Quasi-Deterministic Partially Observable Markov Decision Processes

Camille Besse[1] and Brahim Chaib-draa[1]

Department of Computer Science, Laval University, Quebec, Canada,
{besse,chaib}@damas.ift.ulaval.ca

Abstract. We study a subclass of POMDPs, called quasi-deterministic POMDPs (QDET-POMDPs), characterized by deterministic actions and stochastic observations. While this framework does not model the same general problems as POMDPs, they still capture a number of interesting and challenging problems and, in some cases, have interesting properties. By studying the observability available in this subclass, we show that QDET-POMDPs may fall many steps in the complexity classes of polynomial hierarchy.

1 Introduction

AI planning was initially conceived as a deterministic problem where a sequence of actions has to be decided in order to achieve a goal state with desirable values from an original state. This problem was thoroughly studied in AI with important contributions as A*, GRAPHPLAN, and others [1].

However, numerous problems cannot be modeled using a deterministic formalism. These problems usually involve actions with non-deterministic outcomes or partially observable states that imply to resort to more expressive frameworks like Markov decision processes (MDPs) and partially observable Markov decision processes (POMDPs). While this expressiveness may be useful in many cases, it also comes with a substantial increase in complexity, specially for POMDPs. The gain to use such general frameworks usually does not balance with the cost associated to their resolution. For instance, POMDPs offer one of the most expressive planning under uncertainty model [2], but current algorithms actually scale very poorly as the planning horizon grows and optimal stationary policy does not even exist [3].

Nevertheless, one may not necessary need the complete general framework to model complex problems. Numbers of them are usually deterministic over the actions and over the observations. In fact, these "restrained" models have recently been proposed for planning with incomplete information, e.g. [4], and may also be used for learning partially observable models [5].

Littman briefly discussed these models in his thesis [6], under the name of deterministic POMDPs (DET-POMDPs) for which he obtained some important theoretical results. Littman [6] first showed that a DET-POMDP can be mapped into an MDP with an exponential number of states and then be solved with standard algorithms for MDPs. He also showed that optimal non-stationary policies of polynomial size can be computed in non-deterministic polynomial time and finally that optimal stationary policies can be computed in polynomial space. Since then, up to our knowledge, no publications were made on this subject except [7] that extends these results by defining a specific subclass of DET-POMDPs, that have the so-called *polynomial diameter* property, that can be solved in non-deterministic polynomial time. Bonnet [7] also linked the DET-POMDP framework to the AND/OR tree search algorithms, arguing that this type of algorithm is more efficient than standard POMDP algorithms for this subclass of POMDPs.

According to the role of DET-POMDPs in recent research and motivated by the pursuit of tractable models for decision making under partial observability, we extend the work of Littman and Bonnet in order to bridge a part of the gap between DET-POMDPs and POMDPs, by studying the subclass of POMDPs with deterministic transition but with stochastic observations. We thus present a specific subclass of widely used DET-POMDPs, called quasi-DET-POMDPs (QDET-POMDPs) and show that ε-approximating this subclass falls many steps in complexity in the polynomial hierarchy.

This paper is organized as follows. First, examples of challenging problems are given that motivate our research. Second, a formal definition of the model and the variants are given. In Sect. 4, main theoretical results are described and the complexity of the subclass is presented. Finally, the significance of this work is discussed in Sect. 5.

2 Examples

In this section, examples of problems involving deterministic actions but stochastic observations are presented. While these kinds of problem should have been represented as POMDPs due to their stochastic aspect, reprsenting them as QDET-POMDPs may lead to significant improvements in their resolution.

Robot Navigation: Consider an indoor robot in a $m \times n$ grid that must navigate from an initial position to a goal position while avoiding obstacles using only some noisy sensors on its position. The robot's moves are deterministic but the observation of its current state is distorted by the noise on the sensors. The goal is to find a strategy for guiding the robot to its destination.

Diagnosis: The aim of diagnosis is to identify one of the m states of a system (e.g. a patient) using n noisy binary tests. An instance consists of a $m \times n$ stochastic matrix T where each T_{ij} represent the probability that test j is positive in the state i. The goal is to find the sequence of tests that will identify almost surely the state of the studied system [8].

