For discussion

How to Apportion Blame for a Queue with Arrivals in Bursts?

Bob Briscoe, Independent <research@bobbriscoe.net>

Feb 2023

Roadmap

- Question
- Explanation of approach
- Explain parameter space
- Results
- Evaluation
- Interpretation
- Implications & Next steps

Question

- How to apportion blame for a queue
  with arrivals in bursts?

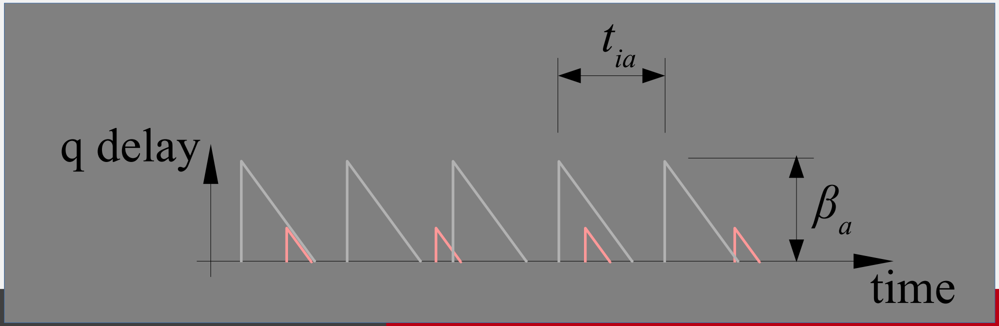
- Only marking packets, with no flow state

- Arrivals in bursts can lead to a queue
    - even if they can be served on average
- Subsequent smoother arrivals sit behind the burst
    - with knock-on impact on the next arrivals
    - even after the initial burst has all departed

* may use flow state

Approach

- Model? Simulate? Testbed?
  - model first: initial goal is understanding
  - noise of reality (packet sizing, timing dither, sync effects) could otherwise obstruct
- Unresponsive? Responsive?
  - unresponsive: cannot assume a response, so marking might solely drive policing*
- Simplest sufficient scenario; 2 flows, a & b, with:
  - constant but different burst sizes, β
  - constant but different capacity shares, λ
  - No need for either to vary (understand bursts first, not bursts of bursts)
- Reduces to 2 (sawtooth) waves with different amplitude β & wavelength (interval t_i)
  - capacity share, $\lambda = \beta / t_i$
  - any 1 of these 3 variables depends on the other 2



* might use flow state

Approach
Normalized metrics

Goal: results applicable to any link rate and any step marking threshold delay

- Burst size β is in units of time (queue delay)
  - normalized to: marking threshold = 1 unit of time
- On time series plots, time is also normalized
  - queue delay at marking threshold = 1 unit of time
- Marking rate, λp, is marked bits per unit time
  - normalized as a dimensionless fraction of link bit rate = 1
- Marking probability, p, and capacity share, λ
  - both dimensionless and bounded within [0,1]
  - so normalized marking rate, λp, also bounded within [0,1]
- Comparison metrics use difference, $p_a - p_b = \Delta p$, not ratio p_a/p_b
  - not inflated as $p_b \rightarrow 0$
  - for visualization, and irrespective of any congestion control assumptions

Approach

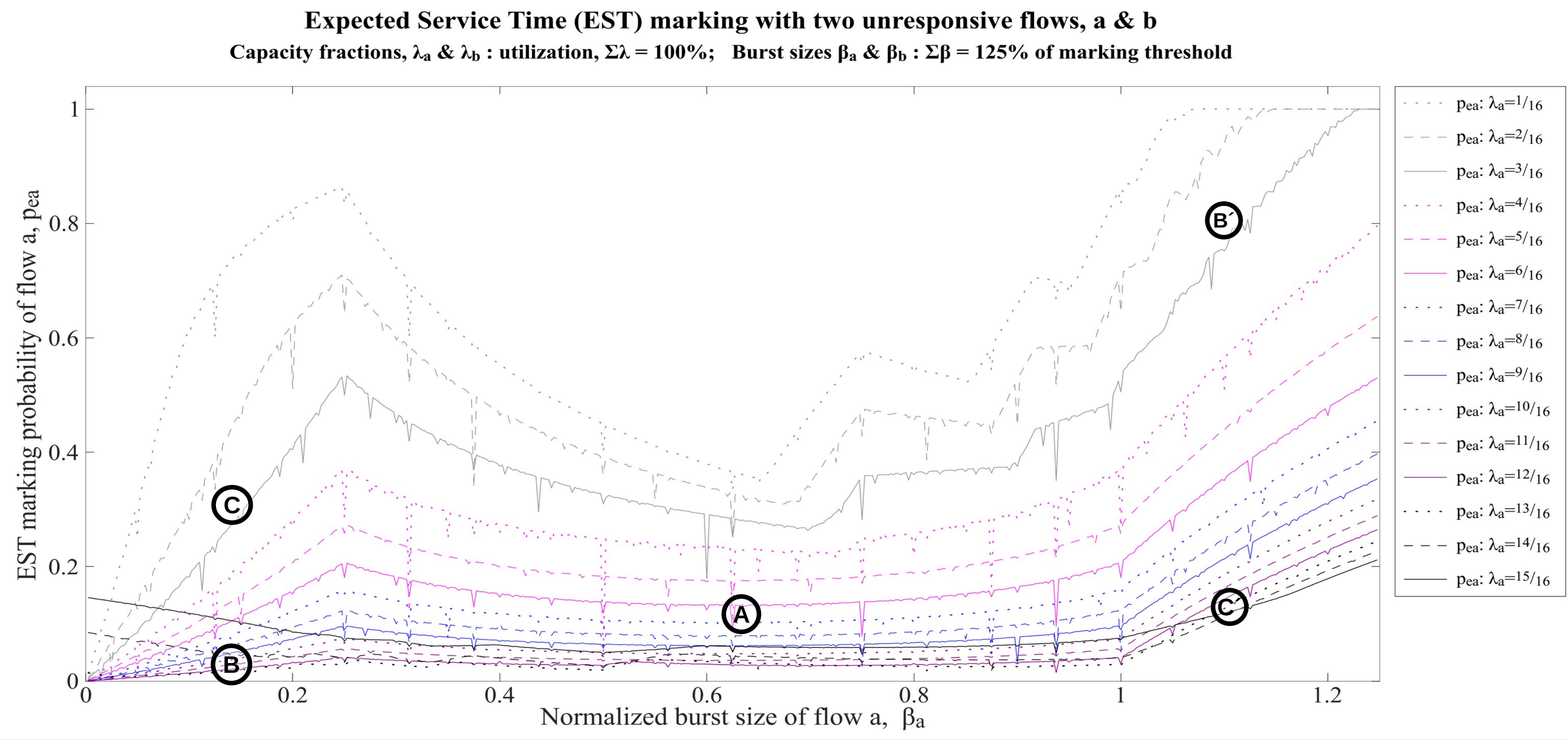
Parameter space

- A full scan of all 4 dimensions:

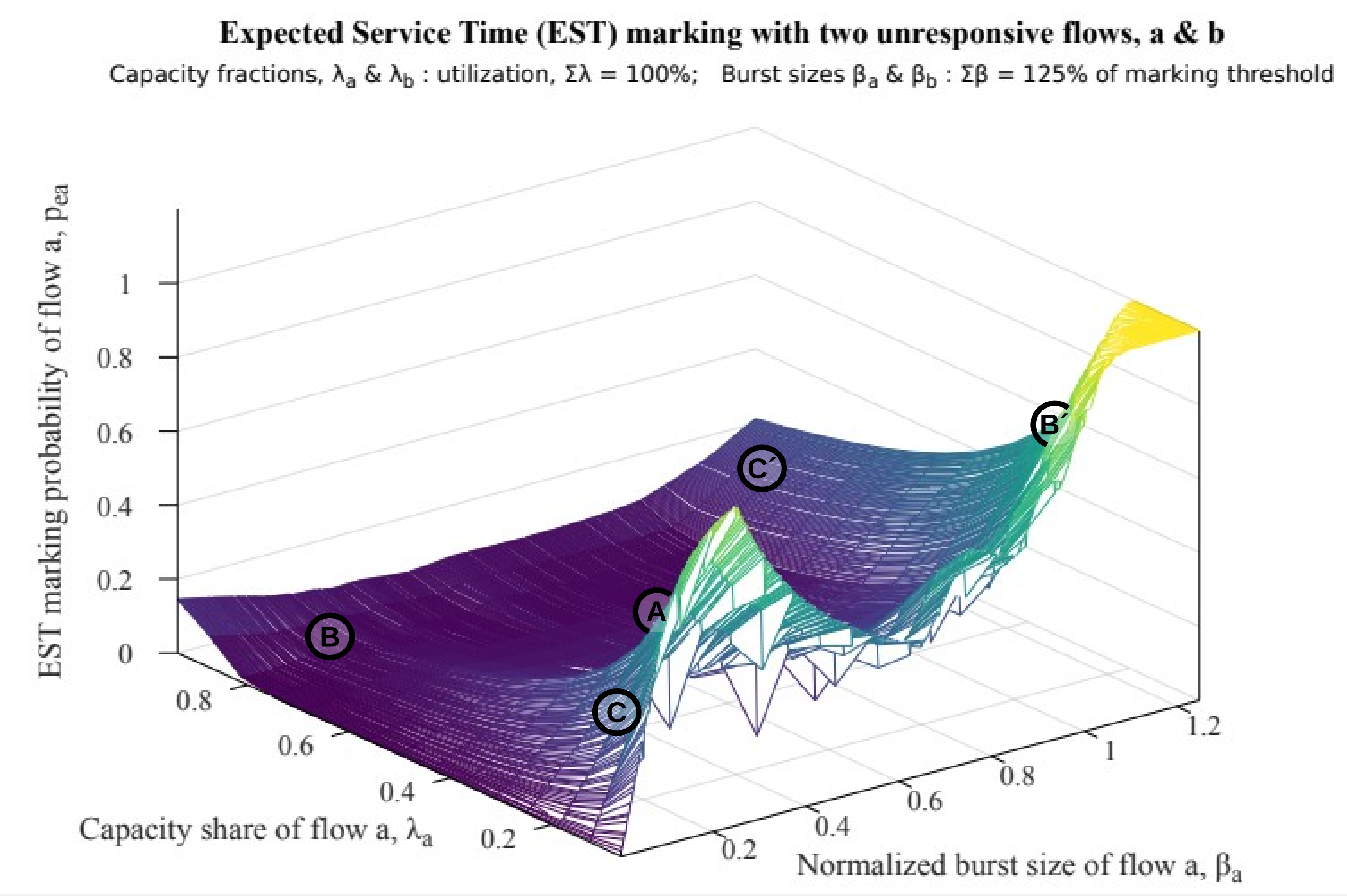
  $\beta_a, \lambda_a, \beta_b, \lambda_b$

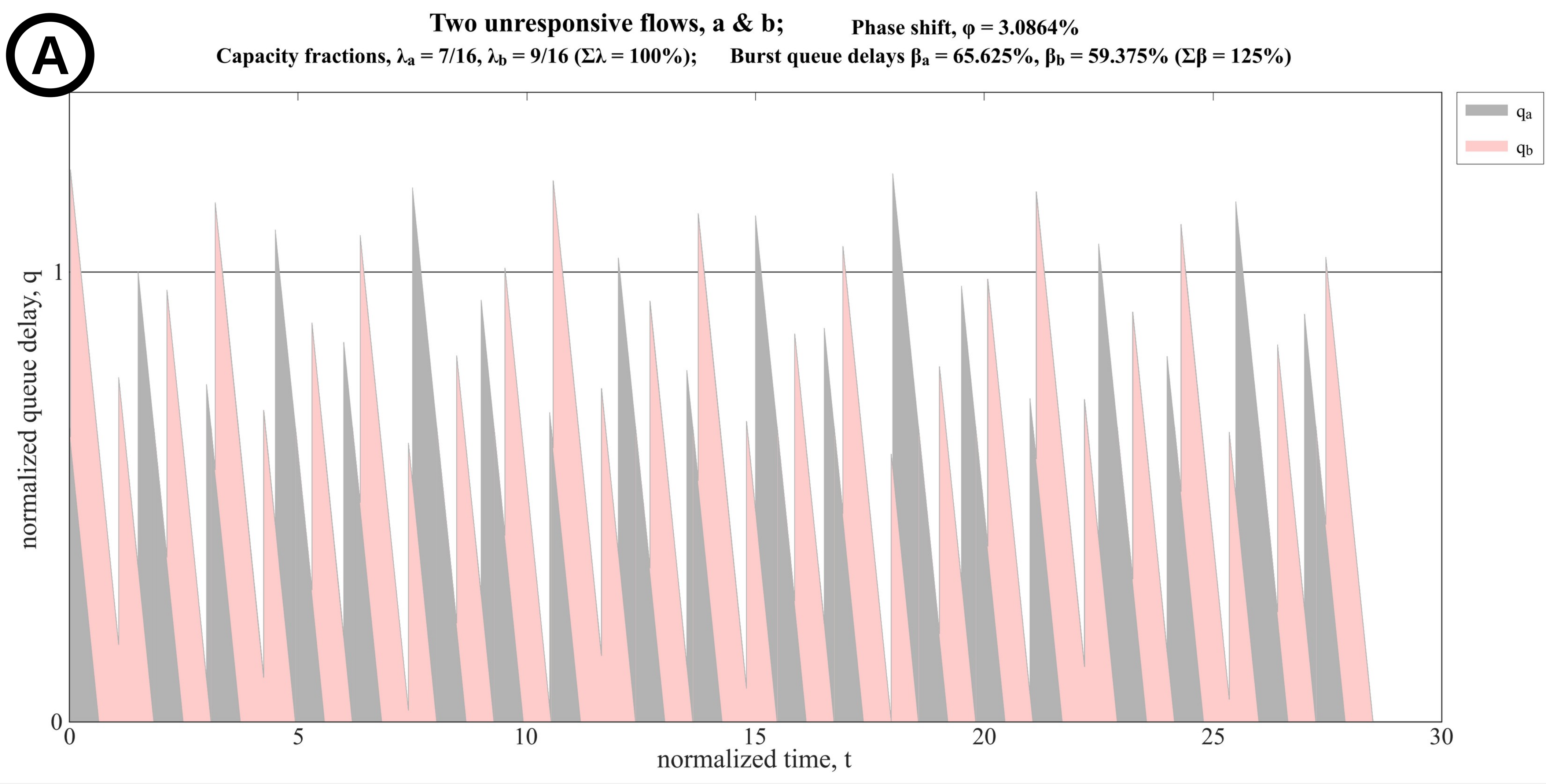
  would generate more heat than light
- To focus on apportioning blame, scan the parameter space of one flow (β_a, λ_a), while trying to keep the whole system constant, i.e.
  - constrain $\Sigma\lambda$ (utilization) to a small selection of constants (assume $\Sigma\lambda \leq 1$)
  - constrain $\Sigma\beta$ (max total burst) to a small selection of constants
- Compare two marking approaches, based on the q delay...
  - ...a packet itself experiences (the queue ahead at enqueue)
  - ...a packet causes to others (the queue behind at dequeue)

