國立陽明交通大學
National Yang Ming Chiao Tung University

Long-Lasting UAV-aided RIS Communications based on SWIPT

IEEE WCNC
Apr. 12th 2022

Haoran Peng, Li-Chun Wang,
National Yang Ming Chiao Tung University
Geoffrey Ye Li, *Imperial College London*
Ang-Hsun Tsai, *Feng Chia University*

UAV: Unmanned Aerial Vehicle **RIS**: Reconfigurable Intelligent Surface **SWIPT**: Simultaneous Wireless Information and Power Transfer

Background

- RIS can transmit signals without any energy loss.
- UAV-aided Communications have a great potential in providing on-demand communications services for dynamic flash crowds outdoor activities, such as Marathon, concert, etc.

Figure 1. Marathon.

Figure 2. Crowded place.

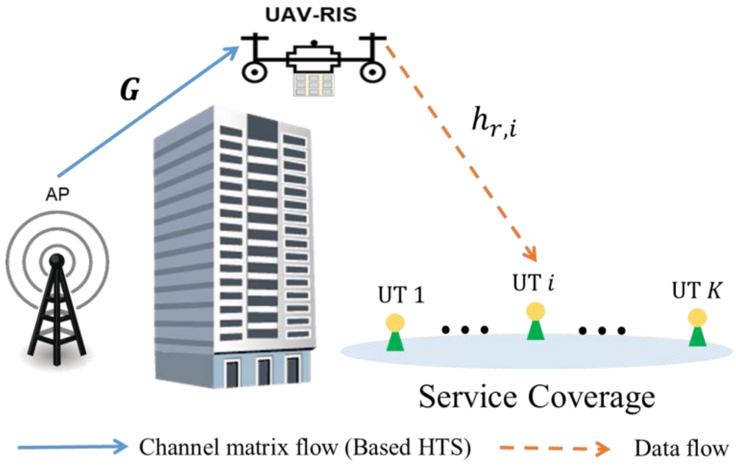

Figure 3. UAV-aided RIS Communications.

UAV mounted with RIS can provide on-demand deployment services in dynamic scenarios.

UT: User Terminal **AP**: Access Point

What's the common limitation in latest research?

	Contribution	Limitation
G. Lee et al. [1]	The RIS system harvests energy from the received RF signals for improving overall energy efficiency.	Dynamic scenarios.
M. Samir et al. [2]	UAVs are integrated with RISs to flexibly deploy RISs in dynamic scenarios.	The limited on-board energy problem of the battery-powered UAV.
Y. Tang et al. [3]	The RIS-aided multiuser MISO SWIPT system was found to enhance the propagation of both the energy signal and the information signal.	The waste of resources when there is a small number of receivers.

[1] G. Lee, M. Jung, A. T. Z. Kasgari, W. Saad, and M. Bennis, "Deep reinforcement learning for energy-efficient networking with reconfigurable intelligent surfaces," in IEEE Int. Conf. Commun. (ICC), Dublin, Ireland, Jul. 2020.

[2] M. Samir, M. Elhattab, C. Assi, S. Sharafeddine, and A. Ghrayeb, "Optimizing age of information through aerial reconfigurable intelligent surfaces: A deep reinforcement learning approach," IEEE Trans. Veh. Technol., vol. 70, no. 4, pp. 3978–3983, Apr. 2021.

[3] Y. Tang, G. Ma, H. Xie, J. Xu, and X. Han, "Joint Transmit and Reflective Beamforming Design for IRS-Assisted Multiuser MISO SWIPT Systems," in IEEE Int. Conf. Commun. (ICC), Dublin, Ireland, Jun. 2020.

Problem statement

Problem

- The onboard battery capacity of UAVs limits the performance and lasting of UAV-RIS communications.
- When there is a small number of UTs, only a few elements are used to transmit signals.