Sensor Management: Consider multiple sensors situated on a single platform where each sensor can be activated solely (e.g. Figure 1). The problem is to track a concealed or distant target by interrogating the sensors. The target is modeled by a set of states, each state representing a contiguous set of target-sensor orientations over which the scattering physics is relatively stationary. The goal is to find a tracking policy for the target while observing only noisy relative sensor angular positions [9].

All of these problems can be modeled as QDET-POMDPs. Let us now see the formal definition of the proposed framework.

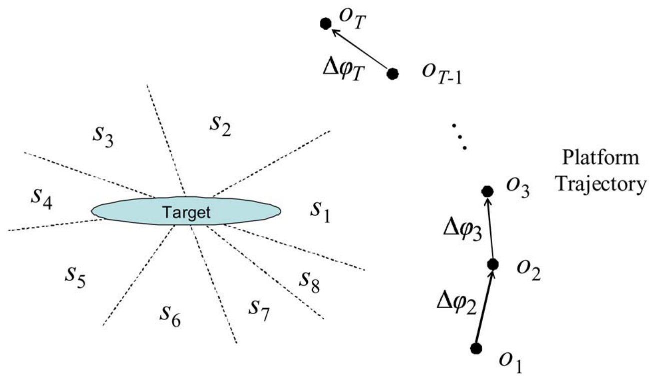

Fig. 1. Multi-aspect sensing of a hidden target. The kth state s_k is a contiguous set of target-sensor orientations over which the scattered fields are approximately stationary ($K = 8$ states are shown). Here T observations are performed, $\{o_1, o_2, ..., o_T\}$, as performed at a sequence of relative sensor angular positions, where $\Delta\varphi_{t+1} = \varphi_{t+1} - \varphi_t$ are orientations. Figure from [9].

3 Model and Variants

Deterministic POMDPs were initially defined as follows [6]:

Definition 1. *[6] A Deterministic Partially Observable Markov Decision Process (*DET-POMDP*) is a tuple $\langle \mathcal{S}, \mathcal{A}, \Omega, \mathcal{T}, \mathcal{O}, \mathcal{R}, \gamma, \boldsymbol{b}^0 \rangle$, where:*

- $\mathcal{S}$ *is a finite set of* states $s \in \mathcal{S}$;
- $\mathcal{A}$ *is the finite set of* actions *of the agent and* $a \in \mathcal{A}$*, denotes an action;*
- Ω *is the finite set of* observations *of the agent and* $z \in \Omega$*, denotes an observation;*
- $\mathcal{O}(z, a, s') : \Omega \times \mathcal{A} \times \mathcal{S} \mapsto \{0, 1\}$ *is the* deterministic observation function *indicating wether or not the agent gets observation z when the world falls in state s' after executing action a;*
- $\mathcal{T}(s, a, s') : \mathcal{S} \times \mathcal{A} \times \mathcal{S} \mapsto \{0, 1\}$ *is the* deterministic transition function *indicating wether or not making action a in state s results in state s';*
- $\mathcal{R}(s, a) : \mathcal{S} \times \mathcal{A} \mapsto \mathbb{R}$ *is the* reward *perceived by the agent when the world falls into state s after agent executes action a;*
- γ *is the discount factor;*
- $\boldsymbol{b}^0$ *is the a priori knowledge about the state, i.e. the initial belief state, assumed non-deterministic.*

The variant proposed by [7] considers a set of absorbing goal states that provide no rewards nor costs but is semantically similar.

Note that the initial *belief state* $\boldsymbol{b}^0$, which describes the different possibilities for the initial state, is crucial. Indeed, if the initial state were known, and since the transition function is deterministic, then all the future states will also be known, and the model reduces to the well studied problem of deterministic planning in AI [1].