How to Interpret the Parameter Space



Same example parameter space in 3-D

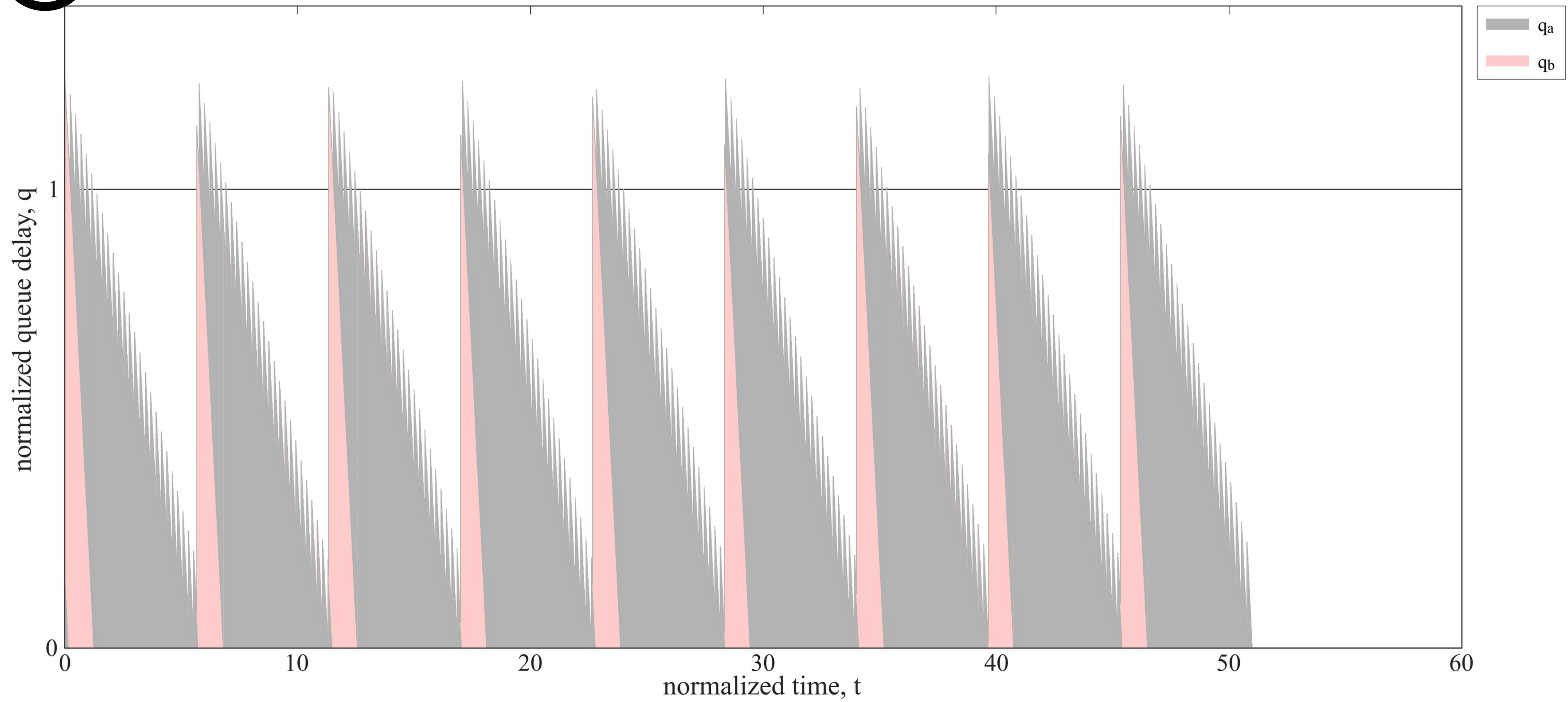



A
Two unresponsive flows, a & b;
Phase shift, φ = 3.0864%
Capacity fractions, λa = 7/16, λb = 9/16 (Σλ = 100%);
Burst queue delays βa = 65.625%, βb = 59.375% (Σβ = 125%)
qa
qb
normalized queue delay, q
1
0
0
5
10
15
20
25
30
normalized time, t

B

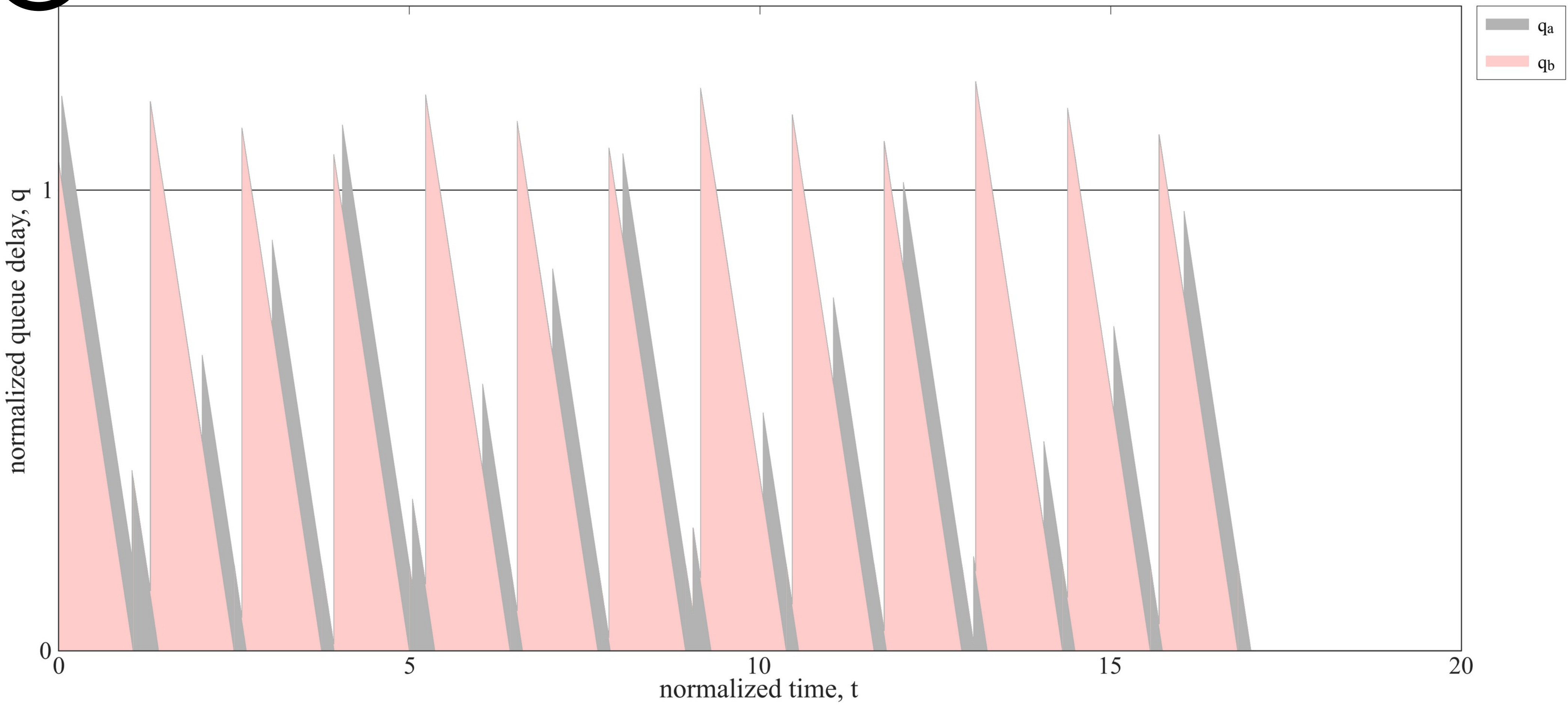
Two unresponsive flows, a & b; Phase shift, $\varphi = 3.0864\%$

Capacity fractions, $\lambda_a = 13/16$, $\lambda_b = 3/16$ ($\Sigma\lambda = 100\%$); Burst queue delays $\beta_a = 18.75\%$, $\beta_b = 106.25\%$ ($\Sigma\beta = 125\%$)



Two unresponsive flows, a & b; Phase shift, $\varphi = 3.0864\%$

Capacity fractions, $\lambda_a = 3/16$, $\lambda_b = 13/16$ ($\Sigma\lambda = 100\%$); Burst queue delays $\beta_a = 18.75\%$, $\beta_b = 106.25\%$ ($\Sigma\beta = 125\%$)



Ideals for apportioning blame

- Marking probability of one flow, p_a
  1) should monotonically increase with its burst size, β_a
  2) should not decrease with its capacity share, λ_a

  assuming the whole system is otherwise constant
- Satisfying both ideals would be robust but probably unattainable, e.g.
  - would fail on #1 if marking saturates, e.g. v large bursts
  - unsure if #2 is even satisfied with equal constant burstiness (see control expt later)
- Some scope to relax either ideal,
  - but unable to quantify precisely, so far

