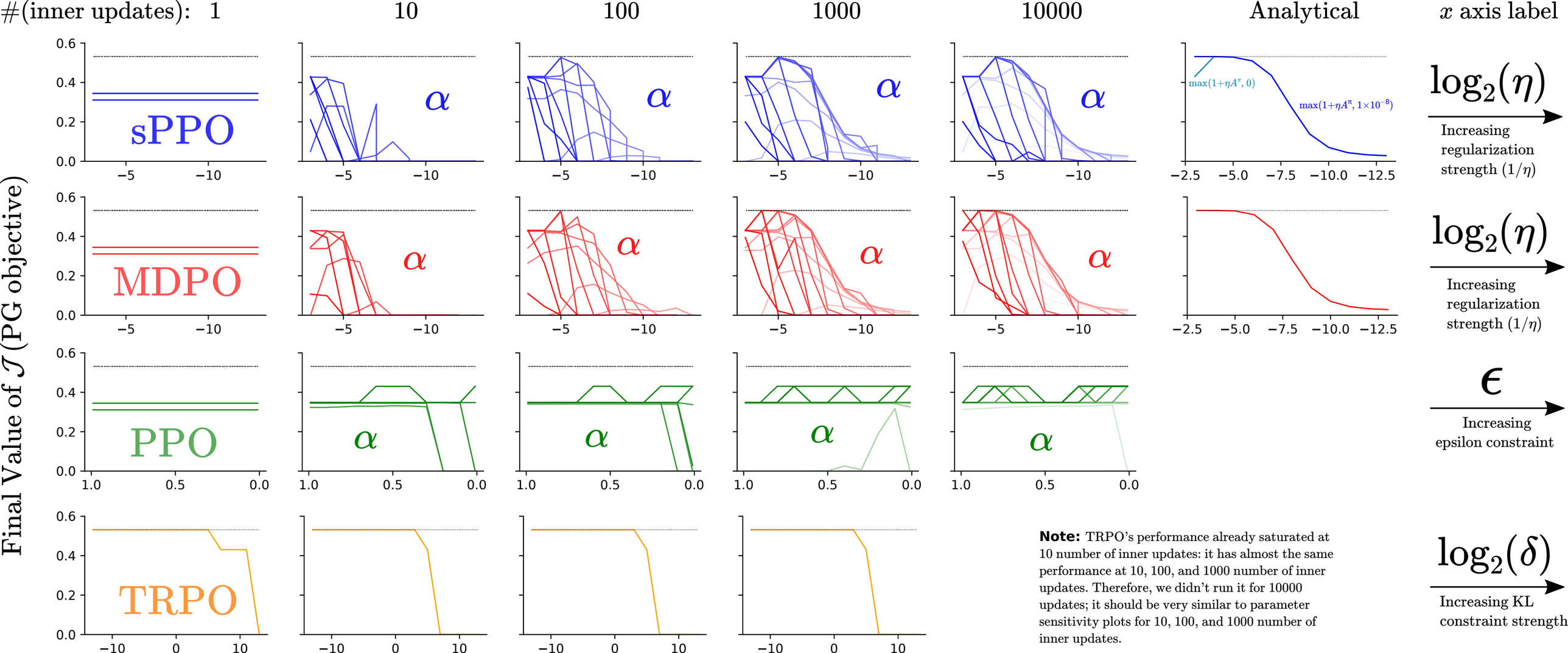

Note: TRPO's performance already saturated at 10 number of inner updates: it has almost the same performance at 10, 100, and 1000 number of inner updates. Therefore, we didn't run it for 10000 updates; it should be very similar to parameter sensitivity plots for 10, 100, and 1000 number of inner updates.