Compared to deterministic POMDPs, the proposed extended model presents changes on the observability function. This model, so called Quasi-deterministic Partially Observable Markov Decision Process, is defined as follows:

Definition 2. *A Quasi-deterministic Partially Observable Markov Decision Process (*QDET-POMDP*) is a tuple* $\langle \mathcal{S}, \mathcal{A}, \mathcal{T}, \mathcal{O}, \mathcal{R}, \boldsymbol{b}^0 \rangle$*, where:*

- $\mathcal{S}, \mathcal{A}, \mathcal{T}, \mathcal{R}, \boldsymbol{b}^0$ *are the same as in Definition 1;*
- $\mathcal{O}(z, a, s') : \Omega \times \mathcal{A} \times \mathcal{S} \mapsto [0, 1]$ *is the* observation function *indicating the probability of getting observation* z *when the world falls in* s' *after executing* a*;* *Moreover,* $\forall\, s' \in \mathcal{S},\ a \in \mathcal{A},\ \exists\, z \in \Omega$*, s.t.* $\mathcal{O}(z, a, s') \geq \theta > \frac{1}{2}$*, i.e. the world is minimally observable and the probability of getting one of the observations is lower bounded in each state by at least one half;*

First, let us notice that θ is just a lower bound on observability of each state and thus that in some states the probability the observation can be greater. Notice also that the planning horizon is not set *a priori*. This is due – as we will see in Section 4 – to an interesting convergence property of this model to a ε-deterministic belief state after a fixed number of steps.

Optimality Criteria and Variants

As our goal is to compute a policy that permits our agent to perform *optimally*, we consider the **maxexp** optimality criteria that maximizes the expected discounted reward of a policy. The value of a policy π is thus computed using:

$$V_\pi(\boldsymbol{b}^0) = \mathop{\mathrm{E}}_{s \sim \boldsymbol{b}^0} \left[\sum_{t=0}^{\infty} \gamma^t \mathcal{R}(s^t, \pi(s^t)) \,\middle|\, s^0 = s,\ \pi \right]$$

The variants of the model are related to the observation model:

Unobservable models in which $|\Omega| = 1$ and thus no information is retrieved about the state. This class is a subclass of the so-called conformant problem in planning [10]

Fully Observable models in which $\Omega = \mathcal{S}$ and $\mathcal{O}(z, a, s') = 1$ iff $z = s'$. This class is exactly the classic fully observable MDPs where only the initial state is unknown.

Non-unobservable models in which $|\Omega| > 1$. This class is exactly the complement of unobservable problems. Among this class of problems, we distinguish:

- **Enough-observable models** in which $\Omega = \mathcal{S}$. This class regroups all the linear but noisy observation problems where the state itself is perceived but with an additive noise. This class regroups for example all control problems where the state is perceived through noisy sensors.
- **Factored-observable models** in which $|\Omega| = |\mathcal{X}| \times |\mathcal{D}_{\mathrm{X}}|$. Where $\mathcal{X}$ is the set of state variables and $\mathcal{D}_{\mathrm{X}}$ is the domain of variable X. The state space is then given by $\mathcal{S} = \bigtimes_{\mathrm{X} \in \mathcal{X}} \mathcal{D}_{\mathrm{X}}$. This class is similar to the previous one using additive noise but restrain the number of observations along the "dimensions" of the state space. Indeed, as the state space is assumed structured, the agent can use this structure to learn about at least one dimension at each time step. The previous class of models is a restriction of this class with only one dimension.

General models which includes previous cases, does not assume anything on the observation function.

As the fully observable, the unobservable and the general cases were extensively studied in the literature [1, 2], we will not consider them in the remaining of the paper. However, the enough-observable and the factored-observable cases present an interesting avenue since many of the quasi-deterministic problems mentioned earlier are very often factored or at least enough-observable. We will show in the next section that these problems actually are easier than the general problems by bounding the history needed to identify almost surely the underlying state.

4 QDET-POMDP Theoretical Analysis

In this section, a lower bound on the number of steps to ensure convergence to a certain belief is given and induced complexity results are explained.

