Late spill scout

> The late spill° on the candidate recipe of the *soft pool scout* ° follows the anti anneal. Holding the anti weight high keeps it near zero, and pausing the anneal delays it until the anneal resumes. Sharpening the slice temperature during the anneal also shrinks it, at no cost we could see. Holding the anti weight lowers removal°.

Spill. The largest drop on another op under the full edit at epoch 400, at any dose, net of the median of the same edit along every other coordinate axis of the run.…

Soft pool scout. Pooling the pull over slices 1 to 3 with the slice temperature held at 0.1 leaves no early spill, but one or two seeds in eight rise late, after epoch 128, to about the criterion; its epoch-64 states are where every run here starts.

Removal. How much of the anchored op the full edit removes at epoch 400, net as for spill.…

The *soft pool scout* left a candidate recipe with no early spill: the pool over slices 1 to 3, with the slice temperature τ_ℓ held at 0.1.[1] But on one of its eight seeds (707) the spill rises after epoch 128 and ends at the criterion, and with more seeds some runs would likely go over it.

1 Slice temperature…

The rise comes while the anti weight λ_s̄ anneals from its peak down to the per-slice holds.[2] By then this is the only loss term still moving. The learning rate comes down over the same epochs on its cosine decay, so a rise that follows one follows the other too, and moving the anti anneal on its own is a way to tell them apart.

2 Anti weight…

Spill-timing° found that the late kind of spill falls on ops the run is still learning. So one reading is that the anti term keeps e_1 closed while an op is learned, and the anneal opens it.

Spill-timing. Spill comes in an early kind, during the rise of the pull, and a late kind, after about epoch 100, on ops a run is still learning; quiet runs scatter by about 0.006 on its score.

This report asks whether a late change to the schedules keeps the late spill down, and what it costs.

Each arm changes one schedule, and only after epoch 64, so each run carries on from the stored epoch-64 state of the same seed and trains for the last 336 of 400 epochs. That makes four sets of eight runs, one of which is the condition unchanged. Only one seed rises under this condition, so the same four changes are also applied to the other finished condition of the *soft pool scout*, `τ_ℓ 0.1, by 40`, at the two seeds that rose under it.

Observations

- Spill along training (E1)°: on the two runs that rise, the rise follows the anti anneal. Held high, the anti weight keeps the spill° near zero; paused until epoch 200, the anneal delays the rise until it resumes. Sharpening τ_ℓ lowers it partway.

Spill along training. The full edit on each checkpoint kept from epoch 96 on (every 32 epochs, and the end), net of the median of eight reference axes, as spill-timing scored it....

Spill. The largest drop on another op under the full edit at epoch 400, at any dose, net of the median of the same edit along every other coordinate axis of the run....

- At the end of training (E2): every arm ends with no run over the criterion, against one for the condition unchanged. Sharpening τ_ℓ keeps removal° and the task where they were.

Removal. How much of the anchored op the full edit removes at epoch 400, net as for spill....

- Alignment and removal (E3)°: holding the anti weight keeps the alignment at the example answers lower. It also lowers removal, a little on most seeds and by about a quarter on one.

Alignment. How far a state points along e₁...

Scope and method

This is an exploratory study, with no preregistration and no gate. It covers one condition of the *soft pool scout*°, `τ_ℓ 0.1, τ_ℓ fixed`, at its eight seeds (704 to 711), and four ways of carrying it on from epoch 64. The same four carry on `τ_ℓ 0.1, by 40` at 705 and 707, the seeds that rose late under it:

***Soft pool scout*.** Pooling the pull over slices 1 to 3 with the slice temperature held at 0.1 leaves no early spill, but one or two seeds in eight rise late, after epoch 128, to about the criterion; its epoch-64 states are where every run here starts.

`base`
: The condition unchanged, with a checkpoint every 32 epochs as the arms have.

`anti late`
: The anti anneal pauses at epoch 64, holding the weight it has reached, and finishes from epoch 200 to 360. That weight is close to its peak (by epoch 64 the anneal has gone about 4% of the way to the holds). So the anti weight stays high while the late ops are learned, and reaches the same per-slice holds by the same epoch. Under `base` the anneal is a quarter of the way to the holds at epoch 128, and two-fifths of the way at 160, where spill-timing saw the late kind begin. So the two arms differ modestly there, and widely from epoch 200 on.