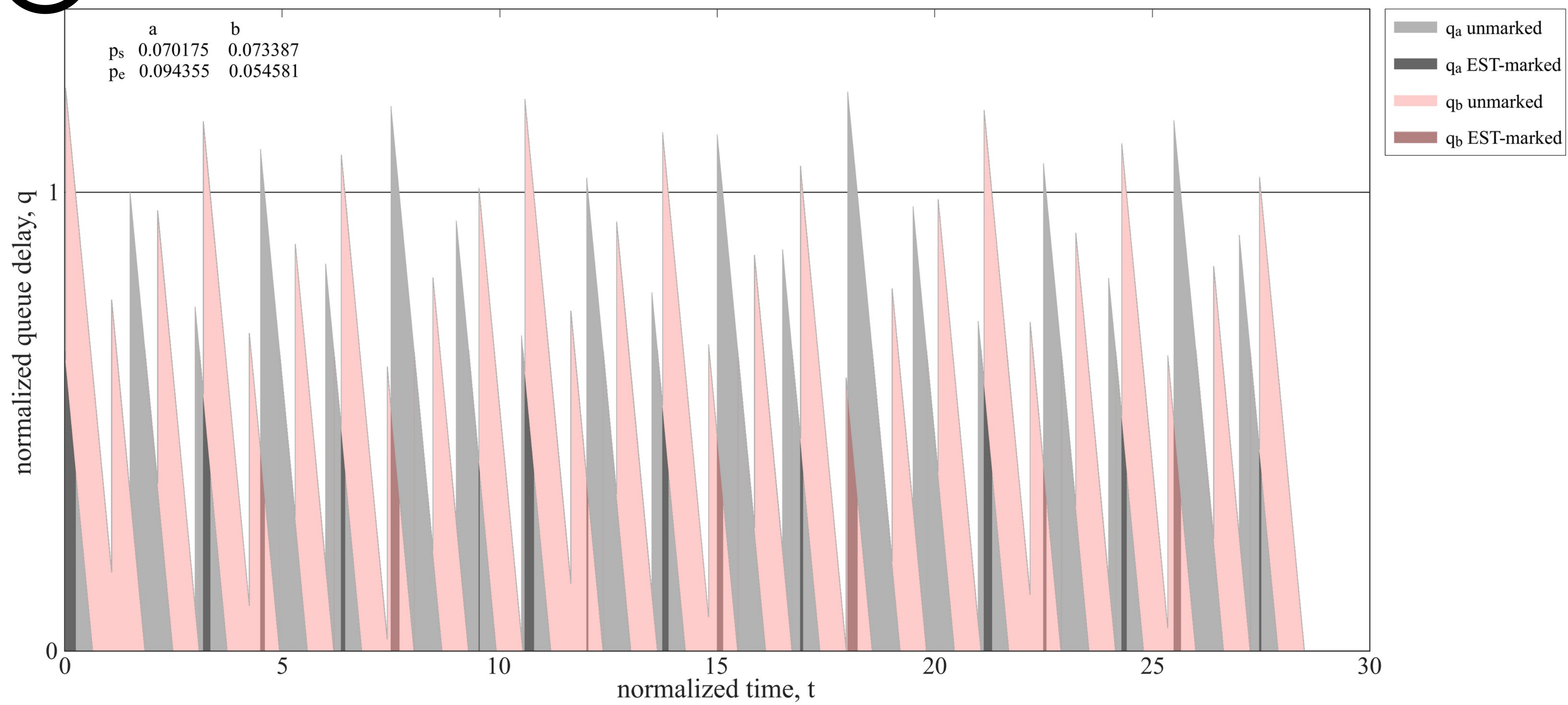
Compare 2 marking approaches

- delay to self
  - sojourn-based marking (s)
  - queue delay ahead at enqueue
  - visualization: colour of flow over threshold
- delay to others
  - expected service time (e)
  - queue delay behind at dequeue
  - visualization: colour of flow dequeued when q over threshold

A

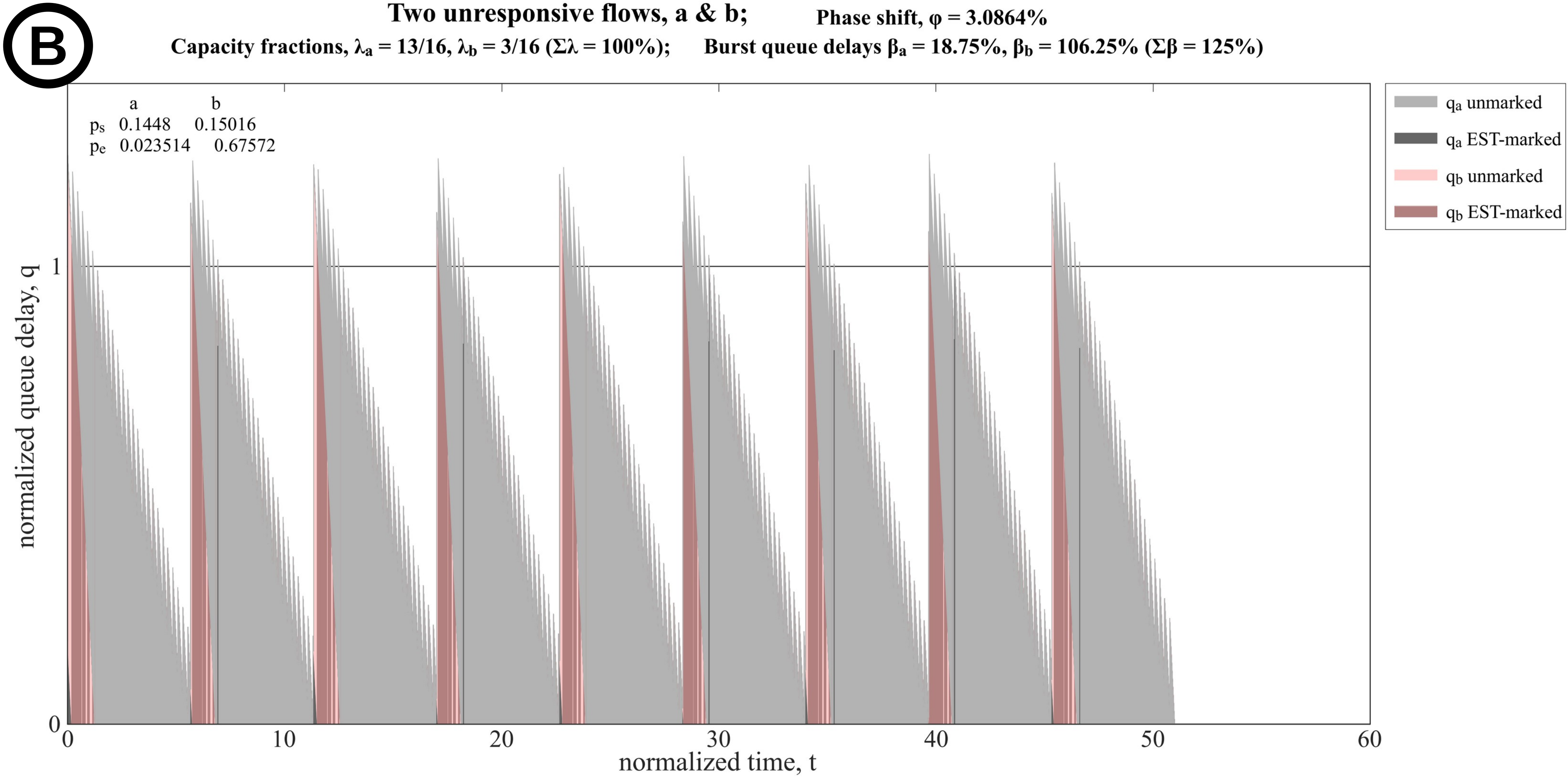
Two unresponsive flows, a & b; **Phase shift, φ = 3.0864%**

Capacity fractions, $\lambda_a = 7/16$, $\lambda_b = 9/16$ ($\Sigma\lambda = 100\%$); Burst queue delays $\beta_a = 65.625\%$, $\beta_b = 59.375\%$ ($\Sigma\beta = 125\%$)



