

# НАСТРОЙКА ВОЛОКОННОГО ЛАЗЕРА ПРИ ПОМОЩИ АЛГОРИТМА ОБУЧЕНИЯ С ПОДКРЕПЛЕНИЕМ

**Куприков Е.А.<sup>1\*</sup>, Серебrenников К.В.<sup>1</sup>, Кохановский А.Ю.<sup>1,2</sup>**

<sup>1</sup> Новосибирский Государственный университет, г. Новосибирск

<sup>2</sup> Институт точной механики и оптики, г. Санкт-Петербург

\* E-mail: e.kuprikov@g.nsu.ru DOI 10.24412/2308-6920-2023-6-433-434

Волоконные лазеры с синхронизацией мод резонатора на основе эффекта нелинейного вращения поляризации (НВП) являются эффективными источниками ультракоротких оптических импульсов. Однако НВП – лазеры подвержены влиянию внешних условий, и процесс настройки является нетривиальной задачей, требующей достаточного опыта экспериментатора. Данные ограничения оставляют НВП-лазер преимущественно лабораторным инструментом для проведения научно-исследовательских задач. Сегодня интерес многих научных групп направлен на создание алгоритмов автоматической настройки НВП-лазеров [1]. Генетические алгоритмы, алгоритмы на основе глубокого обучения уже продемонстрировали принципиальную возможность осуществления самонастройки волоконных источников ультракоротких импульсов, однако, вопрос о их эффективности остается открытым.

В данной работе мы исследуем эффективность алгоритма глубокого обучения с подкреплением для решения задачи настройки импульсного режима НВП-лазера. Алгоритмы глубокого обучения с подкреплением перспективны с точки зрения создания адаптивных систем управления лазерными системами, устойчивыми к влиянию окружающей среды.

Схема экспериментальной установки представлена на рисунке 1. Лазер состоит из активного эрбиевого волокна, электронного контроллера поляризации (ЭКП), выходного ответвителя, отрезка пассивного волокна SMF-28 и оптического изолятора с заблокированной быстрой осью. Активное волокно накачивалось через сумматор одномодовым лазерным диодом, работающим на длине волны 978 нм. Резонатор волоконного лазера помещался на тепловую пластину внутри термостата для выравнивания температуры лазера. Температура пластины контролировалась ПИД-регулятором и была установлена на 35°C. Цель алгоритма обучения с подкреплением заключалась в изучении поведения лазерной системы для получения стабильного режима генерации посредством управления напряжениями на четырех каналах ЭКП.

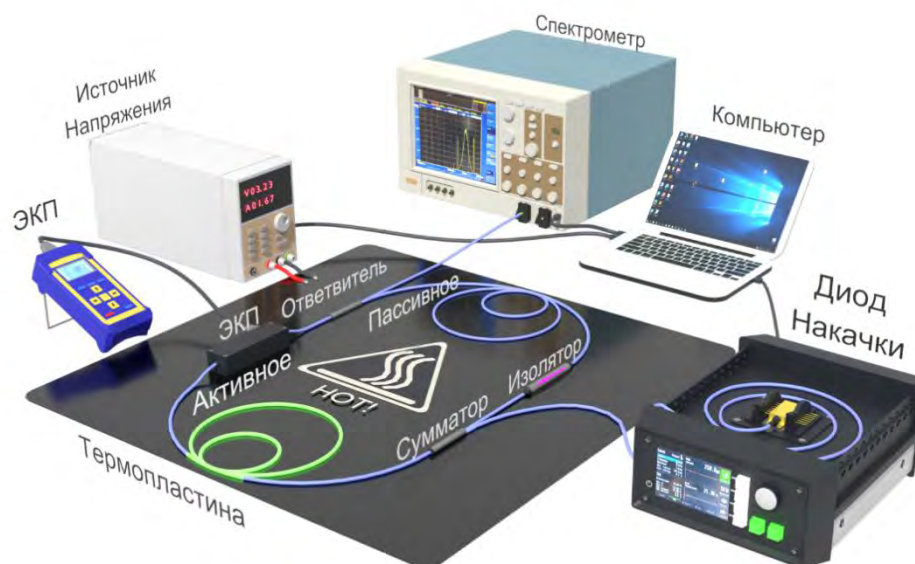


Рис. 1. Экспериментальная установка

Обучение с подкреплением является одним из способов машинного обучения, в ходе которого система (агент) обучается взаимодействовать с некоторой средой [2]. Агент изучает какие действия  $a$  необходимо предпринимать, когда среда находится в состоянии  $s$ , чтобы максимизировать суммарное вознаграждение, выдаваемое средой. В данной работе в качестве среды выступает волоконный лазер с синхронизацией мод на основе эффекта НВП. Состояние  $s$  определяется спектром излучения, который