`anti held`
: The pause runs to epoch 360. So the anti weight never anneals to the holds, and only comes down with the anchor in the last 40 epochs. At slice 3 it stays near 0.24 through epoch 300, against 0.08 under `base`. This is the largest change to the anti term that keeps its first 64 epochs, so it shows the most the anti term can do about the late spill, and what that costs. We expected a cost. An early experiment, ex-2.1.11, found that the shape of the anti schedule mattered (a constant weight, high or low, cost grading), and late-pull found that a higher hold cost removal.

`τ_ℓ late`
: The slice temperature sharpens again, from 0.1 to 0.025 over epochs 64 to 360, alongside the anti anneal. Stage 1 of the soft pool scout found that a sharper `τ_ℓ` lowers the alignment at slice 1 and the early spill. This arm asks whether sharpening further does the same for the late kind. At 0.025 the pool over three slices is close to a max, so the pull is confined almost wholly to the slice where the evidence peaks (slice 3 on every seed).

Every arm uses only schedules and temperatures, nothing keyed to a token role, so each would transfer to real text as the recipe does.

Every run here trained on one GPU type, and every run of a seed shares its first 64 epochs. So a difference between an arm and `base` at the same seed comes from the change after epoch 64 alone. Two arms with the same schedule up to an epoch score the same to the last digit up to it (`anti late` and `anti held` to epoch 192).

But the late rise is small. On the runs it reaches, it ends near the criterion, while quiet runs scatter by about 0.006 around a value a little below zero (one standard deviation, over the stored runs of quiet conditions in *spill-timing*°).

***Spill-timing*.** Spill comes in an early kind, during the rise of the pull, and a late kind, after about epoch 100, on ops a run is still learning; quiet runs scatter by about 0.006 on its score.

The score along training nets against eight reference axes rather than 63, and is at least as noisy. At epoch 96, where the schedules of the arms have barely parted, the four runs of 707 under `τ_ℓ fixed` already spread over 0.017.

So the comparison that counts is the pairings on the runs that rise, judged over the whole of training rather than at one checkpoint: do the arms move them the same way?

Modal had no L4 GPUs free when this ran, so every run carried on from epoch 64 on an A10G, where the *soft pool scout* ran on L4. `base` therefore differs from the finished runs of the *soft pool scout* only in the numerics of the GPU, and that difference is informative in itself.

On the eight seeds of `τ_ℓ 0.1, τ_ℓ fixed`, `base` ends within 0.003 of the L4 run, and the two runs at 707 rise late on both GPUs. But `τ_ℓ 0.1, by 40` at 705, which ended at the criterion on L4, stays below zero on the A10G (second table of E2). So whether a seed rises late is partly a matter of numerics, and whether a single run rises says little by itself. Over a long enough run, a change of GPU likely acts much like a change of seed.

The measurements are those of stage 2 of the *soft pool scout*, netted by the D2.2 rule:

Spill
: The largest drop on another op under the full edit at epoch 400, at any dose, net of the median of the same edit along every other coordinate axis of the run. The criterion is 0.02.

Spill along training
: The full edit on each checkpoint kept from epoch 96 on (every 32 epochs, and the end), net of the median of eight reference axes, as spill-timing scored it. The draw of axes is the same for every run.

Removal
: How much of the anchored op the full edit removes at epoch 400, net as for spill. A run that removes less than 0.5 has lost the anchor.

Task change
: The clean score of each op less that of the control at the same seed, averaged over ops.