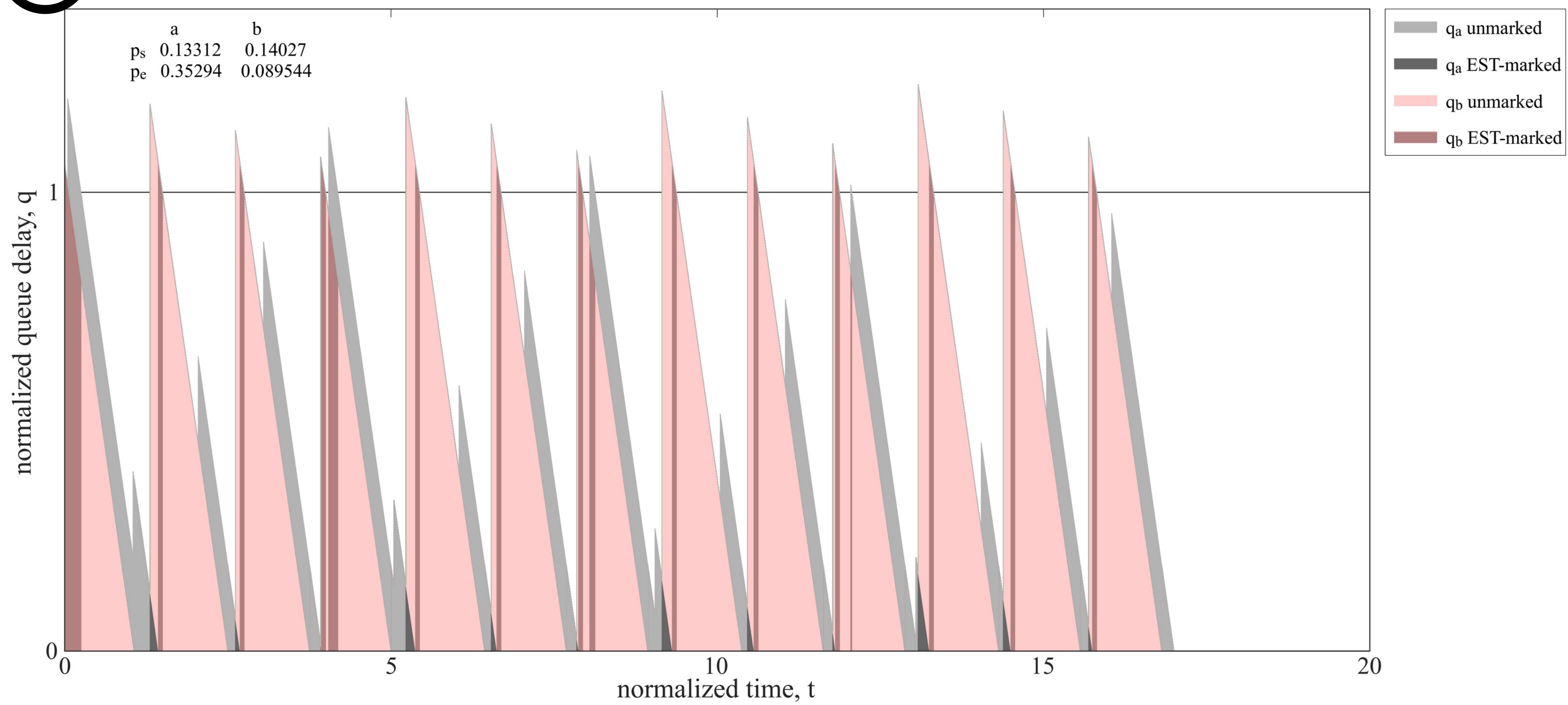
Two unresponsive flows, a & b; **Phase shift, $\varphi = 3.0864\%$**

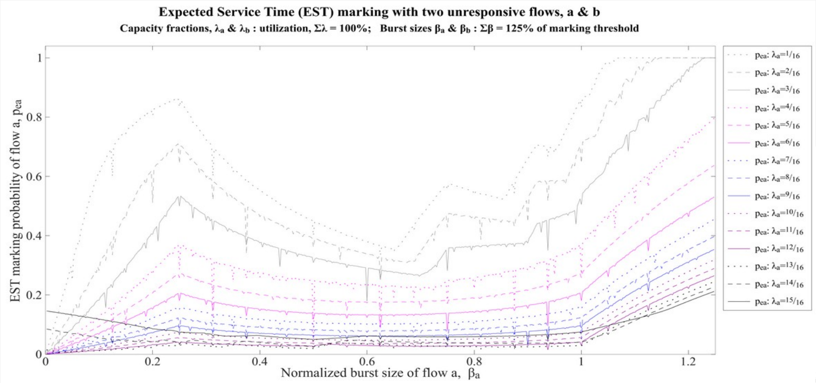
Capacity fractions, $\lambda_a = 13/16$, $\lambda_b = 3/16$ ($\Sigma\lambda = 100\%$); Burst queue delays $\beta_a = 18.75\%$, $\beta_b = 106.25\%$ ($\Sigma\beta = 125\%$)



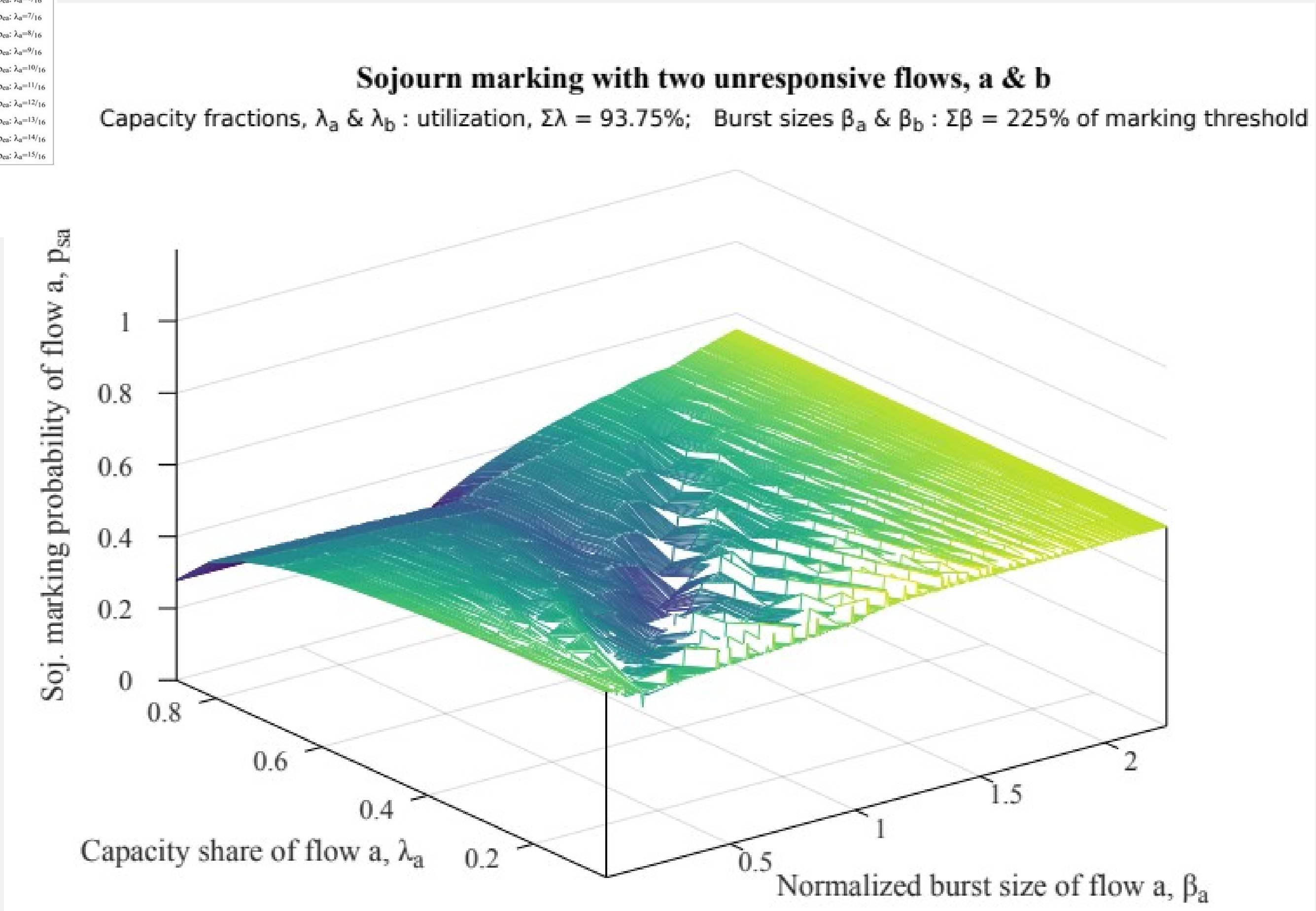
Two unresponsive flows, a & b; Phase shift, φ = 3.0864%

