Simulation results with different sets of parameters

Here we provide raw results for a series of parameter we tried to tune to optimize the POMCP algorithm, but that did not seem to bring additional improvement in our framework. Table 4 shows the impact of the trade off parameter α' in a few scenarios for the particle filter.

Table 1: Raw results for the particle filter varying parameters α', n_{search} and K. For each parameter set, $n = 500$ trajectories were simulated. The Value column is the average cost of the trajectories over the n trajectories, and $\hat{\sigma}$ its empirical variance. We also recorded the runtime of optimizing each trajectory (duration column).

Filter	$\pi_{rollout}$	n_{search}	K	α'	Value	$1.96\hat{\sigma}/\sqrt{n}$	duration	duration s.d
particle	π_{mode}	100	100	0.2	161.93	14.79	1730	982
particle	π_{mode}	100	100	0.5	156.01	13.01	1629	913
particle	π_{mode}	100	100	0.8	165.63	13.51	1714	1053
particle	π_{mode}	100	100	0.99	147.24	6.17	1641	971
particle	π_{mode}	100	500	0.2	141.10	9.83	2727	683
particle	π_{mode}	100	500	0.5	141.56	4.72	2771	699
particle	π_{mode}	100	500	0.8	134.98	4.44	2646	685
particle	π_{mode}	100	500	0.99	133.57	3.56	2640	611
particle	π_{mode}	500	100	0.2	146.30	8.27	4617	660
particle	π_{mode}	500	100	0.5	146.82	11.91	4047	670
particle	π_{mode}	500	100	0.8	145.87	13.01	3708	610
particle	π_{mode}	500	100	0.99	140.91	6.35	3454	614
particle	π_{mode}	500	500	0.2	135.99	4.00	5182	598
particle	π_{mode}	500	500	0.5	132.88	4.25	5068	643
particle	π_{mode}	500	500	0.8	136.14	8.08	5271	698
particle	π_{mode}	500	500	0.99	129.42	5.06	4946	724

To allow adaptive selection of the trade off parameter c we tried three dynamic procedures to exploit or explore more depending on our trust in the current patient state. To do so, we define the state entropy as $E_t = \sum_{m=0}^{2} p_m \log(p_m)$, $p_j = \sum_{s=(m,\zeta,u)\in B} \mathbf{1}\{m = j\}$ and $E_{\max} = \log(1/3)$ and consider the following:

- entropy: $\alpha_t = E_t/E_{\max}$

- rev-entropy: $\alpha_t = 1 - E_t/E_{\max}$

- rev-entropy-2: $\alpha_t = 1 - E_t/2E_{\max}$.

However, none of those procedures improved the results, as illustrated in selected examples in Table 2.

Figure 1 illustrates the distribution of the values of α' over 500 simulated trajectories for each adaptative procedure

Table 2: Simulation results with adaptative choice of the exploration/exploitation parameter α'. For each parameter set, $n = 500$ trajectories were simulated. The Value column is the average cost of the trajectories over the n trajectories, and $\hat{\sigma}$ its empirical variance. We also recorded the runtime of optimizing each trajectory (duration column).

| Filter	n_{search}	α'	Value	$1.96\hat{\sigma}/\sqrt{n}$	duration
conditional	100	entropy	138.63	5.06	776
conditional	100	rev-entropy	131.94	3.70	786
conditional	100	rev-entropy-2	133.75	3.70	770
particles	100	entropy	142.17	9.88	2421
particles	100	rev-entropy	143.27	10.34	2473
particles	100	rev-entropy-2	135.28	3.82	2438
conditional	1000	entropy	131.78	4.70	8313
conditional	1000	rev-entropy	131.41	3.45	8432
conditional	1000	rev-entropy-2	132.73	3.56	8332
particles	1000	entropy	133.39	3.74	9994
particles	1000	rev-entropy	135.64	3.74	10047
particles	1000	rev-entropy-2	131.89	4.24	10028
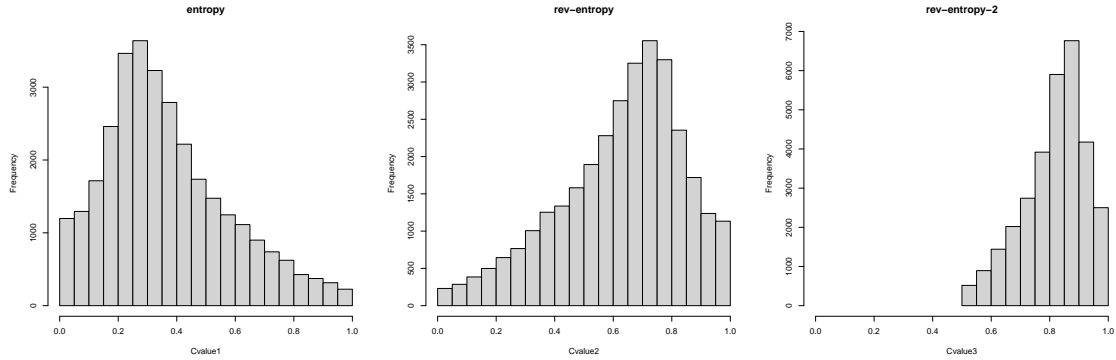
Figure 1: Histogram of values of α' over 500 simulated trajectories for each adaptative procedure