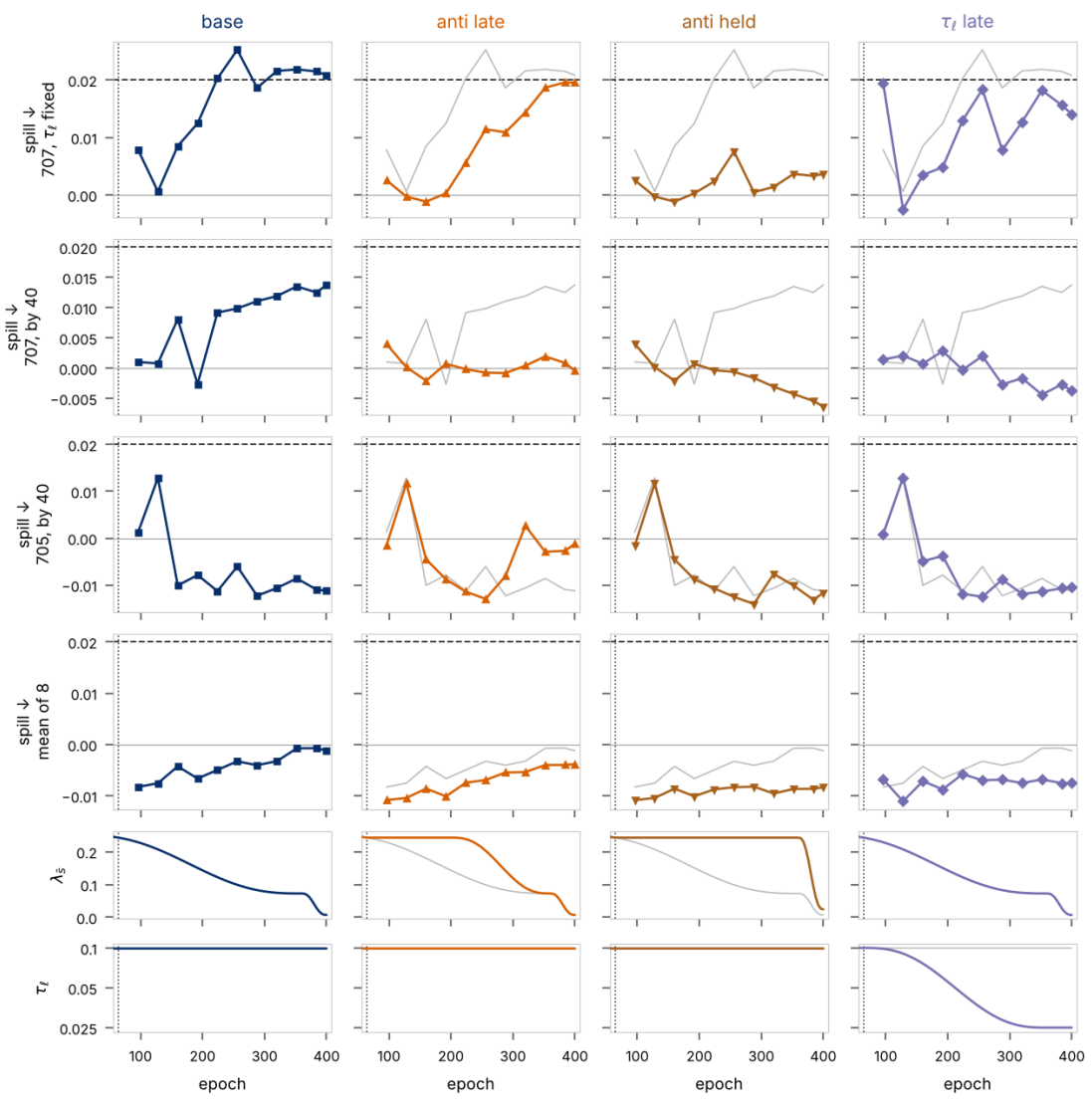
Alignment at the answers
: The mean alignment α with e_1 at the example answers of `difference` contexts, at slice 3, along training.

Spill along training (E1)

Does any arm keep the late rise down on the runs where it appears, and does the rise follow the anti weight in time? The figure has one column per arm, each with its own anti weight (at slice 3) and `τ_ℓ` at the bottom, so a change in the spill° can be read against the schedule just below it. The rows are the three runs that rose late in the *soft pool scout*° and the mean over the eight seeds of `τ_ℓ 0.1, τ_ℓ fixed`, with `base` drawn faintly behind each arm.

Spill. The largest drop on another op under the full edit at epoch 400, at any dose, net of the median of the same edit along every other coordinate axis of the run....

Soft pool scout. Pooling the pull over slices 1 to 3 with the slice temperature held at 0.1 leaves no early spill, but one or two seeds in eight rise late, after epoch 128, to about the criterion; its epoch-64 states are where every run here starts.