As mentioned earlier in the paper, a way to represent compactly the full history of observations during the planning process is the *belief state* [11]. This is a probability distribution over the states that represents the belief of the agent to be in each state through probabilities. We denote by $\boldsymbol{b}^t(s) = \Pr(s|z^t, a^t, \boldsymbol{b}^{t-1})$ the probability of being in state s at step t given that observation z^t was perceived and action a^t was performed in the belief state $\boldsymbol{b}^{t-1}$. This probability is computed using Bayes' rule:

$$\boldsymbol{b}^t(s) = \frac{\mathcal{O}(z^t, a^t, s) \sum_{s' \in \mathcal{S}} \mathcal{T}(s', a^t, s)\boldsymbol{b}^{t-1}(s')}{\sum s'' \in \mathcal{S}\mathcal{O}(z^t, a^t, s'') \sum_{s' \in \mathcal{S}} \mathcal{T}(s', a^t, s'')\boldsymbol{b}^{t-1}(s')} \tag{1}$$

Using a matrix representation, Equation (1) can be rewritten:

$$\boldsymbol{b}^k(s) = \frac{D_k T_{a^k} \cdots D_1 T_{a^1} \boldsymbol{b}^0}{\mathbb{1}^\top D_k T_{a^k} \cdots D_1 T_{a^1} \boldsymbol{b}^0} \tag{2}$$

Where $\boldsymbol{b}^0$ is the initial belief, T_{a^t} are transition matrices according to action a^t, D_i are diagonal matrices with the terms on the diagonal corresponding to the probability to observe z_i given each state, and $\mathbb{1}$ a $|\mathcal{S}|$-dimensional vector of ones.

In order to show the convergence of the belief state to a single state with high probability, let us first state that this probability depends on the number n of succeeded observations among k steps in a non-unobservable context. Nevertheless, non-unobservability is not a sufficient condition to ensure this convergence. Let us now study how n varies regarding to the proposed variants on the observability.

4.1 Enough-Observable models

Enough-observable models ensure that there is only one most likely observation (MLO) in each state and that each state's MLO is not the MLO of any other state:

Definition 3. *An enough-observable* QDET-POMDP *is a* QDET-POMDP *where following assumption holds:*

$$\begin{aligned}
&\exists o_1 \in \Omega,\ \forall a \in \mathcal{A},\ \forall s \in \mathcal{S}^{o_1},\\
&\mathcal{S}^{o_1} = \{s \in \mathcal{S}, o_1 \in \Omega | P(o_1|s, a) > P(o|s, a), \forall o \neq o_1\},\\
&|\Omega| = |\mathcal{S}| \text{ and } |\mathcal{S}^{o_1}| = 1
\end{aligned}$$

Here, $\mathcal{S}^{o_1}$ is the set of states where o_1 is the MLO.

Considering this definition, one can state our first main result:

Theorem 1. *Under the enough-observability assumption,* $\boldsymbol{b}^k(s) \geq 1 - \varepsilon$ *iff*

$$n \geq \frac{1}{2\ln\frac{\nu\theta}{(1-\theta)}} \ln\left[\frac{1-\varepsilon}{\varepsilon}\left(1+\nu^{1-\frac{k}{2}}\right)\right] + \frac{k}{2} \tag{3}$$

Where $\nu = \max_{s,a}\sum_{z\in\Omega} I(\theta > \mathcal{O}(z,a,s) > 0) < |\Omega|$ *the maximum number of "bad" observations that can be perceived in a state.*

Proof (Sketch). In the worst case, the probability of observing the real underlying state is always minimal and equals to θ at each step. Moreover, if the failed observations obtained always support the second most likely state, it results in an increasing of the probability to potentially be in this state. According to Equation (2) and using determinism of transitions, which induces that transition matrices are permutation matrices, one must show that:

$$\frac{\theta^n\frac{(1-\theta)^p}{\nu^p}}{\theta^n\frac{(1-\theta)^p}{\nu^p} + \theta^p\frac{(1-\theta)^n}{\nu^n} + (\nu-1)\frac{(1-\theta)^k}{\nu^k}} \geq 1-\varepsilon \tag{4}$$

Where n is the number of successful observations of the real underlying state and $p = k - n$ the number of failures. The numerator is obtained by obtaining n times a "good" observation and p times a "bad" one during the execution. The denominator sum over all states the same sequence of observation where the first term if for the most likely state, the second term for the second most likely state and the third term for the rest of possible states according to the number of "bad" observations ν. We assume here that the probability to get a "bad" observation is uniform. This assumption is justified by the *maximum-entropy principle* which states that according to the current knowledge, the highest entropy distribution – the uniform in our case – is the best one. Solving[1] this inequality leads to Equation (3). □

Roughly speaking, ν represents also the way the error spreads over the false states.

