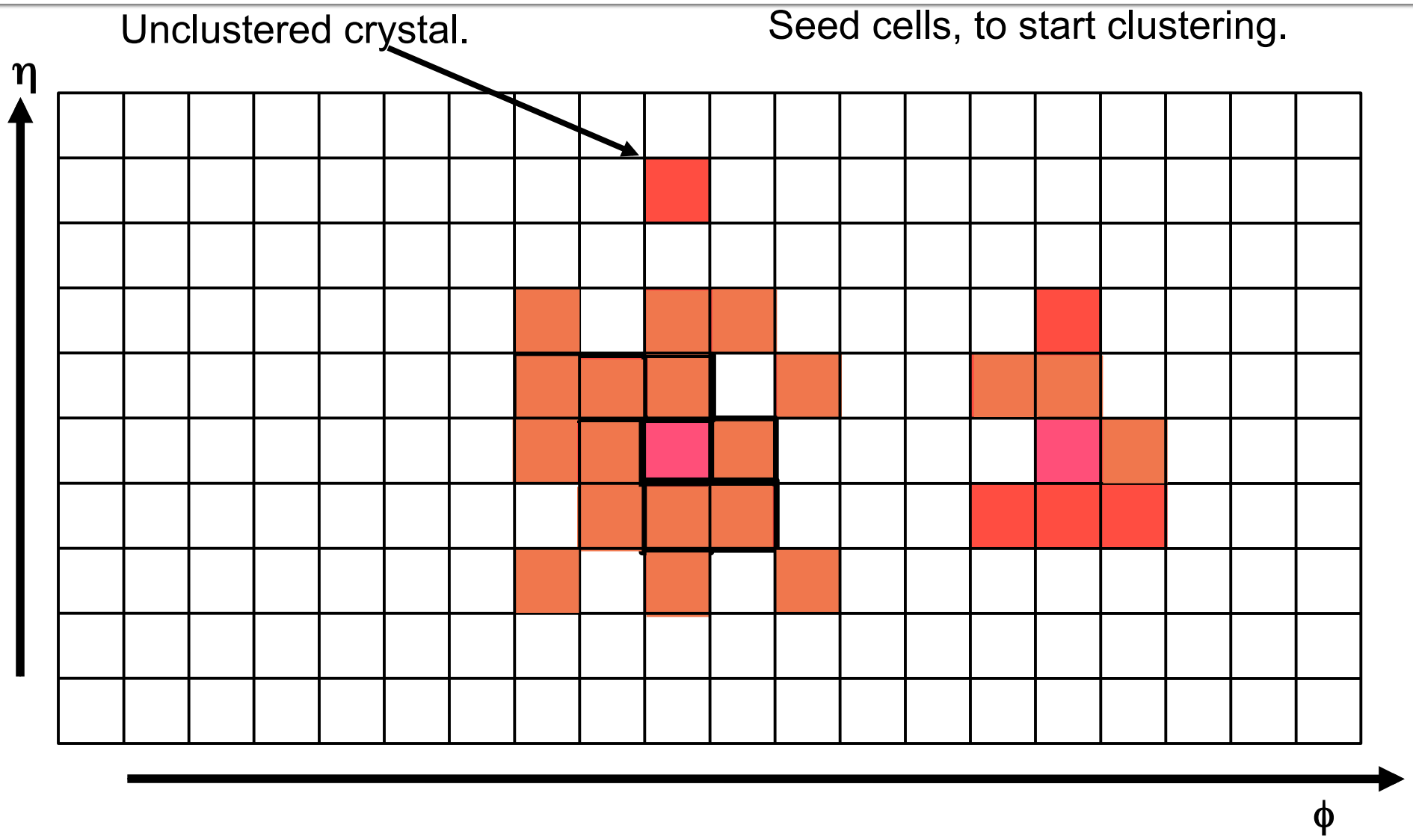
PFlow Clustering:



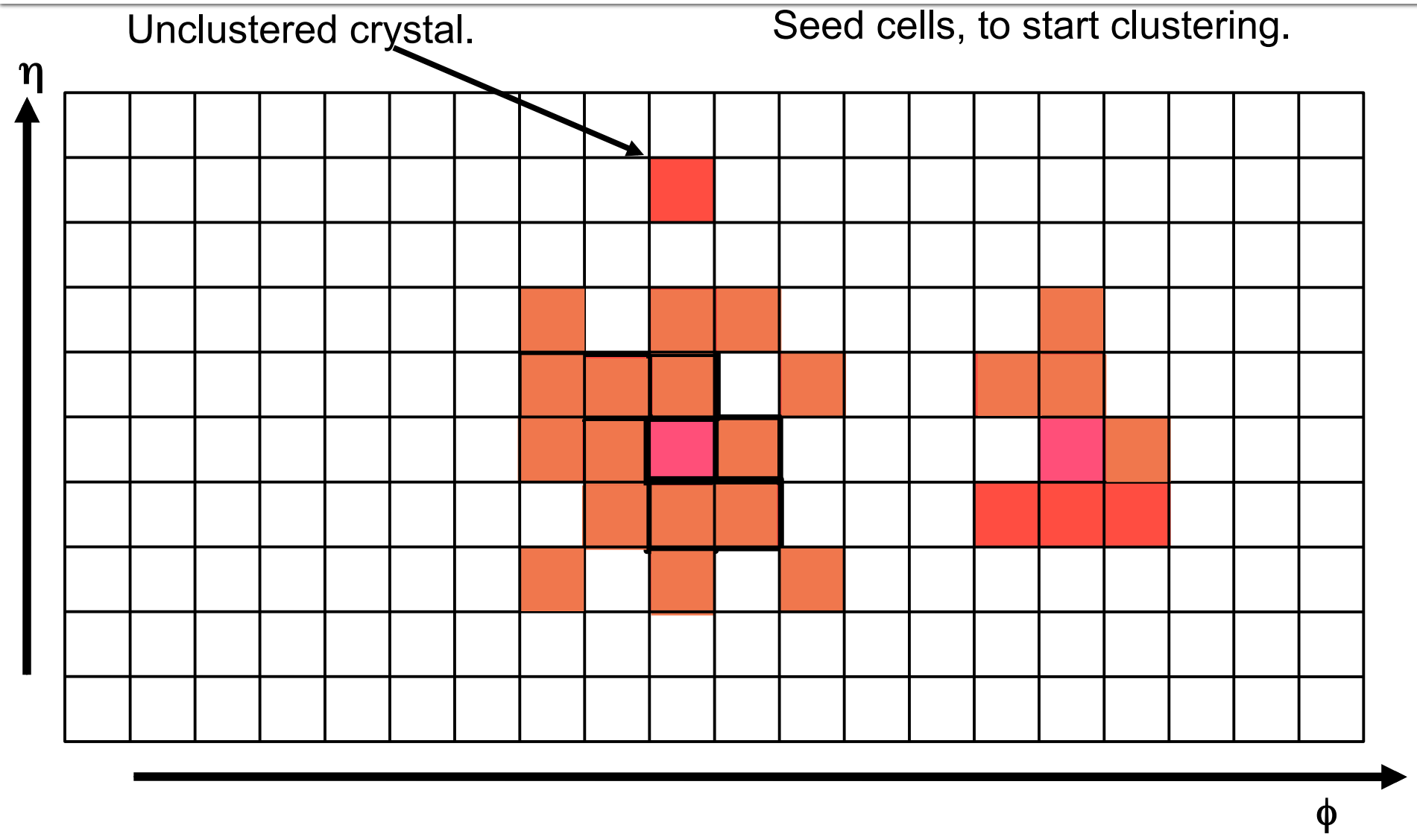
PFlow Clustering:



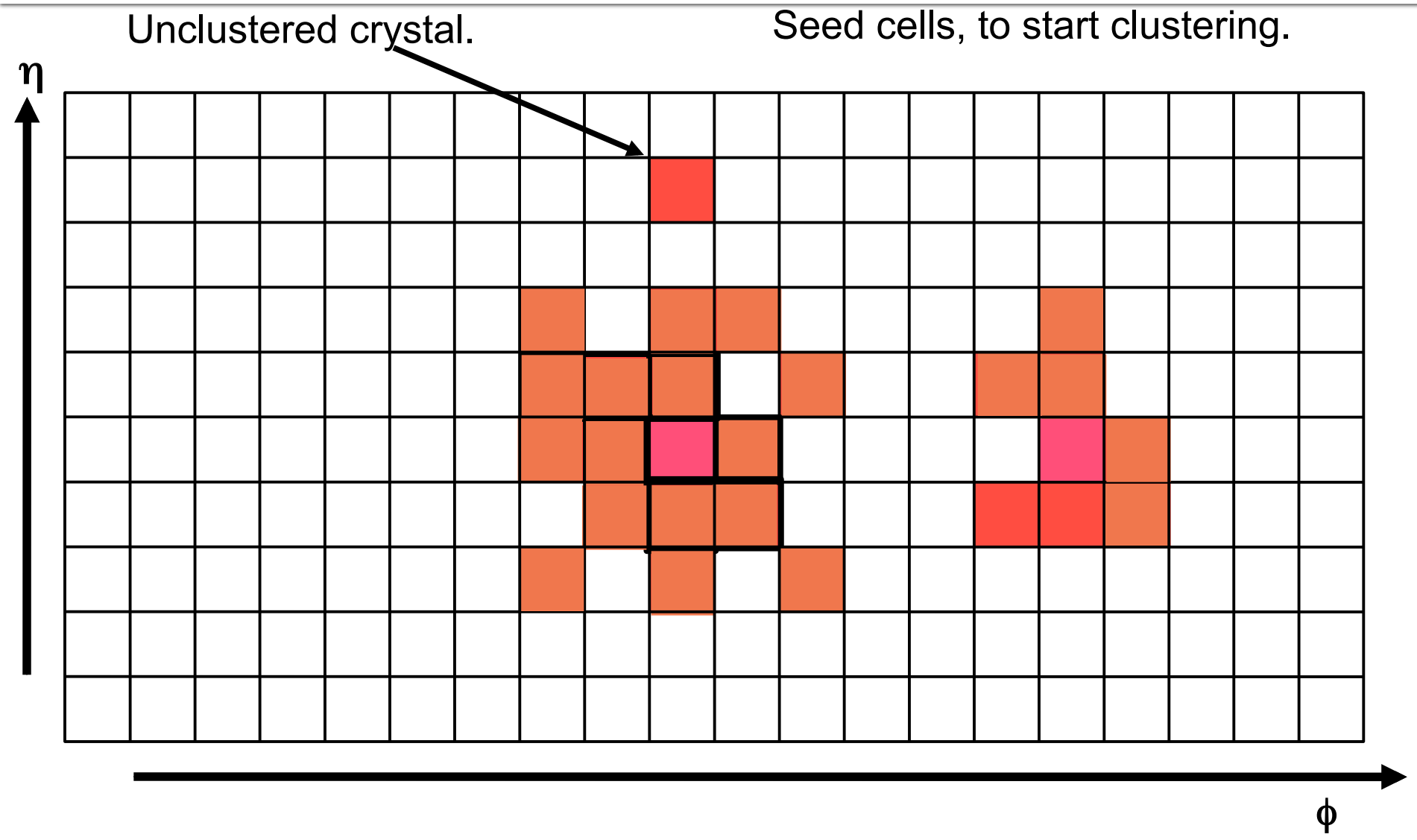
PFlow Clustering:



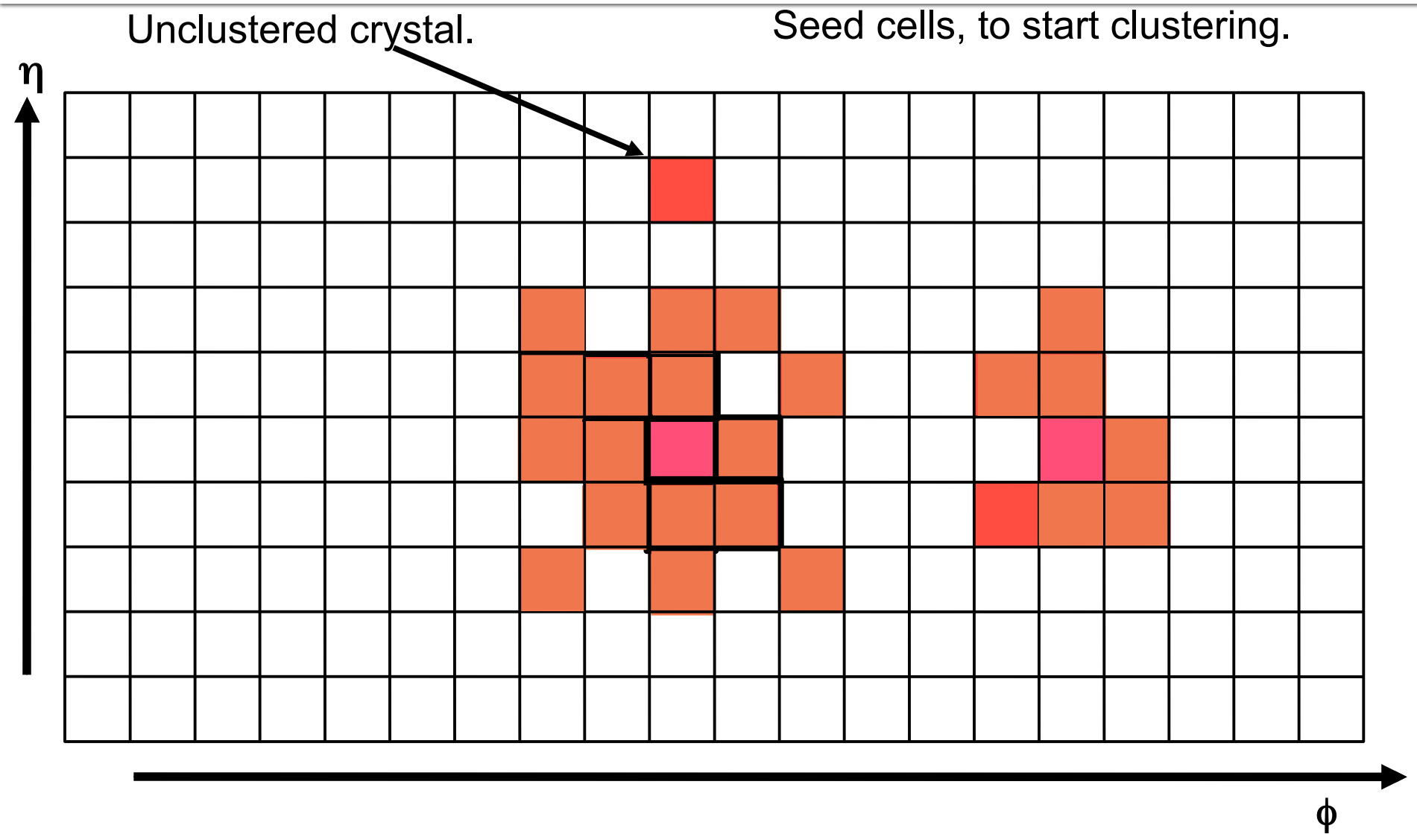
PFlow Clustering:



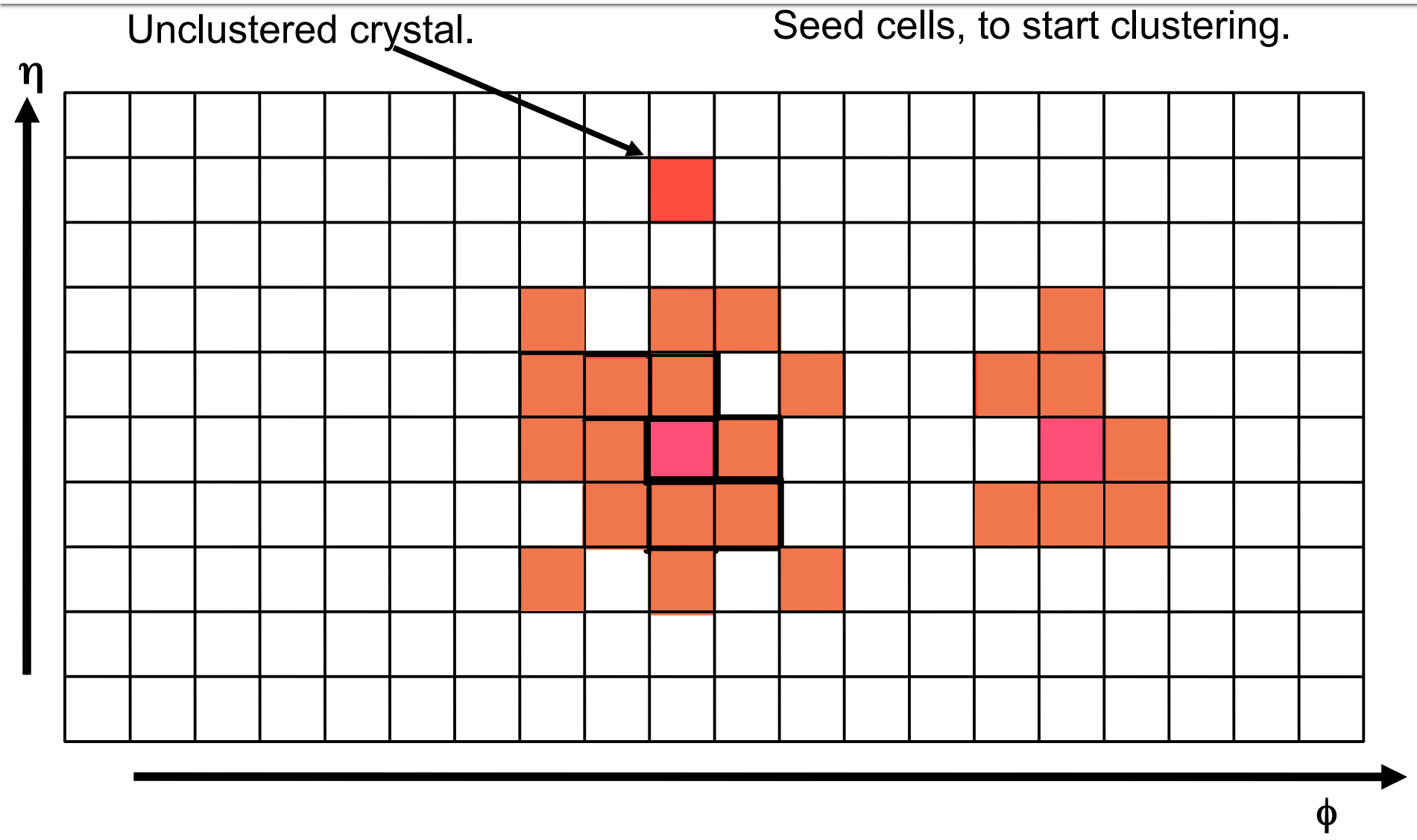
PFlow Clustering:



PFlow Clustering:



PFlow Clustering:



Cleaning:

- All of these algorithms share the cleaning of anomalous ECAL signals, aka "spikes".
  - In fact, Pflow clustering is actually extra strict about which crystals are allowed to seed on this account, since the same algorithm is used on HCAL and ECAL.
- There are two types of cleaning:
  - Topological: No crystal is an island, meaning for each hit that is used (above a threshold of 4 GeV presently), the neighbors must also have a non-negligible amount of energy in comparison.
  - Timing: Crystals above 4 GeV must be within approximately +/- 3 ns of "prompt", which is defined as time equals zero.
    - This is an approximation, really only for high energy crystals, owing to constant term of timing resolution.

That is all prologue:

- The best way to get a feel for how these algorithms will actually make decisions is to take these descriptions and try them out for yourself on the data.
- I've assembled three examples each of the hybrid algorithm in the barrel, the multi5x5 algorithm in the endcap, and three simple particle flow clusters. These are real data events, with all the attendant features.
- Work your way through them!