Spill along training, with the schedules the arms change. *One column per arm, with `base` drawn faintly behind the others. Top four rows: the spill at the checkpoints kept every 32 epochs from epoch 64, where the arms part, and at the end, against the median of eight reference edits; the first three rows are the runs that rise late in the soft pool scout, the fourth the mean over the eight seeds of `τ_ℓ 0.1, τ_ℓ fixed`. The dashed line is the criterion. Below them: the anti weight $\lambda_{\bar{s}}$ at slice 3, and the slice temperature `τ_ℓ` on a log scale.*

On both runs that rise here, the late spill follows the anti weight. Where the weight is held high (`anti held`), the spill stays near zero to the end. The third run, 705 under `by 40`, does not rise on this GPU in any arm, `base` included, so it says nothing about the arms.

Where the anneal pauses until epoch 200 (`anti late`), the spill of 707 under `τ_ℓ fixed` stays flat until the anneal resumes. Then it climbs as the weight comes down, and reaches the level of `base` by the end. On 707 under `by 40` the same arm keeps it flat throughout.

Sharpening `τ_ℓ` during the anneal lowers the spill on both runs that rise: partway on 707 under `τ_ℓ fixed`, and to zero on 707 under `by 40`. On 707 under `τ_ℓ fixed` its spill along training° is bumpier than that of the other arms, with a root-mean-square step between checkpoints of 0.009 against 0.005 for `base`; elsewhere the arms are about as bumpy as each other. We checked that the pull stays put: slice 3 has the highest alignment° at the example answers at every record after epoch 12, under `τ_ℓ late` as under `base`, so the bumps are something else that we did not measure.

Spill along training. The full edit on each checkpoint kept from epoch 96 on (every 32 epochs, and the end), net of the median of eight reference axes, as spill-timing scored it....

Alignment. How far a state points along e_1...

The mean over the eight seeds moves little, since seven of them are quiet under every arm. All three arms end a little below `base` on it.

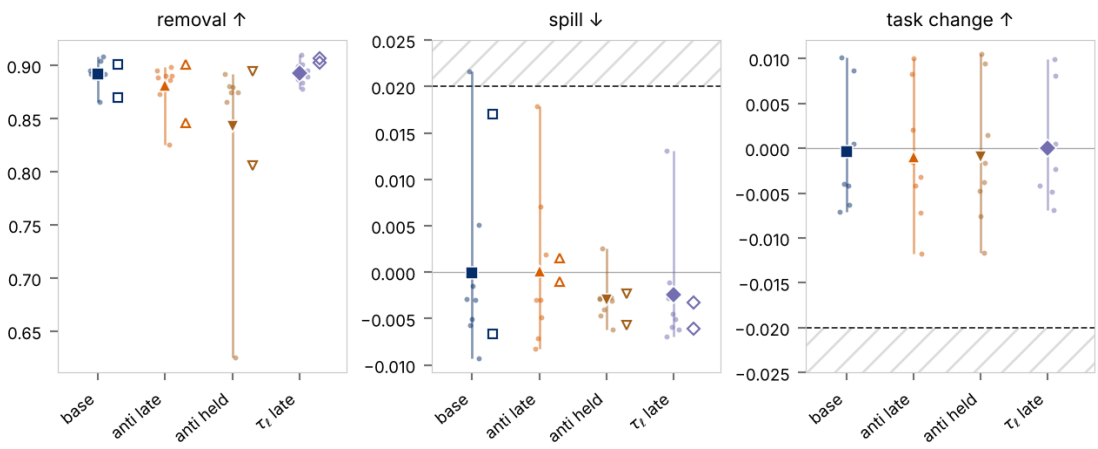
So the timing fits the reading from the introduction, that the anneal opens e_1 to the other ops. The `anti late` run of 707 is the clearest evidence: the same model, from the same state, rises only once the weight comes down. One run is not enough to be sure, but nothing here points the other way.

At the end of training (E2)

How does each arm compare with `base` at epoch 400, on spill°, on removal° and on the task?

Spill. The largest drop on another op under the full edit at epoch 400, at any dose, net of the median of the same edit along every other coordinate axis of the run....

Removal. How much of the anchored op the full edit removes at epoch 400, net as for spill....