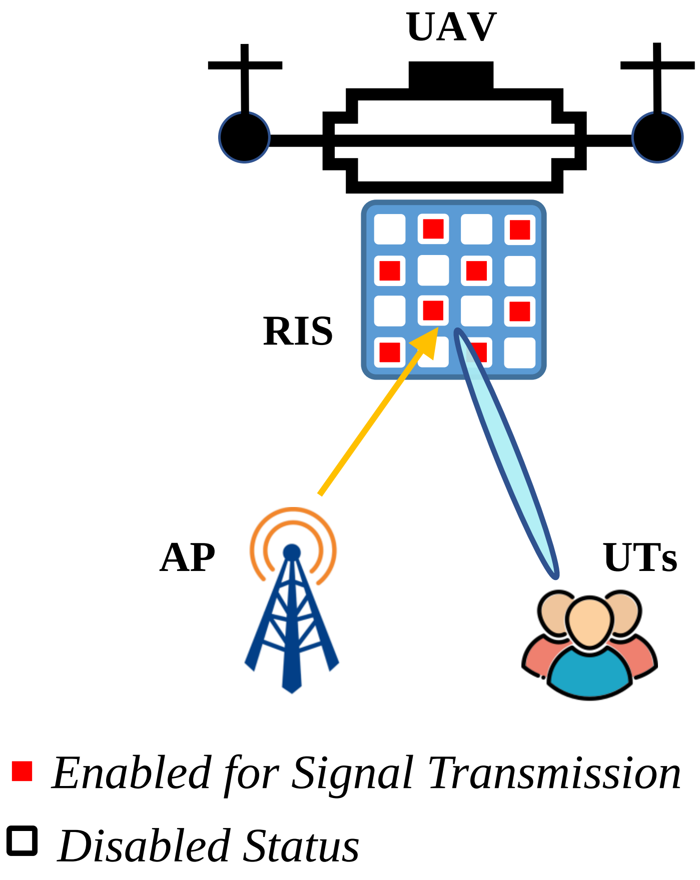

Figure 4. The UAV-aided RIS communications.

Why RIS for lasting the UAV-RIS system?

Motivation

- The RIS system can harvest energy from the received radiofrequency (RF) signal as an extra energy supply for UAVs.

- When the number of UTs is small, using partially reflect-arrays for energy harvesting (EH) can improve energy efficiency.

The proposed dual domains (time & space) EH

Energy Harvesting Phase

All the reflecting units only harvest energy.

Signal Transmission Phase

The array elements in the center for signal reflection, whereas the remainder of the meta-surfaces are used for harvesting energy.

