Emanuele Mengoli

Background

- PhD in Applied Probability for Wireless Networks.
  Supervised by Prof. F. Baccelli, starting in January.
- Pre-doctoral position at Inria, Mathnet team (Current Position).
- MSc in Computer Science at École Polytechnique.

Competencies

- Machine Learning Theory.
- Computer and Wireless Networks.
- Optimization.

Thesis Project

Stochastic Modeling of 6G Networks, focus on Non-Terrestrial Networks (NTNs).

Use cases

- Integration of NTNs into Terrestrial Cellular Networks.
- Joint Sensing and Communication (JCAS) Networks.
- Cellular Networks equipped with Reconfigurable Intelligent Surfaces (RIS).

Current project

Stochastic modeling of JCAS, in the EU project INSTINCT 6G (WP4).

Dynamic Bayesian Optimization for Improving the Performance of Cellular Networks

Master thesis elaborate by Emanuele Mengoli

Supervisors: Patrick Thiran, Anthony Bardou

Host Institution: École Polytechnique Fédérale de Lausanne (EPFL)

13th AEP, Toulouse,
December 3, 2024

Problem framework

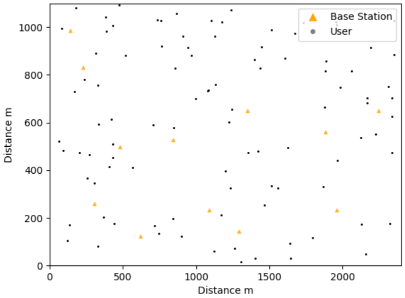


Figure: Example of Dynamic Wireless Cellular Network. BSs position are retrieved from (France Government, 2022).

Key elements

1. Real scenario → Paris. Base Stations (BSs) data (France Government, 2022).
2. Resource Sharing Mechanism (RSM) choice → NOMA.
3. Power control:
   - Task to be optimized.
4. Model User Equipments (UEs):
   - Placement distribution.
   - Birth & Death process.
   - Mobility models.
5. Border effect filtering.

RSM choice

Non-Orthogonal Multiple Access (NOMA)

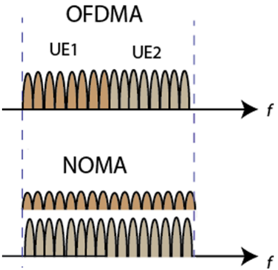


Figure: Spectrum sharing for two users: NOMA vs. OFDMA. Image source: (Kizilirmak and Bizaki, 2016).

Key points

- Full spectrum sharing.
- 2 service region partioning per BS.
- $\mathbf{p}_i = \left(p_i^{(\text{in})}, p_i^{(\text{out})}\right)$
- $P_- \leqslant \|\mathbf{p}_i\|_1 \leqslant P_+$ (BS power interval)
- $p^{(\text{in})} \leqslant p^{(\text{out})}$.
- Successive Interference Cancellation (SIC):
  - Inner UEs decode the signal for outer UEs first and cancel their interference.
  - Outer UEs decode only their own signal and treat the inner UEs' signal as noise.
- Literature contribution (Kizilirmak and Bizaki, 2016; Sen et al., 2010; Gorce et al., 2021; Clerckx et al., 2021).

Optimization framework (1)

Objective

Goal: The natural objective of the model is to **maximize the overall downlink capacity**, yielding the **highest Quality of Service (QoS) for the end-users**.

How? (Let BSs $\mathcal{N} = \{1, \ldots, n\}$, UEs $\mathcal{M} = \{1, \ldots, m\}$)

Ingredient 1: NOMA power constraint for efficient SIC → Define the **parameter space** of the Problem:

$$\underbrace{\left(p_i^{(\text{in})}, p_i^{(\text{out})}\right)}_{\mathbf{p}_i} \rightarrow \underbrace{\left(p_i^{(\text{in})} + p_i^{(\text{out})}, \frac{p_i^{(\text{out})}}{p_i^{(\text{in})} + p_i^{(\text{out})}}\right)}_{\mathbf{x}_i} \tag{1}$$

$$\mathcal{D} = \bigtimes_{w=1}^{n} \mathcal{D}^{(w)} = \left([P_-, P_+] \times \left[\frac{1}{2}, 1\right)\right)^n \tag{2}$$

Ingredient 2: Include the notion of **α-fairness** (Mo and Walrand, 2000) for scalar mapping (Eq. (3)):

$$F_\alpha(c_t) = \begin{cases} \sum_{j=1}^{m} \log c_t^{(j)} & \text{if } \alpha = 1, \\ \sum_{j=1}^{m} \frac{\left(c_t^{(j)}\right)^{1-\alpha}}{1-\alpha} & \text{otherwise,} \end{cases} \tag{3}$$

where $c_t^{(j)}$ denotes $c^{(j)}(\mathbf{x}_t)$, i.e. the **downlink Shannon capacity** for UE j.