4.2 Factored models

In a more general way than enough-observable models, factored-observable models ensure that each value of each variable is sufficiently often observed so that the factored state can de determined in a finite number of steps:

Definition 4. *A factored-observable* QDET-POMDP *is a* QDET-POMDP *where following assumption holds:*

- *The state space is* factored *in* μ *state variables:* $\mathcal{S} = \bigtimes_{\mathrm{X}\in\mathcal{X}} \mathcal{D}_{\mathrm{X}}$ *and observations possible are* $\Omega = \bigcup_{\mathrm{X}\in\mathcal{X}} \mathcal{D}_{\mathrm{X}}$.

[1] An extensive derivation of the equations is given in Appendix A.

– *The sum of probabilities of observing one state's variables' real values is lower bounded by $\theta > \frac{1}{2}$.*

This definition implies that, in the worst case, for each state variable, there is a probability $\frac{\theta}{\mu}$ to observe its real value and a probability $\frac{1-\theta}{|\Omega|-\mu}$ to observe anything else. Note also that this definition is a generalization of Definition 3 which is the case $\mu = 1$. This statement leads to the following theorem:

Theorem 2. *Under the factored-observability assumption, $\boldsymbol{b}^k(s) \geq 1 - \varepsilon$ iff*

$$n \geq \frac{1}{2 \ln \frac{(|\Omega|-\mu)\theta}{\mu(1-\theta)}} \ln \left[\frac{1-\varepsilon}{\varepsilon} \left(1 + |\mathcal{S}| - \mu\right) \right] + \frac{k}{2} \tag{5}$$

Proof (Sketch). The proof follows exactly the same arguments as in Theorem 1. □

Once the number n of most likely observation is lower bounded, finding the probability to achieve at least this number is simply an application of the binomial distribution to have at least n successes on k trials:

Corollary 1. *In any* QDET-POMDP *and under Theorem 1* ***or*** *Theorem 2 assumptions, the probability that a belief state $\boldsymbol{b}^k(s)$ is ε-deterministic after k steps is:*

$$\exists s, \Pr(\boldsymbol{b}^k(s) \geq 1 - \varepsilon) = \sum_{i=n}^{k} \binom{k}{i} \theta^i (1-\theta)^{k-i} \tag{6}$$

θ	ν	k	$n \geq$
0.6	3	75	40
0.6	10	59	31
0.6	100	50	26
0.7	3	22	13
0.7	10	19	11
0.7	100	14	8
0.8	3	9	6
0.8	10	6	4
0.8	100	6	4

Table 1. Enough-Observable bound.

θ	μ	$\lvert\mathcal{D}\rvert$	$\lvert\mathcal{S}\rvert$	k	$n \geq$
0.6	2	10	100	84	44
0.6	3	5	125	98	52
0.6	10	6	10^6	112	60
0.7	2	10	100	30	17
0.7	3	5	125	33	19
0.7	10	6	10^6	39	23
0.8	2	10	100	13	8
0.8	3	5	125	16	10
0.8	10	6	10^6	20	13

Table 2. Factored-observable bound.

4.3 Experimentations

To give an idea of the efficiency of the proposed PAC bound, we define $\delta > 0$ such that $\Pr(\boldsymbol{b}^K(s) \geq 1 - \varepsilon) \geq 1 - \delta$. Table 1 and 2 give, for $\varepsilon = 10^{-3}$, $\delta = 10^{-1}$ and different values of θ, ν, μ and the domains' size of variables, the value of the bound on the horizon k and the number of successes needed n given that the probability of having both is above $1 - \delta$. As expected, horizons needed to converge are greater in the factored case than in the enough-observable case for similar state and observation spaces since the agent, at each time step, get less information about the current state. Actually, observations discriminate among subsets of states but not among states themselves like in the previous case. However, as the number of observations is much less than the previous case, current algorithms may have much less difficulty in this type of problems. An empirical study of their difference should be interesting as a research avenue. Let us now derive the worst case complexity from these bounds.