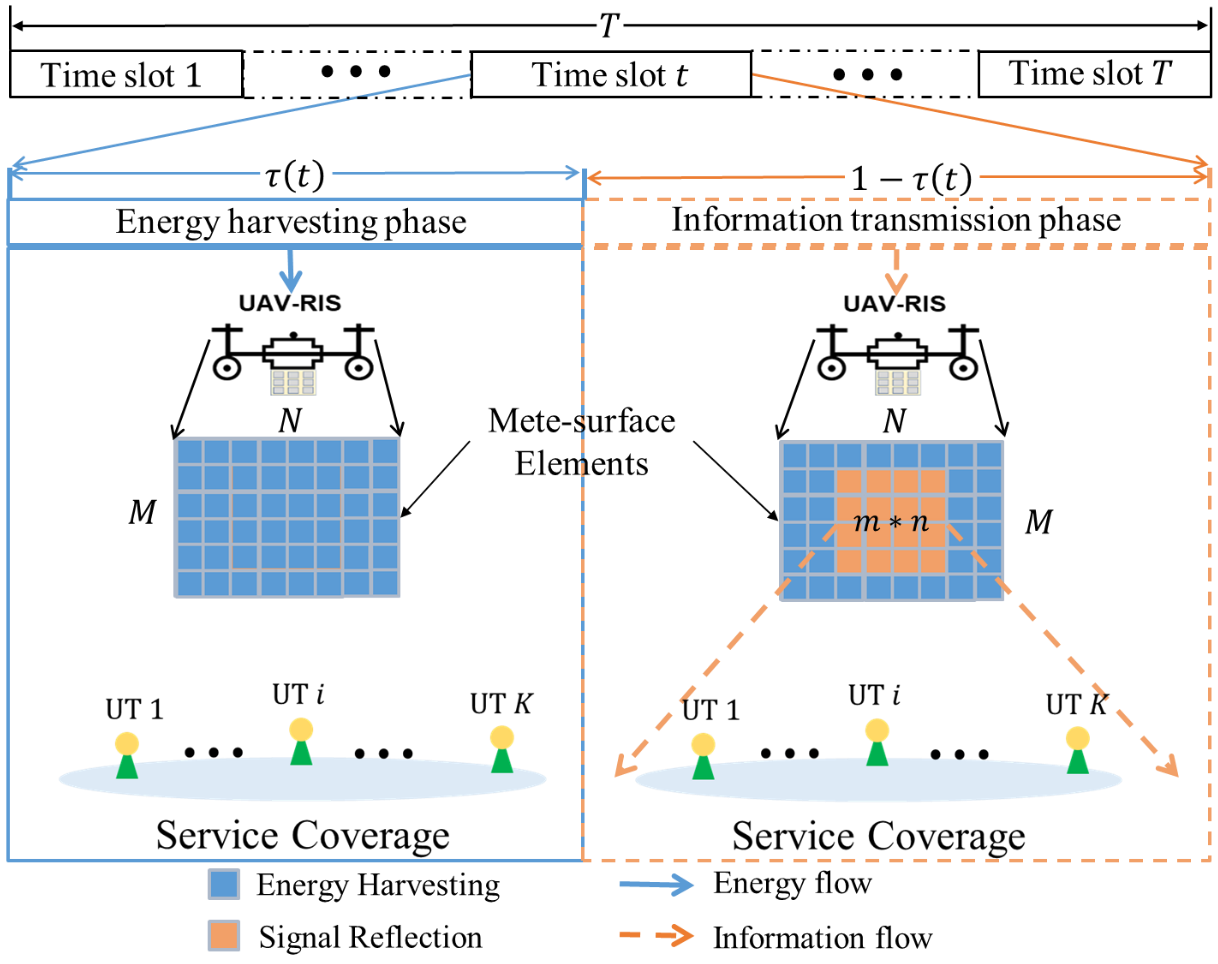

Figure 5. The proposed dual-domain (time and space) EH model.

System model

Resources allocation-based EH model

The harvested energy $E(t)$ of the UAV-RIS system at the time slot t can be expressed as follows:

$h_{i,j}^{H}$ is the channel between the AP and the reflection unit $\mathcal{R}_{i,j}$

$\eta \in (0,1)$ is the EH efficiency.

$$E(t) = \tau(t) \sum_{i=1}^{M} \sum_{j=1}^{N} \eta \left\| h_{i,j}^{H} g_{i,j} \boldsymbol{G} \right\|^2 + (1 - \tau(t)) \sum_{i=1}^{M} \sum_{j=1}^{N} \omega_{i,j} \eta \left\| h_{i,j}^{H} g_{i,j} \boldsymbol{G} \right\|^2$$

$g_{i,j}$ is the channel power gain between the AP and the reflection unit $\mathcal{R}_{i,j}$

G presents the AP's transmit signals.

$$\omega_{i,j} = \begin{cases} 0, \mathcal{R}_{i,j} \in \mathcal{L} & \text{, the element } \mathcal{R}_{i,j} \text{ is adopted to reflect signals} \\ 1, \mathcal{R}_{i,j} \notin \mathcal{L} & \text{, the element } \mathcal{R}_{i,j} \text{ is adopted to harvest energy} \end{cases}$$

System model

AP-RIS-UT channel model

The received signal-to-interference-plus-noise ratio (SINR) at the k-th UT is given by

Θ is the reflection-coefficients matrix of the UAV-RIS system

V_k is the precoding vector for the k-th UT

$$SINR_k = \frac{|\sum_{i=1}^{M}\sum_{j=1}^{N}(1-\omega_{i,j})h_{i,j}^{H}(k)\Theta Z V_k|^2}{\sum_{k'\neq k}|\sum_{i=1}^{M}\sum_{j=1}^{N}(1-\omega_{i,j})h_{i,j}^{H}(k')\Theta Z V_k|^2 + \sigma_k^2}, k, k' \epsilon \mathcal{K}$$

Z represents the channel from the AP to the UAV-RIS system

σ_k^2 is the additive white Gaussian noise

Problem formulation and objective

Objective

This work aims to maximize the total harvested energy of the UAV-RIS system while satisfying communication QoS requirements.

Problem Formulation

Nonconvex optimization problem.

$$\bar{E} = \max_{\tau(t),p,\lambda} \sum_{t=1}^{T} E(t),$$

$s.t.\ C1 : SINR_k(t) \geq SINR_{min}, \forall k \in \mathcal{K}, t \in \mathcal{T},$ → The required minimum SINR on each UT.