Optimization framework (2)

Ingredient 3: Optimal Scheduling $s_j^{(k)}$, $j \in \mathcal{A}^{(k)}$, w.r.t the k-th network partition $(k = 1, \cdots, 2n)$.

$$s_j^{(k)} = \begin{cases} \mathbb{1}_{j=j*} & \text{if } \alpha = 0, \\ \frac{\left(c_t^{(j)}\right)^{(1-\alpha)/\alpha}}{\sum_{i\in\mathcal{A}^{(k)}} \left(c_t^{(i)}\right)^{(1-\alpha)/\alpha}} & \text{otherwise,} \end{cases} \quad (4)$$

with $j* = \arg\max_{j\in\mathcal{A}^{(k)}} c_t^{(j)}$.

All together: the **objective function** (Eq. (5)):

$$f_\alpha\left(\mathbf{x}_t\right) = \sum_{i=1}^{n} \sum_{\mathcal{A}^{(k)}\in\mathcal{A}_i} \sum_{j\in\mathcal{A}^{(k)}} F_\alpha\left(\mathbf{s}^{(k)} \odot \mathbf{c}^{(k)}\left(\mathbf{x}_t^{(\mathcal{N}_i)}\right)\right) \quad (5)$$

with $\mathcal{A}_i$ the regions associated to BS i, $\mathbf{x}_t^{(\mathcal{N}_i)} = \left(\mathbf{x}_t^{(i')}\right)_{i'\in\mathcal{N}_i}$, $\mathcal{N}_i$ indicates the neighborhood of BS i.

Optimization

How to optimize $f_\alpha\left(\mathbf{x}_t\right)$ (Eq. (5)) ? → **Power Allocation Strategies**.

Bayesian Optimization (BO)

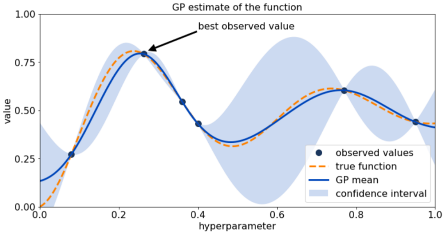


Figure: Example of BO procedure using GP. Image source (Kiss, 2019).

How?

Input: objective function f, acquisition function φ, time horizon T

Init dataset $\mathcal{D} = \varnothing$ and surrogate model $\mathcal{G} = \mathcal{GP}(\mu(\mathbf{x}), k((\mathbf{x}), (\mathbf{x}')))$ where $\mu_0(\mathbf{x}) = 0$ w.l.o.s.

for $t \in [1, T]$ **do**

$\mu_t(\boldsymbol{x}) = \mathbf{k}(\boldsymbol{x}, \mathbf{X})^\top \mathbf{K}^{-1}\mathbf{z}$
$\sigma_t^2(\boldsymbol{x}) = \sigma_0^2(\boldsymbol{x}) - \boldsymbol{k}(\boldsymbol{x}, \mathbf{X})^\top \mathbf{K}^{-1}\mathbf{k}(\boldsymbol{x}, \mathbf{X})$

where $\mathbf{k}(\mathbf{x}, \mathbf{X}) = (k(\mathbf{x}, \mathbf{x}_i))_{i \in [1,t]}$, $\mathbf{K} = (k(\mathbf{x}_i, \mathbf{x}_j))_{i,j \in [1,t]}$, $\mathbf{X} = (\boldsymbol{x}_i)_{i=1}^t$, $\mathbf{z} = (z_i)_{i=1}^t$.