Removal, spill, and task at epoch 400, by arm. *One column per arm on `τ_ℓ 0.1, τ_ℓ fixed`, the 8 seeds as small dots over a thin bar for their range, and the seed mean as the large mark; the hollow marks beside are the same arm on `τ_ℓ 0.1, by 40` at seeds 705 and 707.* ***Left:*** *removal at the full dose.* ***Middle:*** *spill, the largest net drop on another op at any dose; the dashed line is the criterion, with the side over it hatched. Both are net of the median over the other 63 axes of the run.* ***Right:*** *the task change, the mean clean score over ops less that of the control at the seed.*

arm	removal ↑	less base	spill ↓	less base	over criterion ↓	task change ↑
`base`	0.89 (0.87 to 0.91)	–	−0.000 (−0.009 to 0.022)	–	1 of 8	−0.000 (−0.007 to 0.010)
`anti late`	0.88 (0.83 to 0.90)	−0.01 (−0.04 to 0.00)	0.000 (−0.008 to 0.018)	0.000 (−0.004 to 0.009)	0 of 8	−0.001 (−0.012 to 0.010)
`anti held`	0.84 (0.63 to 0.89)	−0.05 (−0.24 to −0.01)	−0.003 (−0.006 to 0.003)	−0.003 (−0.019 to 0.005)	0 of 8	−0.001 (−0.012 to 0.011)
`τ_ℓ late`	0.89 (0.88 to 0.91)	0.00 (−0.01 to 0.01)	−0.002 (−0.007 to 0.013)	−0.002 (−0.009 to 0.002)	0 of 8	0.000 (−0.007 to 0.010)

Removal, spill and task at epoch 400 on `τ_ℓ 0.1, τ_ℓ fixed`, *seed mean and range over 8 runs, net of the median drop over the other 63 coordinate axes of the run. "Less base" is the arm less `base` at the same seed. The task change is the mean over ops of the clean score less that of the control at the seed.*

run	stage 2 (L4)	`base`	`anti late`	`anti held`	`τ_ℓ late`
`τ_ℓ fixed`,707	0.020	0.022	0.018	0.003	0.013
`by 40`,705	0.020	−0.007	0.002	−0.006	−0.006
`by 40`,707	0.019	0.017	−0.001	−0.002	−0.003

Spill at the end on the runs that rose late in the soft pool scout. *The first column is the finished run of the soft pool scout, trained on L4 from the same epoch-64 state; the others ran on A10G.*

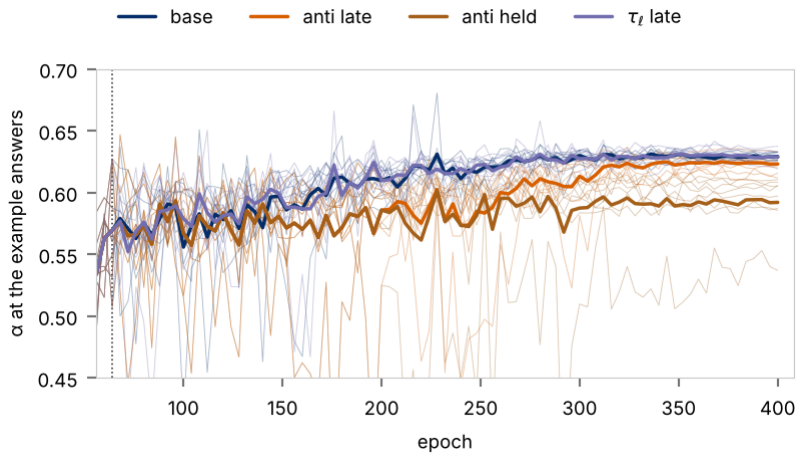
No arm has a run over the criterion, against one for `base` (707, at 0.022). On that run the spill falls to 0.003 with the anti weight held, and to 0.013 with `τ_ℓ` sharpened. `anti late` ends at 0.018, just under the line, as E1 showed. On 707 under `by 40`, all three arms end below zero, against 0.017 for `base`.

`τ_ℓ late` keeps removal where `base` has it, on the mean and seed by seed. The two anti arms lower removal: a little on most seeds, and on one seed by much more under `anti held` (E3). None of the arms changes the task by more than a few thousandths.

Alignment and removal (E3)

Holding the anti weight keeps the late spill° down, but the anti term pushes on e_1 everywhere, the anchored op included. How much does holding it cost?

Spill. The largest drop on another op under the full edit at epoch 400, at any dose, net of the median of the same edit along every other coordinate axis of the run....