$C2 : 0 \leq \tau(t) \leq 1, \forall t \in \mathcal{T},$ → The time constraint of EH phase on each time slot t.

$C3 : 0 \leq p = \sum_{k\in\mathcal{K}} \|V_k\|^2 \leq p_{max},$ → The maximum power control constraint of the AP.

$C4 : 0 \leq \lambda \leq 1.$ → The ratio of the signal reflection units to all the elements in transmission phase.

The proposed DRL-based EH method

The key technology of DDPG

- The actor and critic nets:

$$\pi(s|\delta^{\pi}) \text{ and } Q(s,a|\delta^{Q})$$

- The target actor and target critic nets:

$$\pi'(s|\delta^{\pi'}) \text{ and } Q'(s,a|\delta^{Q'})$$

- DDPG creates an exploration policy via adding a noise sampled from the stochastic noise process N

$$\pi'(s) = \pi(s|\delta^{\pi}) + \mathcal{N}$$

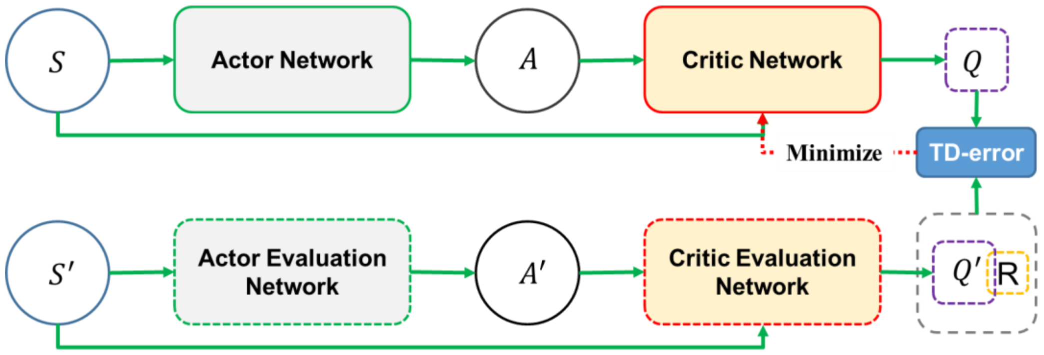

Figure 6. The architecture of the DDPG algorithm.

DRL: Deep Reinforcement Learning **DDPG**: Deep deterministic policy gradient

The proposed DRL-based EH method

The nets update of DDPG

- The actor net updates its policy using the following approximation

$$\nabla_{\delta^{\pi}} J \approx \frac{1}{N_b} \sum_i \left[\nabla_a Q\left(s, a \mid \delta^Q\right) |_{s_i, a=\pi(s_i)} \nabla_{\delta^{\pi}} \pi(s \mid \delta^{\pi}) |_{s_i} \right]$$

- The critic net updates its policy to minimize the loss by

$$L\left(\delta^Q\right) = \frac{1}{N_b} \sum_{i=1}^{N_b} \left(y_i - Q\left(s_i, a_i \mid \delta^Q\right)\right)^2$$

- The DDPG updates weights of the target nets as

$$\delta^{Q'} \leftarrow \psi \delta^Q + (1-\psi)\delta^{Q'},$$

$$\delta^{\pi'} \leftarrow \psi \delta^{\pi} + (1-\psi)\delta^{\pi'},$$

where $\psi \ll 1$ is the learning rate for soft updating actor and critic networks.

Observation and action space

Observation

The transmitted signal from the AP $\boldsymbol{G}$;

The channel power gain $\{g_{i,j}\}_{i,j=1}^{M,N}$;

The RIS reflection-coefficients matrix $\boldsymbol{\Theta}$.

The distance between RIS and the UT d_u.

The baseband equivalent channels from the AP to the UAV-RIS Z.

The baseband equivalent channels from the UAV-RIS to the UT $h_{i,j}^H$.

Action Space

- The length of the EH phase $\tau(t) \in [0,1]$.
- The ratio of the signal reflection units to all the meta-surfaces units $\lambda \in [0,1]$.