снимается в диапазоне от 1550 нм до 1615 нм с разрешением 0.1 нм. Действием  $a$  является четыре напряжения ЭКП, которые находятся в диапазоне от -5 до 5 В. В качестве вознаграждения использовалась  $r = \Delta w^2$ , где  $\Delta w$  - ширина спектра излучения. Для обучения агента использовался алгоритм обучения с подкреплением мягкий актер-критик (МАК) [3].

Процесс обучения агента разбит на эпизоды. Каждый эпизод начинается со случайной конфигурации ЭКП и состоит из 100 шагов. На каждом шаге агент получает состояние  $s$  на основе которого выбирает следующее действие  $a$ . Выполнив действие  $a$  агент получает вознаграждение  $r$  и следующее состояние  $s'$ . На каждом шаге алгоритма набор значений  $(s, a, r, s')$  записывается в буфер воспроизведения опыта. Когда размер буфера превышает 256 записей, на каждом шаге также добавляется этап обучения агента на накопленном опыте.

На рисунке 2 представлены результаты работы алгоритма. На рисунке 2а изображено значение награды, которое получает агент на каждом шаге, а на рисунке 2б представлена эволюция спектра в процессе настройки. Из рисунка видно, что для получения режима импульсной генерации агенту потребовалось 25 шагов. И даже в случае потери режима на 34 шаге, агент смог восстановить режим импульсной генерации и продолжал поддерживать его до конца эпизода.

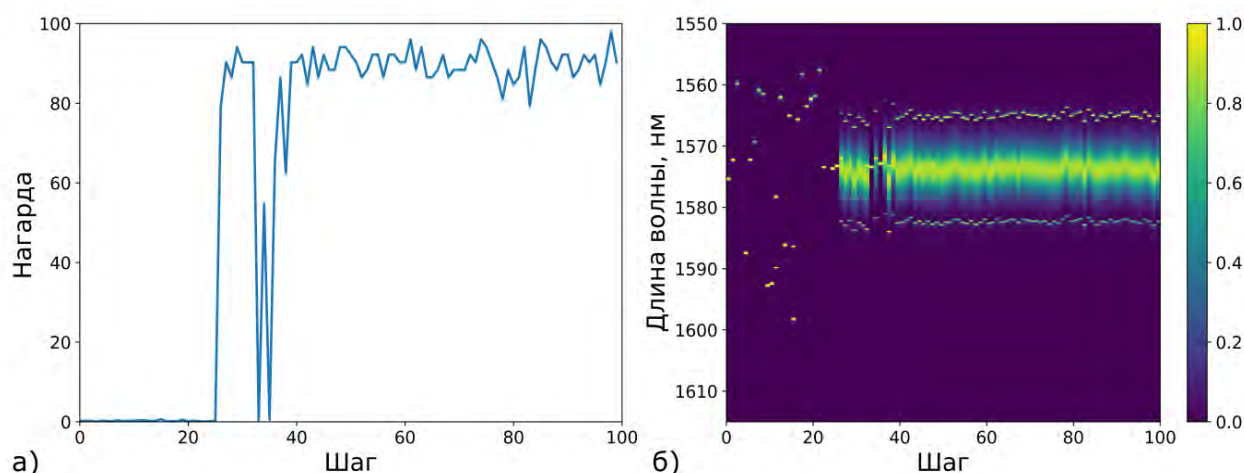


Рис. 2. Настройка волоконного лазера. а) Значение награды на каждом шаге алгоритма. б) Эволюция спектра в процессе настройки

Для тестирования алгоритма мы провели настройку лазера 100 раз со случайных значений контроллера поляризации. В каждом из запусков агенту удалось достичь режима импульсной генерации лазера. Среднее количество действий на настройку лазера составило 21.4. Время одного шага агента составляет 200 мс, что обусловлено временем снятия спектра. Таким образом среднее время настройки составляет 4.28 с.

Исследование выполнено за счет гранта Российского научного фонда (проект 17-72-30006-П).

## Литература

1. Pu G. et al. *Science China Information Sciences*. 63, 1-24 (2020)
2. Sutton R.S., et al, MIT press. (2018)
3. Haarnoja T., et al, *International conference on machine learning*. 1861-1870 (2018)