# SACRIFICING SOME OPTIMALITY IN AN OVERLOAD CONTROL TO ACHIEVE RAPID RECOVERY

**Ohad Perry**, Northwestern University, Evanston, IL, ohad.perry@northwestern.edu
**Ward Whitt**, Columbia University, New York, NY, ww2040@columbia.edu

We consider how two large service pools, each primarily dedicated to one class of customers but is capable of serving both classes, can help each other in a **time-varying environment** involving periods of overloads. In particular, we assume that the arrival rates and number of agents in either service pool are time dependent, and we seek an optimal routing policy ("sharing" of customers among the two pools), assuming a holding cost is incurred on both queues. We suggest a Fixed-Queue Ratio with Activation-and-Release Thresholds (FQR-ART) control, whose aim is to: (i) Activate sharing in one direction when it is optimal to do so, and then keep the two queues at one of two fixed ratios (depending on the direction of sharing); (ii) switch the order of sharing when the direction of overload changes; and (iii) stop sharing when the system returns to normal loading. The FQR-ART control has the desired properties of being automatic and simple to apply. However, we show that there is a need to sacrifice some optimality so as to avoid undesirable sharing and a resulting oscillatory behavior which can lead to severe degradation in the system's performance. To study this nonstationary system we employ a deterministic fluid approximation, which is described implicitly via a time-inhomogeneous ordinary differential equation.