Reward design

Reward Design

The positive reward represents the objective of the proposed framework, aiming to maximize the total harvested energy of the UAV-RIS system.

$$\mathcal{R}_t \begin{cases} \bar{E}(t), & SINR(t) \geq SINR_{min}. \\ 0, & SINR(t) < SINR_{min}. \end{cases}$$

The cumulated reward is given by

$$\mathcal{T} = \sum_t \gamma \mathcal{R}_t$$

where γ is the discount factor.

Simulation Parameters

Table 1. Partial values of the simulation parameters.

Symbol	Description	Value
$\mathcal{A}$	Environment constants of LoS	9.61
$\mathcal{B}$	Environment constants of LoS	0.16
η	EH efficiency	0.7
κ	Path-loss at the distance of 1m	-30dB
σ_k^2	Noise power of the k-th UT	-102dBm
φ	Attenuation factor of NLoS	20dB
P_i	Transmit power for the k-th UT	500mW
P_{max}	Maximal AP transmit power	500W
$\bar{\alpha}$	Path-loss exponent for RIS-UT	2.5
α	Path-loss exponent for AP-RIS	3

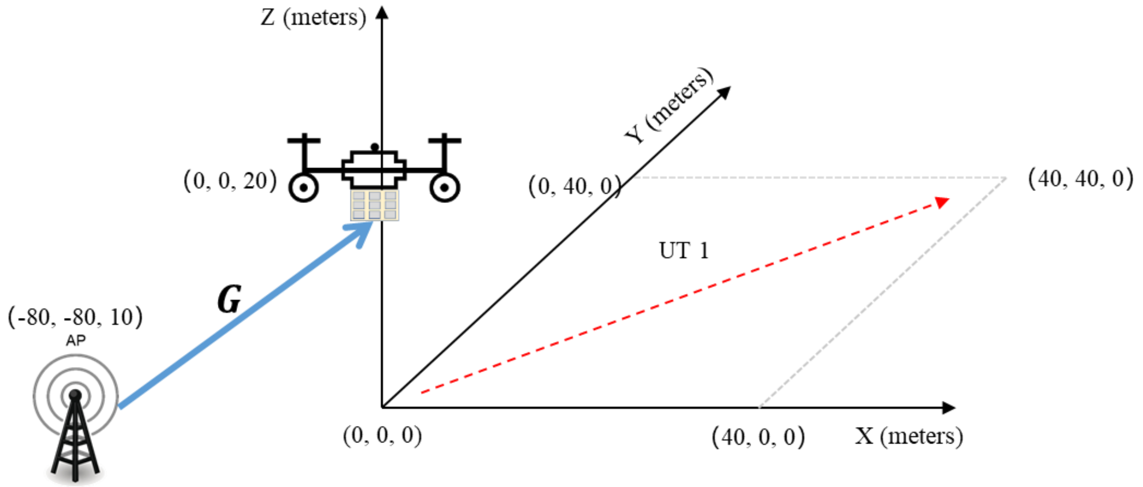

Figure 7. The simulation scenario and the UT trajectory.

The source code: https://github.com/Haoran-Peng/UAV-RIS_EH_DDPG

Performance Comparison of DDPG Under different EH Policy

Table 2. The ratio of the total harvested energy to the overall received RF signal in terms of different optimization methods.

	SWIPT	The proposed dual-domain SWIPT
The Proposed DDPG-based Method	55.32%	57.03%
Exhaust	64.7%	65.1%

Table 2 illustrates the overall ratio of the harvested energy to the received RF signal for the eighty-one steps shown in ***Figure 8***.

Figure 8 shows the ratio of the harvested energy to the received RF signal in each interaction.