4.4 Impact on complexity

A major implication of Theorems 1 and 2 is the reduction of the complexity of general POMDPs problems when a QDET-POMDP is encountered. Indeed, [12] have shown that finite-horizon POMDPs are PSPACE-complete. However, fixing the horizon T to be constant, causes to complexity to fall down many steps in the polynomial hierarchy [13]. In the case of constant horizon POMDP, one can state:

Proposition 1. *Finding a policy for a finite-horizon-k POMDP, that leads to an expected reward at least C is $\Sigma^{\mathrm{P}}_{2k-1}$.*

Proof. To show that the problem is in $\Sigma^{\mathrm{P}}_{2k-1}$, the following algorithm using a $\Sigma^{\mathrm{P}}_{2k-2}$ oracle can be used: guess a policy for $k - 1$ steps with the oracle and then verify that this policy leads to an expected reward at least C in polynomial time by verifying the $|\Omega|^k$ possible histories, since k is a constant. □

As QDET-POMDPs are a subclass of POMDPs and since fixing δ induces a constant horizon under Theorem 1 or Theorem 2 assumptions:

Corollary 2. *Finding a policy for a QDET-POMDP, under Theorem 1 or 2 assumptions, that leads to an expected reward at least C with probability $1 - \delta$, is $\Sigma^{\mathrm{P}}_{2k-1}$.*

Practically, finding a probably approximatively correct ε-optimal policy for a QDET-POMDP thus implies using a k-QMDP algorithm that computes exactly k exact backups of a POMDP and that then uses the policy of the underlying MDP for the remaining steps (eventually infinite).

5 Conclusion and Future Work

To summarize, we proposed in this paper an extension of the DET-POMDP framework to stochastic observability, called QDET-POMDP, that bridges a part of the gap between DET-POMDPs and general POMDPs. A study of their convergence properties leads to a significant improvement in terms of computational complexity. A sketch of an algorithm is also proposed, opening the avenue of multiple applications. As future work, many avenues can be explored. First, efficient and specific algorithms could be developed that exploit the determinism of transitions and error bounds can be found using the presented bounds on the horizon. Second, an extension to the multiagent case can also lead to major improvements in terms of complexity. Finally, adding some white noise on the transition may help to find more general but still tractable models.

References

1. Nau, D., Ghallab, M., Traverso, P.: Automated Planning: Theory & Practice. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA (2004)
2. Kaelbling, L.P., Littman, M.L., Cassandra, A.R.: Planning and Acting in Partially Observable Stochastic Domains. Artif. Intell. **101**(1-2) (1998) 99–134
3. Madani, O., Hanks, S., Condon, A.: On the Undecidability of Probabilistic Planning and Infinite-Horizon Partially Observable Markov Decision Problems. In: Proc. of Assoc. for the Adv. of Artificial Intelligence. (1999) 541–548
4. Palacios, H., Geffner, H.: From Conformant into Classical Planning: Efficient Translations that May Be Complete Too. In: Proc. of the 17th Int. Conf. on Automated Planning and Scheduling. (2007) 264–271
5. Amir, E., Chang, A.: Learning Partially Observable Deterministic Action Models. J. Artif. Intell. Res. **33** (2008) 349–402
6. Littman, M.: Algorithms for Sequential Decision Making. PhD thesis, Department of Computer Science, Brown University (1996)
7. Bonnet, B.: Deterministic POMDPs Revisited. In: Proc. of Uncertainty in Artificial Intelligence. (2009)
8. Pattipati, K., Alexandridis, M.: Application of Heuristic Search and Info. Theo. to Sequential fault Diagnosis. IEEE Trans. on Syst., Man and Cyb. **20**(4) (1990) 872–887
9. Ji, S., Parr, R., Carin, L.: Nonmyopic multiaspect sensing with partially observable markov decision processes. Signal Processing, IEEE Transactions on **55**(6) (June 2007) 2720–2730
10. Goldman, R.P., Boddy, M.S.: Expressive planning and explicit knowledge. In: Proc. of the 3rd Inter. Conf. on Artif. Intel. Planning Systems. (1996) 110–117
11. Sondik, E.J.: The optimal control of Partially Observable Markov Processes. PhD thesis, Stanford University (1971)
12. Papadimitriou, C., Tsisiklis, J.: The Complexity of Markov Decision Processes. Math. Oper. Res. **12**(3) (1987) 441–450
13. Stockmeyer, L.J.: The Polynomial-Time Hierarchy. Theor. Comput. Sci. **3**(1) (1976) 1–22

