D-ALPS:
Incorporating non-reversible temperature swapping to the Annealed Leap-Point Sampler

Fernando A. Zepeda Herrera Supervisors: Dr. Gareth O. Roberts and Dr. Nick Tawn

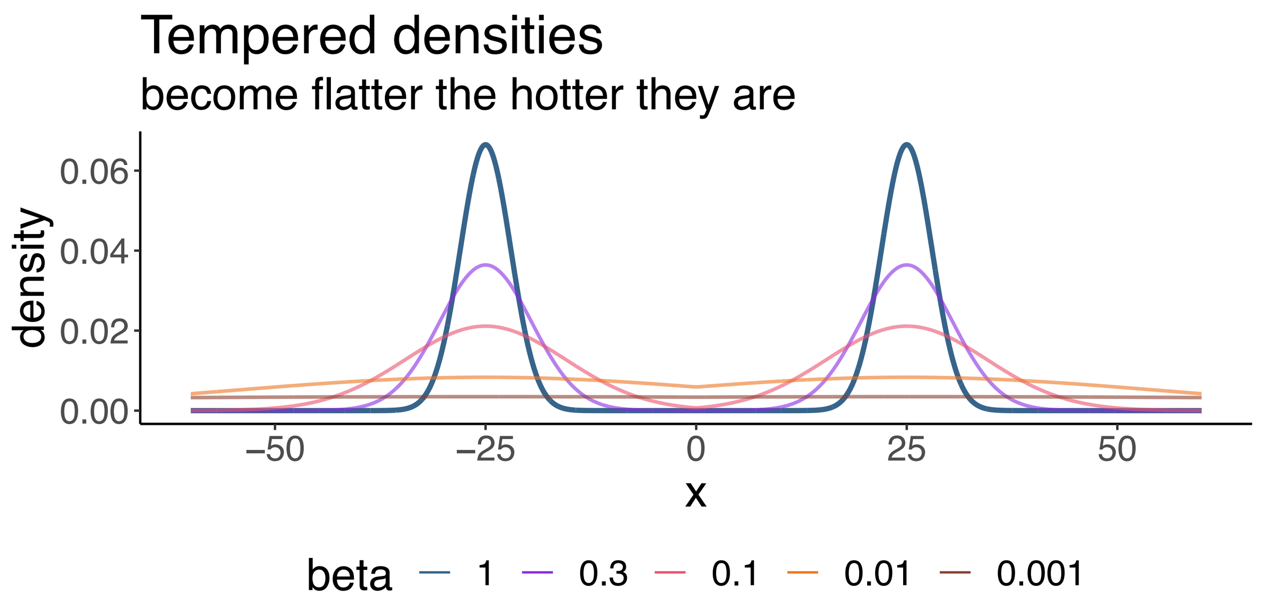
Electronic version at www.fazepher.me/posters/d-alps

WARWICK
THE UNIVERSITY OF WARWICK

Warming Up

- Traditional MCMC algorithms with localized proposals tend to get trapped in a single modal region when facing a multimodal target $\pi(x)$
- **Parallel Tempering (PT) methods** (Geyer, 1991) introduce an (inverse) *temperature schedule* $\beta_1 = 1 > \beta_2 > \cdots > \beta_K > 0$ and an extended target of K *replicas* of x