Capacity fractions, $\lambda_a = 3/16$, $\lambda_b = 13/16$ ($\Sigma\lambda = 100\%$); Burst queue delays $\beta_a = 18.75\%$, $\beta_b = 106.25\%$ ($\Sigma\beta = 125\%$)



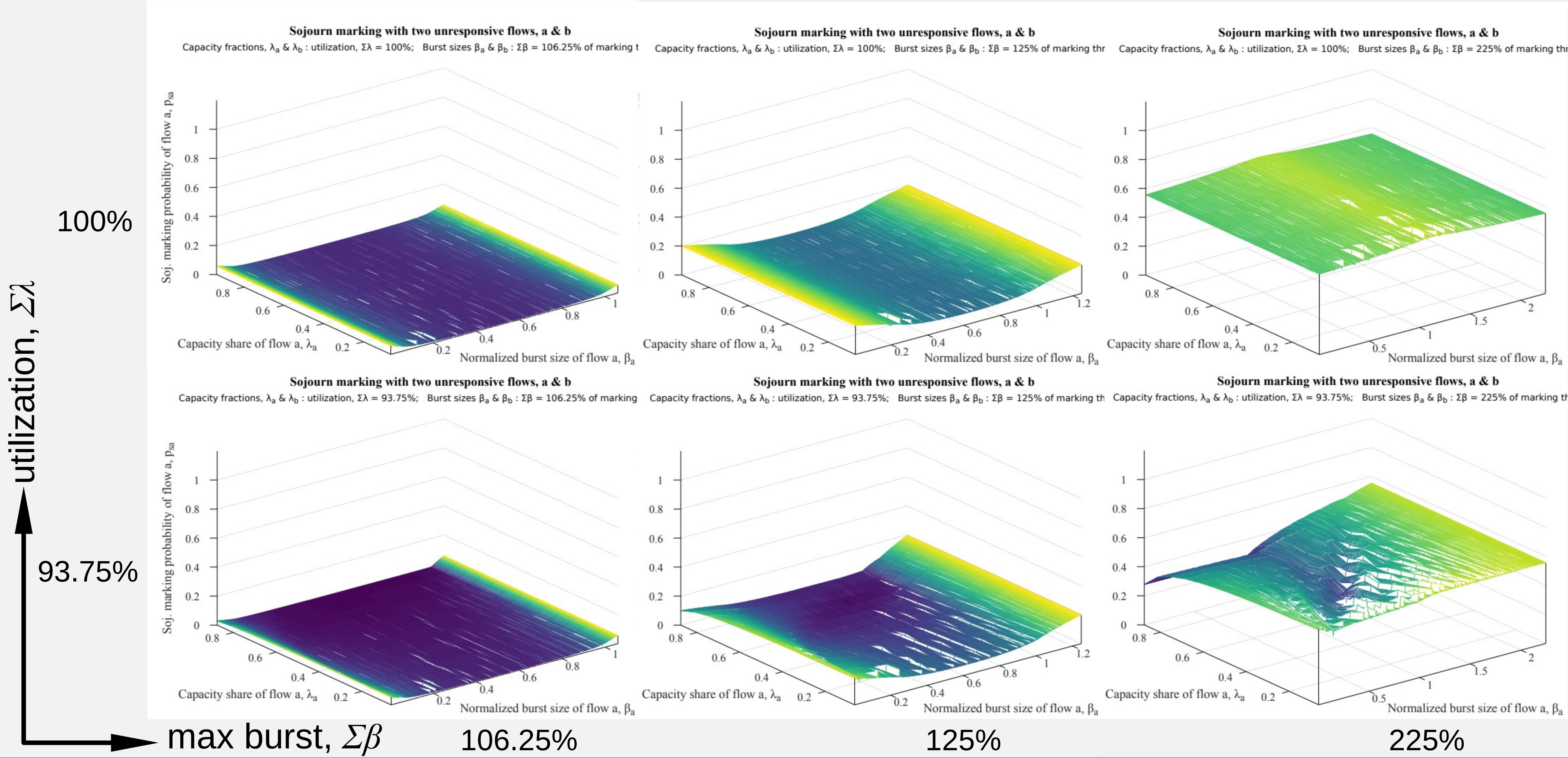


Results – Examples

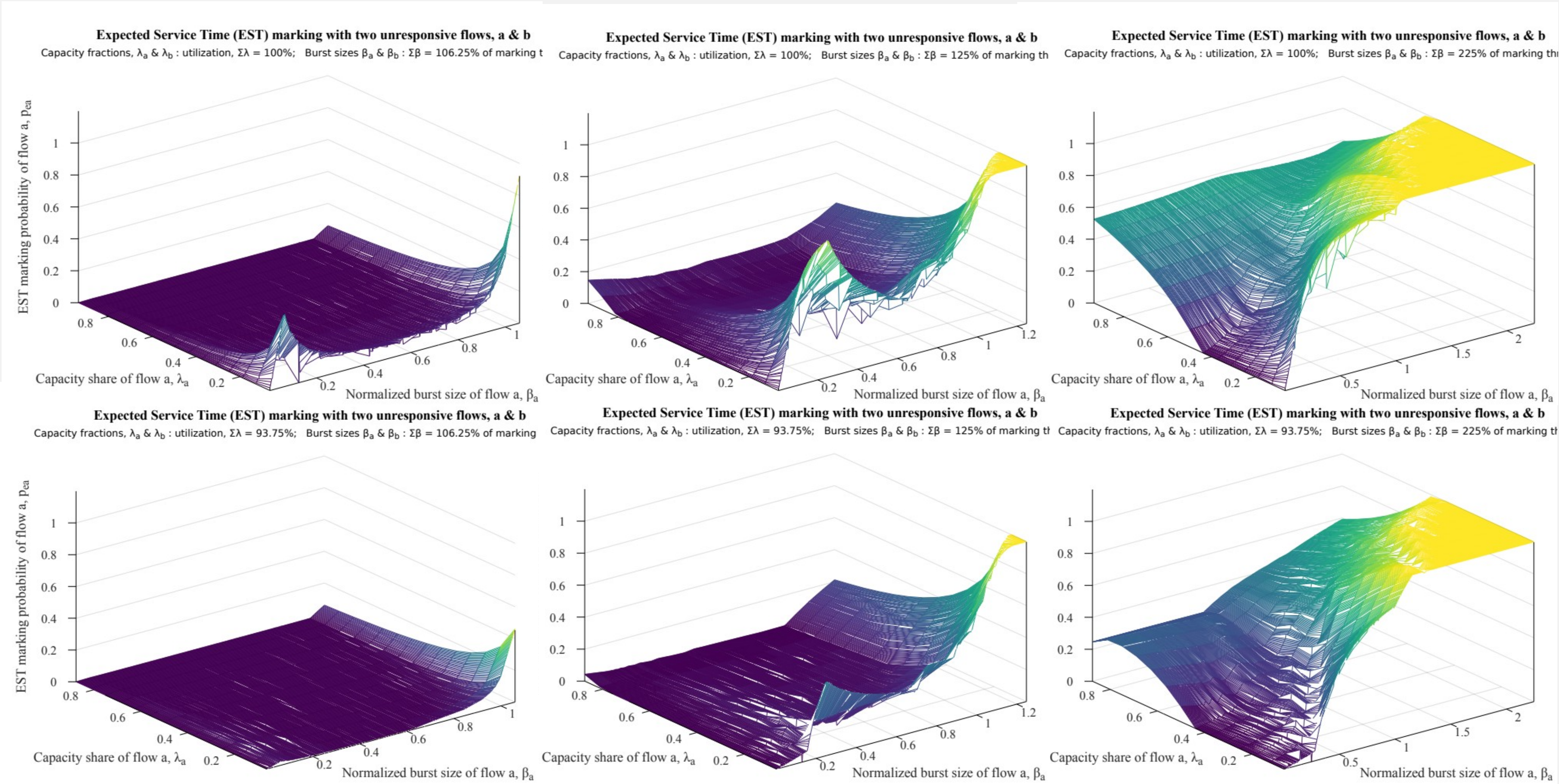


- Detailed 2-D plots (like above)
  - in 4 complementary slide packs
  - 1 for each metric
- choice of 4 metrics
  - p_a : marking probability of flow a
  - $\Delta p = p_a - p_b$
  - $\lambda_a p_a$: marking rate of flow a
  - $\Delta(\lambda p) = \lambda_a p_a - \lambda_b p_b$
- Next 2 slides: 3-D plots
  - using first metric only (p_a)
  - axes will be too small to read,
    but all like the example to the right