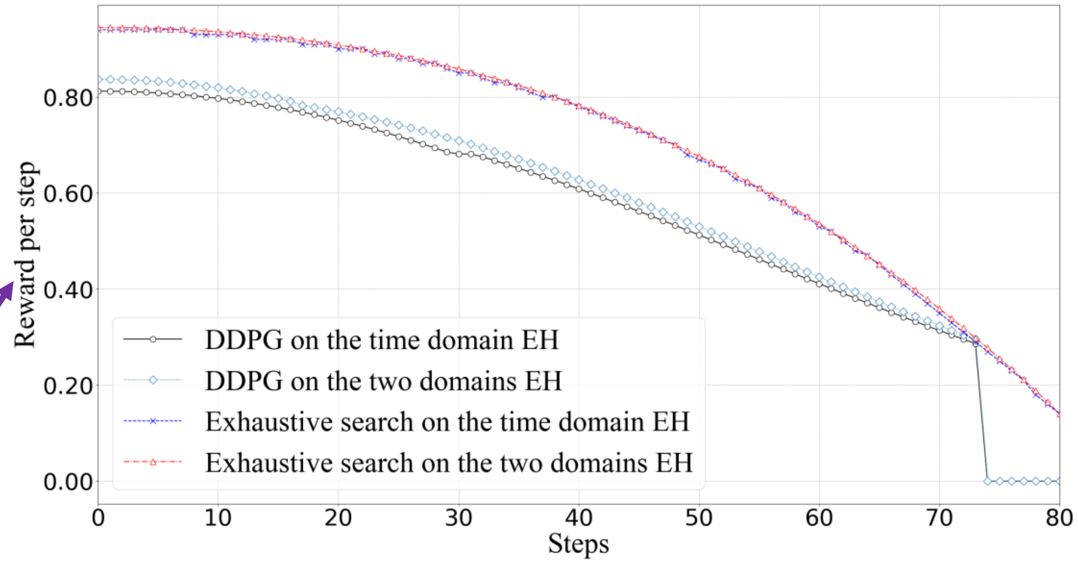

Figure 8. The ratio of the harvested energy to the received RF signal per step.

The proposed dual-domain SWIPT scheme outperforms the conventional SWIPT with respect to different optimization approaches.

Simulation Results

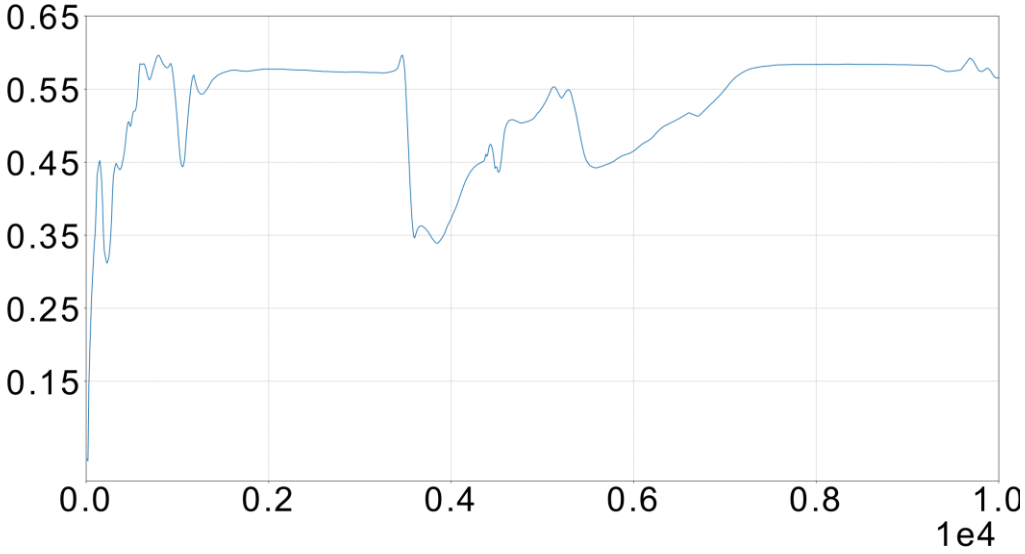

Figure 9. The cumulative rewards per training episode with increasing iterations of the proposed DDPG-base EH method.

The convergence behavior

- ✓ Convergence fast and stable;
- ✓ High performance.

Conclusion and Future Work

The limited energy issue of the UAV-RIS system was studied.

- ✓ The designed dual-domain (Time and Space) SWIPT scheme can improve the EH capability of the UAV-RIS system.
- ✓ The proposed DDPG-based EH method effectively trade-off the time-efficiency and energy-efficiency of the designed dual-domain SWIPT.

Future Work

- ✓ Expanding the proposed long-lasting UAV-RIS system from the single-UT case to the multi-UT case.
- ✓ Finding the fast charging technique and thus improve the performance of the EH scheme.

Q & A

thank you!

Haoran Peng, Li-Chun Wang, Geoffrey Ye Li, and Ang-Hsun Tsai “Long-Lasting UAV-aided RIS Communications based on SWIPT”, IEEE Wireless Commun. Netw. Conf. (WCNC), Austin, TX, Apr. 2022.