Compute $x_t = \arg\max_{x \in \mathcal{M}} \varphi(x; \mathcal{G})$ *, being $\mathcal{M}$ the domain

Observe $z_t = f(x_t)$ and add (x_t, z_t) to $\mathcal{D}$

end for

* e.g. GP-UCB (Srinivas et al., 2012):
$\varphi_t(\mathbf{x}) = \mu_t(\mathbf{x}) + \beta_t^{1/2}\sigma_t(\mathbf{x})$

Key Components (Bardou et al., 2024)

- **Objective function**: $f : \mathcal{S} \times \mathcal{T} \to \mathbb{R}$, s.t. $\mathcal{S} \subseteq \mathbb{R}^d$, $\mathcal{T} \subseteq \mathbb{R}$.
- **Surrogate Model**: $\mathcal{G} \sim \mathcal{GP}(\mu(\mathbf{x}, t), k((\mathbf{x}, t), (\mathbf{x}', t')))$
  - **Init**: Mean $\mu(\mathbf{x}, t) = 0$, Covariance $k((\mathbf{x}, t), (\mathbf{x}', t'))$
  - Let $k : \mathcal{S} \times \mathcal{T} \times \mathcal{S} \times \mathcal{T} \to \mathbb{R}_+$
- **Kernel Decomposition**:

$$k((\mathbf{x}, t), (\mathbf{x}', t')) = \lambda k_S(\|\mathbf{x} - \mathbf{x}'\|, l_S) k_T(|t - t'|, l_T)$$

  - **Spatial Kernel**: $k_S(\cdot, l_S)$, s.t. $k_S : \mathbb{R}_+ \to [0, 1], l_S > 0$
  - **Temporal Kernel**: $k_T(\cdot, l_T)$, s.t. $k_T : \mathbb{R}_+ \to [0, 1], l_T > 0$
  - **Lengthscale parameters**: λ, l_S, l_T; $\lambda > 0$

W-DBO

Spatio-Temporal BO (2/2)

Key Components (Bardou et al., 2024)

- **W-DBO characteristic**:
  - **MLE Estimation**: Infer (λ, l_S, l_T)
  - **Wasserstein Distance**:

$$W_2\left(\mathcal{GP}_{\mathcal{D}}, \mathcal{GP}_{\tilde{\mathcal{D}}}\right) = \left(\oint_{\mathcal{S}} \int_{t_0}^{\infty} W_2^2\left(\mathcal{N}_{\mathcal{D}}(\mathbf{x}, t), \mathcal{N}_{\tilde{\mathcal{D}}}(\mathbf{x}, t)\right) \, d\mathbf{x} \, dt \right)^{\frac{1}{2}}$$

Measures deviation between two GPs, differing by one removed observation. $\mathcal{D}$ is the original dataset, and $\tilde{\mathcal{D}}$ is the dataset with one observation removed.

Power Allocation Strategies

Random Power Pick

At each optimization round t the vector $\mathbf{x}_{\text{rndm}}^{(t)} \sim \mathcal{U}(\mathcal{D})$.

Constant Power Pick

$\mathbf{x}_{\text{const}} \sim \mathcal{U}(\mathcal{D})$ is sampled just once, during the first round.

W-DBO (Bardou et al., 2024)

- W-DBO instances define different power level suggestions for a common neighborhood: $\mathbf{x}_{\text{wdbo},i}^{(t)} \in \mathcal{D}^{(i)}$, $\forall i \in \mathcal{N}$.
- Median is computed to determine BS power levels: $\mathbf{p}_i = \text{MEDIAN}_{i' \in \mathcal{N}_i}(\mathbf{p}_{i'})$, $\forall i \in \mathcal{N}$.
- Power levels $\mathbf{p}_i$ are retrieved via the bijection in Eq. (1).

UEs Generation

Placement

- UEs distributed uniformly over the network following:
  - Arbitrary number of UEs (defined by testbed user)
  - Poisson Point Process (Baccelli and Blaszczyszyn, 2009)
  - Log-Gaussian Cox process (Møller et al., 1998)

Birth & Death Process

- M/M/1 queue (Kleinrock, 1975; Sankararaman and Baccelli, 2017)

UEs Mobility

Mobility - Trajectory Based Motion

- **if O/D Traffic-flow data available**:
  - Mean-of-transport based probabilistic mobility model (Thesis, Appendix A.2.1)
- **Cluster-based Motion**:
  - hybrid-GMM (Thesis, Appendix A.2.2)
- **User-based Motion**:
  - Biased Random Walk
  - Random Waypoint (RPW) (Panisson, 2014)
  - Lévy flight (Mantegna, 1994)
  - Truncated Lévy flight (Mantegna and Stanley, 1994, 1995; Dybiec et al., 2017)
  - Brownian motion

Our testbed

For our tesbed we used **hybrid-GMM**, **Biased Random Walk**, **RPW**.
How these models work?

Mobility models description

RPW

RPW Model is used in simulations to model the movement of nodes. It involves:

- Moving towards a randomly selected waypoint (Uniformly sampled).
- Pausing at waypoints (Pause time uniformly sampled).
- Selecting new waypoints after pausing (Uniformly sampled).