6 Appendix A

Proof (Proof of Theorem 1).

$$\begin{aligned}
&\frac{\theta^n \frac{(1-\theta)^p}{\nu^p}}{\theta^n \frac{(1-\theta)^p}{\nu^p} + \theta^p \frac{(1-\theta)^n}{\nu^n} + (\nu-1)\frac{(1-\theta)^k}{\nu^k}} \geq 1-\varepsilon \\
\Leftrightarrow\ &\frac{\nu^n\theta^n(1-\theta)^p}{\nu^n\theta^n(1-\theta)^p + \nu^p\theta^p(1-\theta)^n + (\nu-1)(1-\theta)^k} \geq 1-\varepsilon \\
\Leftrightarrow\ &\frac{\nu^p\theta^p(1-\theta)^n}{\nu^n\theta^n(1-\theta)^p} + \frac{(\nu-1)(1-\theta)^k}{\nu^n\theta^n(1-\theta)^p} \leq \frac{1}{1-\varepsilon} - 1 \\
\Leftrightarrow\ &\nu^{k-2n}\theta^{k-2n}(1-\theta)^{2n-k} + (\nu-1)\frac{\theta^{-n}\nu^{-n}}{(1-\theta)^{-n}} \leq \frac{\varepsilon}{1-\varepsilon} \\
\Leftrightarrow\ &\frac{\nu^{k-2n}\theta^{k-2n}}{(1-\theta)^{k-2n}}\left[1 + (\nu-1)\frac{\nu^{-p}\theta^{-p}}{(1-\theta)^{-p}}\right] \leq \frac{\varepsilon}{1-\varepsilon} \\
\Leftrightarrow\ &(k-2n)\ln\frac{\nu\theta}{(1-\theta)} + \ln\left[1 + (\nu-1)\frac{\nu^{-p}\theta^{-p}}{(1-\theta)^{-p}}\right] \leq \ln\frac{\varepsilon}{1-\varepsilon} \\
\Leftrightarrow\ &(k-2n)\ln\frac{\nu\theta}{(1-\theta)} \leq \ln\frac{\varepsilon}{1-\varepsilon} - \ln\left[1 + (\nu-1)\frac{(1-\theta)^p}{\nu^p\theta^p}\right] \\
\Leftrightarrow\ &(2n-k)\ln\frac{\nu\theta}{(1-\theta)} \geq \ln\frac{1-\varepsilon}{\varepsilon} + \ln\left[1 + (\nu-1)\frac{(1-\theta)^p}{\nu^p\theta^p}\right] \\
\Leftrightarrow\ &(2n-k) \geq \frac{\ln\frac{1-\varepsilon}{\varepsilon}}{\ln\frac{\nu\theta}{(1-\theta)}} + \frac{1}{\ln\frac{\nu\theta}{(1-\theta)}}\ln\left[1 + (\nu-1)\frac{(1-\theta)^p}{\nu^p\theta^p}\right] \\
\Leftrightarrow\ &n \geq \frac{\ln\frac{1-\varepsilon}{\varepsilon}}{2\ln\frac{\nu\theta}{(1-\theta)}} + \frac{1}{2\ln\frac{\nu\theta}{(1-\theta)}}\ln\left[1 + (\nu-1)\frac{(1-\theta)^p}{\nu^p\theta^p}\right] + \frac{k}{2} \qquad (7)
\end{aligned}$$

but since $2 \leq \nu \leq |\mathcal{S}| - 1$, $n > \frac{k}{2}$ and $\frac{1-\theta}{\theta} < 1$,

$$\begin{aligned}
\ln\left[1 + \frac{|\mathcal{S}|-2}{\nu^{k-n}}\frac{(1-\theta)^{k-n}}{\theta^{k-n}}\right] &\leq \ln\left[1 + \frac{|\mathcal{S}|-2}{\nu^{\frac{k}{2}}}\left(\frac{1-\theta}{\theta}\right)^{\frac{k}{2}}\right] \\
&\leq \ln\left[1 + \nu^{1-\frac{k}{2}}\right]
\end{aligned}$$

From which, ensuring Eqn. (3) also ensures Eqn. (7),