Similarly, we considered an adaptative procedure for selecting the number of exploratory simulations within the POMCP algorithm ($n_s earch$) based on the same entropy. The procedure is as follows: for a fixed $\mathcal{D}$ ($= 10^{-1}$) and a fixed number of particles K ($= 500$), we select two parameters n_{base} and n_{ent} and select n_{search} as $n_{search} = n_{base} + n_{ent} + E_t/E_{\max}$. In this case, we will allocate more simulation budget to cases where our belief on the current patient state is not strong. Table 3 illustrates the unconclusive results.

Belief	n_{base}	n_{ent}	Value	$1.96\hat{\sigma}/\sqrt{n}$	duration	n_{sim}^-
conditional	20	80	136.07	4.66	183	48.2
conditional	50	200	137.96	4.33	511	123.7
particles	20	80	140.33	4.38	1926	48.5
particles	50	200	140.38	8.66	2190	123.8

Table 3: Simulation results (based on 500 trajectories) for the adaptative choice of n_{search} with different initial conditions.

Finally, tables below show results on varying the standard parameters of POMCP for the particle filter (Table **??**) and for the conditional filter (Table 5). The last table shows the impact of the discretisation paramter $\mathcal{D}$.

Belief update	$\pi_{rollout}$	n_{search}	K	α'	Value	$1.96\hat{\sigma}/\sqrt{n}$	duration	duration s.d
particles	π_{mode}	100	100	0.2	161.93	14.79	1730	982
particles	π_{mode}	100	100	0.5	156.01	13.01	1629	913
particles	π_{mode}	100	100	0.8	165.63	13.51	1714	1053
particles	π_{mode}	100	100	0.99	147.24	6.17	1641	971
particles	π_{mode}	100	500	0.2	141.10	9.83	2727	683
particles	π_{mode}	100	500	0.5	141.56	4.72	2771	699
particles	π_{mode}	100	500	0.8	134.98	4.44	2646	685
particles	π_{mode}	100	500	0.99	133.57	3.56	2640	611
particles	π_{mode}	500	100	0.2	146.30	8.27	4617.67	660.87
particles	π_{mode}	500	100	0.5	146.82	11.91	4047.85	670.85
particles	π_{mode}	500	100	0.8	145.87	13.01	3708.93	610.07
particles	π_{mode}	500	100	0.99	140.91	6.35	3454.05	614.83
particles	π_{mode}	500	500	0.2	133	3.70		
particles	π_{mode}	500	500	0.5	131	3.92		
particles	π_{mode}	500	500	0.8	132	3.75		
particles	π_{mode}	500	500	0.99	131	3.70		
particles	π_{mode}	500	500	0.2	135.99	4.00	5182.73	598.84
particles	π_{mode}	500	500	0.5	132.88	4.25	5068.57	643.97
particles	π_{mode}	500	500	0.8	136.14	8.08	5271.29	698.09
particles	π_{mode}	500	500	0.99	129.42	5.06	4946	724.09
particles	π_{mode}	500	500	0.2	145	11		
particles	π_{mode}	500	500	0.5	135	3.82		
particles	π_{mode}	500	500	0.8	139	8.52		
particles	π_{mode}	500	500	0.99	144	12		
particles	π_{mode}	1000	500	0.5	132.52	3.68	12505	914
particles	π_{mode}	500	1000	0.2	133.59	3.41	8600	905
particles	π_{mode}	500	1000	0.5	129.13	3.94	8415	899
particles	π_{mode}	500	1000	0.8	126.88	3.62	8301	917
particles	π_{mode}	500	1000	0.99	131.78	3.87	8245	940
particles	π_{mode}	1000	1500	0.2	130.94	3.88	15526	1257
particles	π_{mode}	1000	1500	0.5	131.79	3.74	15341	1296
particles	π_{mode}	1000	1500	0.8	128.09	3.74	15247	1265
particles	π_{mode}	1000	1500	0.99	132.77	4.84	15217	1288