Biased Random Walk

Biased Random Walk limits erratic movement typical of Random Walks, by introducing a biased direction.

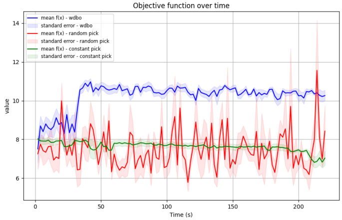
hybrid-GMM

hybrid-GMM combines Gauss-Markov Mobility model and RPW. Modelize transition:

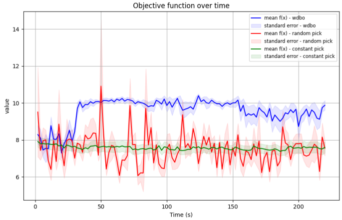
- Temporal-Spatial correlated movement (GMM contribution)
- Pausing time during motion
- Destinations uniformly sampled *

* In our implementation to define way-points, we considered the BS areal density (computed through DBSCAN algorithm (Ester et al., 1996)).

Figure: W-DBO (blue), Random pick (red) and Constant pick (green), with 7 replicates and 12 BSs in an area of 2.64 km^2, approx. 110 optimization rounds. α-fairness equal to 1.



(a) Biased Random Walk

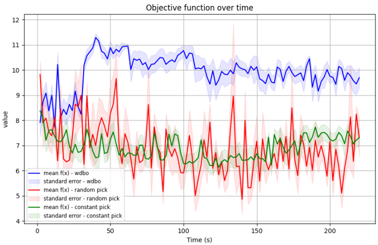


(b) Random Waypoint

Method	**App. W-DBO Time**	**App. Pick Time (Random,Constant)**
Biased Random Walk	85-90 mins	35-40 mins
Random-waypoint	85-90 mins	35-40 mins
Hybrid-GMM	85-90 mins	35-40 mins

Table: Computational Time Comparison - MacBook Pro with a 2.9 GHz Dual-Core Intel Core i5 processor and 16 GB of 2133 MHz LPDDR3 memory. - 2 concurrent threads.

Experimental Results
Obj. Comparison (2/2)

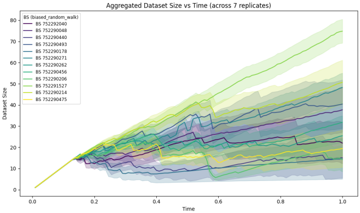


(a) Hybrid-GMM

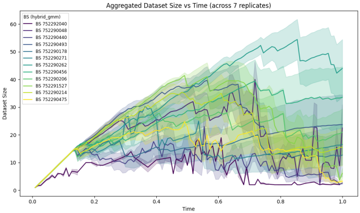
- The Hybrid-GMM model exhibits more dynamic objective function behavior than Biased Random Walk and RPW.
- This is due to moving cluster dynamics, impacting UE density.
- Objective function behavior depends on W-DBO's hyper-parameters, i.e., dataset size, thus temporal lengthscale.

Hyper-parameters

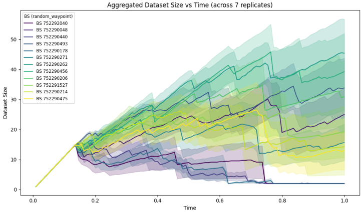
Comparison on disaggregated view - Dataset Size



(a) Biased random walk.



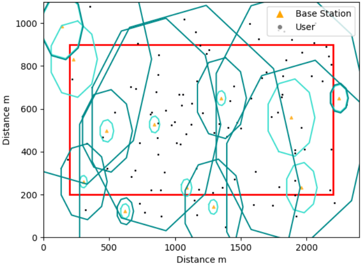
(b) Hybrid-GMM.



(c) Random waypoint.

Hybrid-GMM oscillatory evolutions. RPW and Biased Random Walk smoother curves.

(a) Interactive GIF-visualization. W-DBO power allocation strategy and Biased Random Walk mobility model.



(b) Interactive GIF-visualization. Const-pick power allocation strategy and RPW mobility model.

Figure: Full set of available visualization per motion model and power allocation strategy at NOMA Simulator - Visualizations.

Conclusion

Outcomes

- Developed a simulator to test W-DBO algorithm for power allocation in a dynamic wireless cellular networks.
- Benchmarked W-DBO against Random Pick and Constant Pick strategies.
- Demonstrated that **W-DBO consistently outperforms** these simpler strategies.