Sojourn marking



EST marking



utilization, $\Sigma\lambda$

100%

93.75%

max burst, $\Sigma\beta$

106.25%

125%

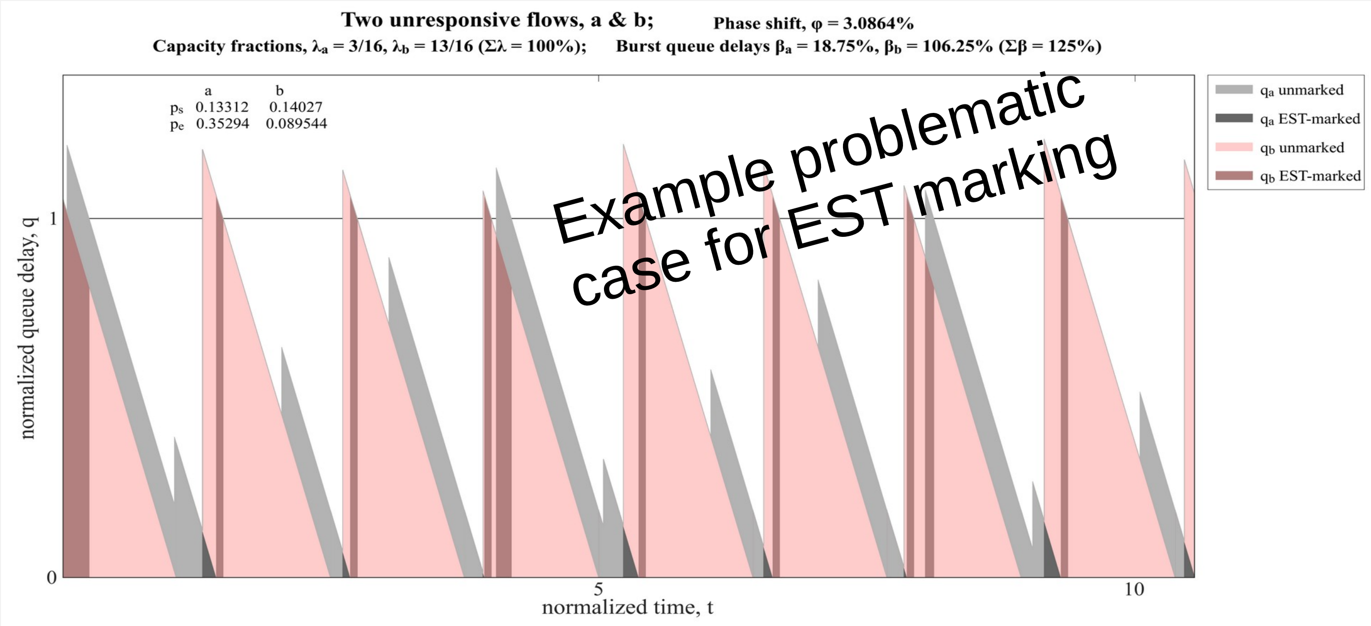
225%

Evaluation

ideal:	increase with burst size β_a?	not decrease with capacity share λ_a?
sojourn[1]	N[2]	Y[1]
EST	Y & N[3]	N

[1] in general, other than high max burst and low utilization
[2] symmetric about the avg burst size
[3] decreases from a peak at $\beta_a = \Sigma\beta - 1$ if low capacity share

Interpretation



- for a grey flow with low capacity share and smaller bursts than pink
- even though grey arrives smoothly, it can only depart in the gaps betw. pink bursts
- variation in these gaps and in how many grey bursts arrive between pink ones is high relative to average grey traffic
- EST marks the residual grey in the q when the next pink burst arrives
- and the pink burst gets pushed back, closing subsequent gaps
- **on average grey fits between pink bursts, but EST punishes grey for all variance**

Next steps

- Design a better marking approach?
  - sthg like $p_e - p_s$?
- Validate model against:
  - ns3 simulation
  - testbed
- Design and evaluate an aggregate policer?

How to Apportion Blame for a Queue with Arrivals in Bursts?

Discussion

and spare slides

Compare 2 marking approaches: Sojourn (s) & EST (e)

Experiment plans

Revisit Original

- ✔Expt 1.1:
  - For a set of fixed capacity shares $\lambda_a + \lambda_b = \Sigma\lambda$ (constant)
  - burst size β: increase β_a, decrease β_b, with $\beta_a+\beta_b = \Sigma\beta$ (constant)
  - measure both marking probabilities, p_s & p_e
  - for each approach, report mean, max & min of each marking metric over a range of phase shifts
- ✔Control expt 1.2:
  - Same as #1.1, with Σλ and Σβ constant
  - but with $\beta_a = \beta_b$ increase λ_a
  - marking should not depend on capacity share, λ
  - (**can visualize this on 3-D plots of expt 1.1**)

- Expt 2.1:
  - Same as #1.1, except hold β_b, while increasing β_a
- ✔Expt 3.1:
  - Same as #1.1 except increase β_a with λ_a
  - (**can visualize this on 3-D plots of expt 1.1**)

- Expt 4? Model packetization or use ns3
- Redesign marking?
- Design & Model aggregate policer