The alignment at the example answers along training. *The mean alignment α with e_1 at the example answers of* `difference` *contexts, at slice 3, recorded every four epochs from epoch 64, where the arms part: the seed mean of each arm on* `τ_ℓ 0.1, τ_ℓ fixed` *as a line, the eight seeds as hairlines (a few dip below the lower edge).*

Alignment. How far a state points along e_1...

Under `base` the alignment° at the example answers keeps growing after epoch 64, as the anti weight comes down. Under both anti arms it grows more slowly while the weight is held. `anti held` levels off below `base` (a seed mean of 0.59 at epoch 400, against 0.63), and `anti late` catches up most of the way once its anneal resumes. `τ_ℓ late` stays with `base` throughout.

Removal. How much of the anchored op the full edit removes at epoch 400, net as for spill....

Removal° follows. Under `anti held` it is lower than under `base` on every seed: by a few hundredths on seven of them, a change of −0.02 (−0.04 to −0.01) as mean and range, and by about a quarter on 710 (0.87 to 0.63), the seed with the lowest alignment.

The run at 710 keeps the anchor by our rule, but at that size it is closer to a lost run than any of the others. `anti late` lowers removal less, and on one seed (710 again) by a few hundredths.

So holding the anti weight high shows what the anti term can do about the late spill, but it is not a recipe. It trades the late spill for removal, and on one seed in eight that cost is large, as we expected.

Discussion

The late spill° on the candidate recipe follows the anti anneal, which fits the reading in the late-spill backlog item: the anti term keeps e_1 closed to the other ops, and the anneal opens it.

Spill. The largest drop on another op under the full edit at epoch 400, at any dose, net of the median of the same edit along every other coordinate axis of the run.…

Spill-timing° left open which way the link between spill and learning runs: learning an op late may leave it free to land on e_1, or whatever draws it to e_1 may also slow its learning. The second would not tie the spill to the anti schedule, and here the schedule alone moves the rise. But the two need not exclude each other, since we did not measure when the spilled op (`mix`) is learned on these runs. The evidence rests on two rising runs, so it is a strong hint, short of a settled result.

Spill-timing. Spill comes in an early kind, during the rise of the pull, and a late kind, after about epoch 100, on ops a run is still learning; quiet runs scatter by about 0.006 on its score.

Of the three changes, sharpening τ_ℓ during the anneal is the one that costs nothing we measured. It lowers the late spill on both rising runs, and keeps removal° and the task where they were. It adds no term or weight to the recipe, only a second step in a temperature the recipe already has.

Removal. How much of the anchored op the full edit removes at epoch 400, net as for spill.…

Why it helps is less clear than for the anti arms. A sharper slice pool puts the pull almost wholly on slice 3, which may leave slices 1 and 2 freer for the anti term to clear, but nothing here measures that.

Paused, the anti anneal delays the spill without preventing it under `τ_ℓ fixed`, so the pause alone changes little at the end of training.

The GPU difference is a reminder of how small this effect is. A run that reaches the criterion late on one GPU can stay below zero on another, from the same state with the same schedule. So a count of runs over the criterion at a handful of seeds cannot tell these arms apart. If one of them carries on to more seeds, the paired spill along training° is the measurement to watch.

Spill along training. The full edit on each checkpoint kept from epoch 96 on (every 32 epochs, and the end), net of the median of eight reference axes, as spill-timing scored it.…

References

Soft pool scout

Pooling the pull over slices 1 to 3 with the slice temperature held at 0.1 leaves no early spill, but one or two seeds in eight rise late, after epoch 128, to about the criterion; its epoch-64 states are where every run here starts.

Spill-timing

Spill comes in an early kind, during the rise of the pull, and a late kind, after about epoch 100, on ops a run is still learning; quiet runs scatter by about 0.006 on its score.

ex-2.1.11

The shape of the anti schedule matters: holding the anti weight constant, at its peak or at its hold, cost grading.

Late-pull

A higher anti hold (0.2) kept every run within the spill criterion, at some cost to removal.

---

1. Slice temperature: how sharply the anchor pull picks among the pooled slices. A lower τ_ℓ puts more of the pull on the one slice where the evidence peaks. ↩
2. Anti weight: the strength of the anti term, which pushes states off e_1. It anneals (ramps down on a schedule) from a peak to a hold set for each slice. ↩