Table 4: Simulation results for the particle filter

Belief update	$\pi_{rollout}$	n_{search}	K	α'	Value	$1.96\hat{\sigma}/\sqrt{n}$	duration	duration s.d
conditional	π_{mode}	100	100	0.2	133.32	3.53		
conditional	π_{mode}	100	100	0.5	133.09	3.71		
conditional	π_{mode}	100	100	0.8	133.08	3.45		
conditional	π_{mode}	100	100	0.99	133.39	3.60	1076	80
conditional	π_{mode}	100	100	0.2	132.13	3.68		
conditional	π_{mode}	100	100	0.5	131.04	3.46		
conditional	π_{mode}	100	100	0.8	132.90	3.39		
conditional	π_{mode}	100	100	0.99	132.78	4.23		
conditional	π_{mode}	100	500	0.2	133.53	3.61		
conditional	π_{mode}	100	500	0.5	133.34	4.17		
conditional	π_{mode}	100	500	0.8	134.55	4.07		
conditional	π_{mode}	100	500	0.99	129.70	3.34		
conditional	π_{mode}	100	500	0.2	138.17	3.62		
conditional	π_{mode}	100	500	0.5	136.63	5.05		
conditional	π_{mode}	100	500	0.8	133.15	3.60		
conditional	π_{mode}	100	500	0.99	134.97	5.09		
conditional	π_{mode}	500	100	0.2	132.24	3.34		
conditional	π_{mode}	500	100	0.5	130.31	3.47		
conditional	π_{mode}	500	100	0.8	131.45	3.79	5286	413
conditional	π_{mode}	500	100	0.99	131.79	4.46		
conditional	π_{mode}	500	500	0.2	130	3.48		
conditional	π_{mode}	500	500	0.5	129	3.43		
conditional	π_{mode}	500	500	0.8	132	3.54		
conditional	π_{mode}	500	500	0.99	131	3.44		
conditional	π_{mode}	500	500	0.2	132.42	3.54		
conditional	π_{mode}	500	500	0.5	131.49	3.44	5505	403
conditional	π_{mode}	500	500	0.8	132.39	3.50		
conditional	π_{mode}	500	500	0.99	135.21	3.62		
conditional	π_{mode}	500	500	0.2	136	4.49		
conditional	π_{mode}	500	500	0.5	136	4.97		
conditional	π_{mode}	500	500	0.8	132	3.59		
conditional	π_{mode}	500	500	0.99	132	3.39		
conditional	π_{mode}	500	1000	0.2	132.93	3.57		
conditional	π_{mode}	500	1000	0.5	131.47	3.38		
conditional	π_{mode}	500	1000	0.8	133.12	3.48	5476	372
conditional	π_{mode}	500	1000	0.99	130.90	3.60		
conditional	π_{mode}	1000	500	0.2	132.72	4.88		
conditional	π_{mode}	1000	500	0.5	129.99	3.50		
conditional	π_{mode}	1000	500	0.8	130.39	3.50	10994	803
conditional	π_{mode}	1000	500	0.99	132.11	3.67		
conditional	π_{mode}	1000	1000	0.2	133.92	5.71		
conditional	π_{mode}	1000	1000	0.5	129.65	3.59		
conditional	π_{mode}	1000	1000	0.8	128.41	3.38	10987	820
conditional	π_{mode}	1000	1000	0.99	128.61	3.28		

Table 5: Simulation results for the conditional filter

Belief update	n_{search}	K	$\mathcal{D}$	Value	$1.96\hat{\sigma}/\sqrt{n}$
conditional	500	500	0.001	136.25	4.97
conditional	500	500	0.01	131.49	3.44
conditional	500	500	0.1	129.42	3.43
conditional	500	500	0.2	134.43	3.44
conditional	500	500	0.5	129.56	3.32
conditional	500	500	1	132.17	4.16
particles	500	500	0.001	135.22	3.82
particles	500	500	0.01	132.88	4.25
particles	500	500	0.1	131.32	3.92
particles	500	500	0.2	131.26	3.67
particles	500	500	0.5	134.58	4.22
particles	500	500	1	1322	82
conditional	1000	500	0.001	136.00	4.82
conditional	1000	500	0.01	129.99	3.50
conditional	1000	500	0.1	131.30	3.46
conditional	1000	500	0.2	131.20	3.35
conditional	1000	500	0.5	129.12	4.49
conditional	1000	500	1	128.28	4.41
particles	1000	500	0.001	137.80	4.13
particles	1000	500	0.01	132.52	3.68
particles	1000	500	0.1	131.42	3.76
particles	1000	500	0.2	137.73	10.63
particles	1000	500	0.5	135.52	5.13
particles	1000	500	1	1166	79
conditional	500	1000	0.01	131.47	3.38
conditional	500	1000	0.2	130.83	3.22
conditional	500	1000	0.5	130.53	3.15
conditional	500	1000	1	131.4	4.63
particles	500	1000	0.01	131.47	3.38
particles	500	1000	0.2	139.59	9.57
particles	500	1000	0.5	133.36	10.22
particles	500	1000	1		
conditional	1000	1000	0.01	129.65	3.59
conditional	1000	1000	0.2	134.00	5.63
conditional	1000	1000	0.5	128.34	2.95
conditional	1000	1000	1	131.37	3.21
particles	1000	1000	0.2	130.82	3.86
particles	1000	1000	1	965	78

Table 6: Impact of the discretisation paramter $\mathcal{D}$ on the performance of POMCP for a fixed parameter $\alpha' = 0.5$.