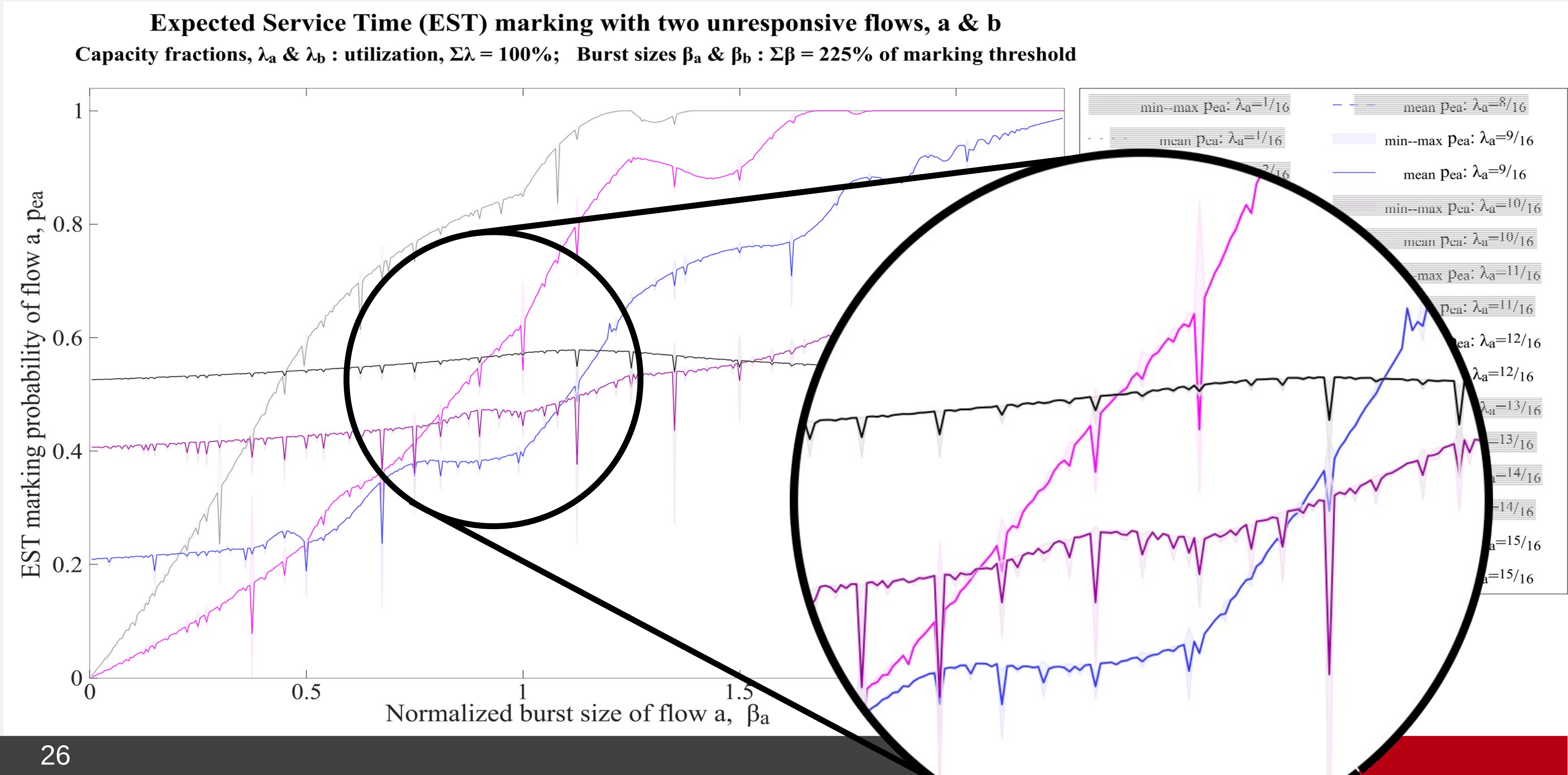
Approach – more detail

Phase Shift

- At each point in the parameter space, start from multiple different phase shifts
- Avoid always including zero as one phase shift
- Record mean, max & min* of marking metrics

* variation is not symmetric, so std. dev. not applicable

Typical Spread of Results over 8 phase samples



Interpretation: phase shift results

- points where spread increases are where the pattern repeats after a few bursts
  - i.e. lowest common integer multiple of the two burst intervals is low
  - then "law of large numbers" doesn't apply
  - unusual coincidences more likely,
    e.g. bursts never precisely coincide
- flows are unlikely to get stuck at these points
  - lower marking causes flow to increase window
- recommend a little randomization of burst sizing – just in case

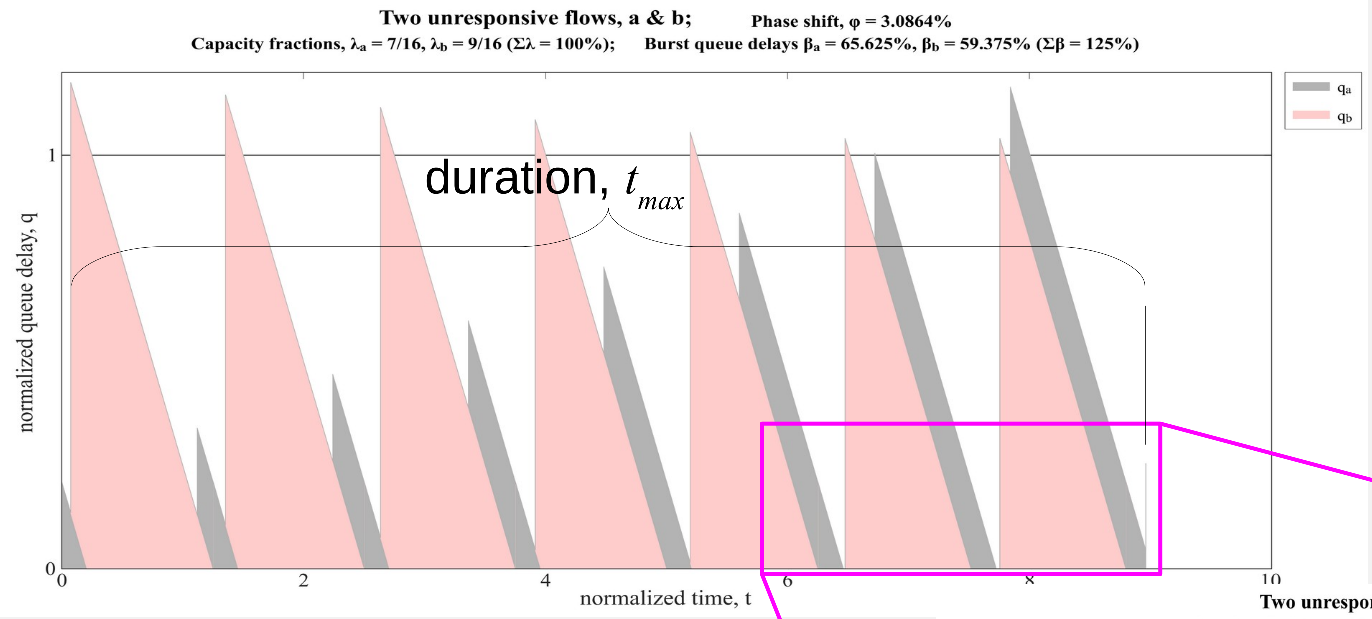
Approach – more detail

Minimal Repeating Pattern

- Duration: lowest common integer multiple
  - of the two burst intervals (not integers themselves)
- Find where to start
  - assume a sufficient standing queue to never go idle
  - start 2nd pass where standing q is smallest
  - challenge: 2 passes without doubling the run time

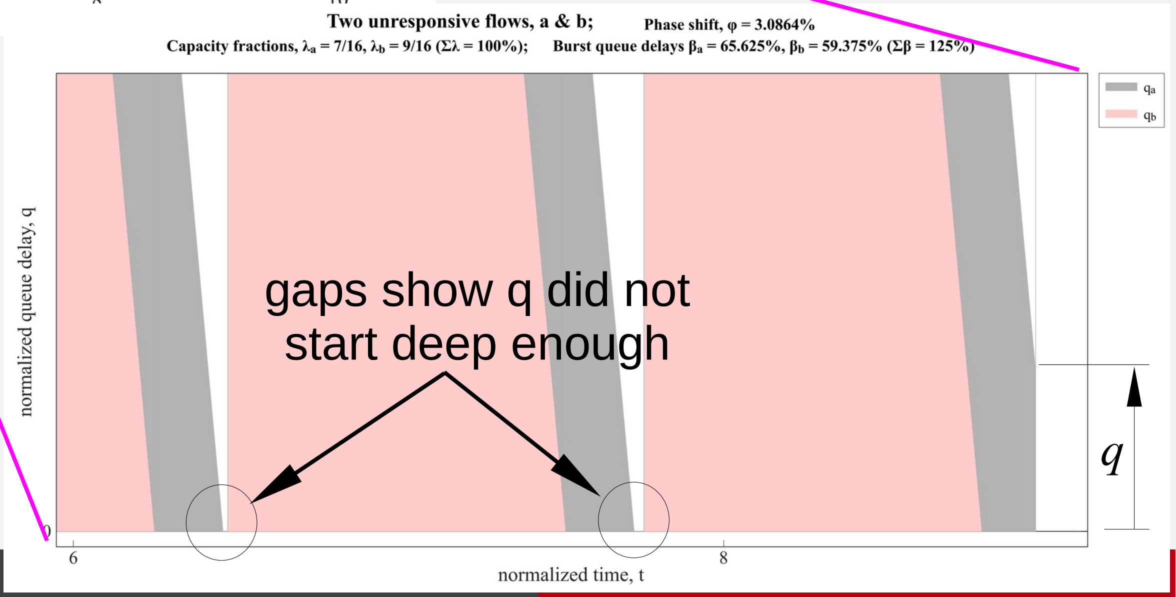
* variation is not symmetric, so std. dev. not applicable

Example of q=0 at the wrong burst



- if set $q = 0$ at wrong burst,
  $q > 0$ at end of duration

  and starting with a slightly larger q
  has a knock-on effect next round



Approach

What if's

- Check the validity of the approach, by investigating alternative avenues
  - increase burstiness of flow a, β_a, while holding β_b at a small selection of const. values
  - increase λ_a & β_a together, related by a selection of factors, e.g. $\Delta\lambda = k\Delta\beta$
  - investigate including zero in the range of phase shifts