Future Extensions

- Expand simulations to **larger-scale settings** with over 30 BSs and more **complex cluster-based motion dynamics**.
- Integrate **real urban mobility traces** (Zhang et al., 2015) for more realistic simulations.
- Extend simulation duration to over 30 minutes for better analysis of real-world scenarios.
- Study a W-DBO variant with **power adjustments constrained** within a confidence interval.

Q&A

Suggestions

Emanuele Mengoli

emanuele.mengoli@inria.fr

References I

F. Baccelli and B. Blaszczyszyn. *Stochastic Geometry and Wireless Networks, Volume I - Theory*, volume 1 of *Foundations and Trends in Networking Vol. 3: No 3-4, pp 249-449*. NoW Publishers, 2009. doi: 10.1561/1300000006. URL https://inria.hal.science/inria-00403039. Stochastic Geometry and Wireless Networks, Volume II - Applications; see http://hal.inria.fr/inria-00403040.

A. Bardou, P. Thiran, and G. Ranieri. This too shall pass: Removing stale observations in dynamic bayesian optimization, 2024. URL https://arxiv.org/abs/2405.14540.

B. Clerckx, Y. Mao, R. Schober, E. A. Jorswieck, D. J. Love, J. Yuan, L. Hanzo, G. Y. Li, E. G. Larsson, and G. Caire. Is noma efficient in multi-antenna networks? a critical look at next generation multiple access techniques. *IEEE Open Journal of the Communications Society*, 2:1310–1343, 2021.

B. Dybiec, E. Gudowska-Nowak, E. Barkai, and A. A. Dubkov. Lévy flights versus lévy walks in bounded domains. *Physical Review E*, 95(5):052102, 2017.

M. Ester, H.-P. Kriegel, J. Sander, X. Xu, et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In *kdd*, volume 96, pages 226–231, 1996.

France Government. Arcep Mobile Data - Base station sites. https://files.data.gouv.fr/arcep_donnees/mobile/sites/2022_T3/2022_T3_sites_Metropole.csv, 2022.

J.-M. Gorce, P. Mary, D. Anade, and J.-M. Kélif. Fundamental limits of non-orthogonal multiple access (noma) for the massive gaussian broadcast channel in finite block-length. *Sensors*, 21(3):715, 2021.

References II

O. Kiss. arxiv: Bayesian optimization for machine learning algorithms in the context of higgs searches at the cms experiment. Technical report, 2019.

R. C. Kizilirmak and H. K. Bizaki. Non-orthogonal multiple access (noma) for 5g networks. *Towards 5G Wireless Networks-A Physical Layer Perspective*, 83:83–98, 2016.

L. Kleinrock. Queuing systems, volume i: Theory, 1975.

R. N. Mantegna. Fast, accurate algorithm for numerical simulation of levy stable stochastic processes. *Physical Review E*, 49(5):4677, 1994.

R. N. Mantegna and H. E. Stanley. Stochastic process with ultraslow convergence to a gaussian: the truncated lévy flight. *Physical Review Letters*, 73(22):2946, 1994.

R. N. Mantegna and H. E. Stanley. Scaling behaviour in the dynamics of an economic index. *Nature*, 376(6535):46–49, 1995.

J. Mo and J. Walrand. Fair end-to-end window-based congestion control. *IEEE/ACM Transactions on networking*, 8(5):556–567, 2000.

J. Møller, A. R. Syversveen, and R. P. Waagepetersen. Log gaussian cox processes. *Scandinavian journal of statistics*, 25(3):451–482, 1998.

A. Panisson. Pymobility v0. 1—python implementation of mobility models, 2014.

A. Sankararaman and F. Baccelli. Spatial birth–death wireless networks. *IEEE Transactions on Information Theory*, 63(6):3964–3982, 2017.

References III

S. Sen, N. Santhapuri, R. R. Choudhury, and S. Nelakuditi. Successive interference cancellation: A back-of-the-envelope perspective. In *Proceedings of the 9th ACM SIGCOMM Workshop on Hot Topics in Networks*, pages 1–6, 2010.

N. Srinivas, A. Krause, S. M. Kakade, and M. W. Seeger. Information-theoretic regret bounds for gaussian process optimization in the bandit setting. *IEEE transactions on information theory*, 58(5):3250–3265, 2012.

D. Zhang, J. Zhao, F. Zhang, and T. He. Urbancps: A cyber-physical system based on multi-source big infrastructure data for heterogeneous model integration. In *Proceedings of the ACM/IEEE Sixth International Conference on Cyber-Physical Systems*, pages 238–247, 2015.