$$\pi_{PT}(x_1, x_2, \ldots, x_K) \propto \prod_{k=1}^{K} \pi(x_k)^{\beta_k}$$

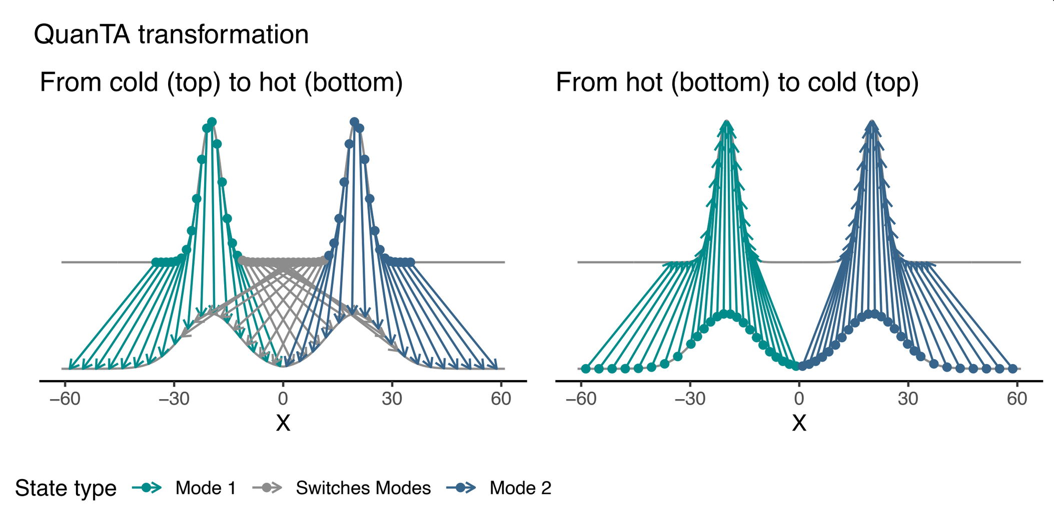


- Each PT cycle contains two types of moves
  1. Parallel within-level exploration (RWM, HMC, etc.)
  2. Temperature Swap according to Metropolis ratio:

$$\frac{\pi_{PT}(x_1, \ldots, x_{k+1}, x_k, \ldots, x_K)}{\pi_{PT}(x_1, \ldots, x_k, x_{k+1}, \ldots, x_K)} = \exp\{(\beta_k - \beta_{k+1})[\log \pi(x_{k+1}) - \log \pi(x_k)]\}$$

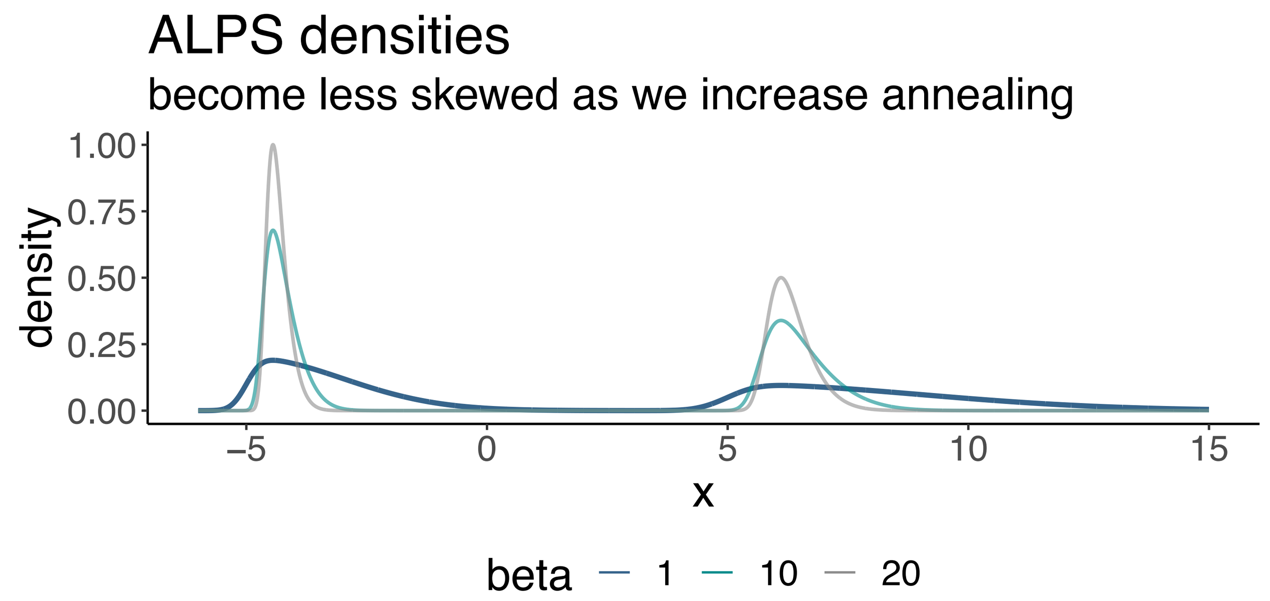
QuanTA

Tawn and Roberts (2019) $T(x, \beta_i, \beta_j; \mu) = \mu + (x - \mu)\left[\frac{\beta_i}{\beta_j}\right]^{1/2}$



Now to the Cold

- Tawn et al (2021) introduced **ALPS**:
  1. **Mode finder** exploration component at a hot temperature $\beta_H \ll 1$
  2. **Parallel Annealing** component with colder schedule $\beta_1 = 1 < \beta_2 < \cdots < \beta_K$
- At colder temperatures ALPS targets[1] look increasingly Gaussian around the mode points $\{\mu_1, \mu_2, \ldots, \mu_m\}$, so use **Independent Sampler at β_K**

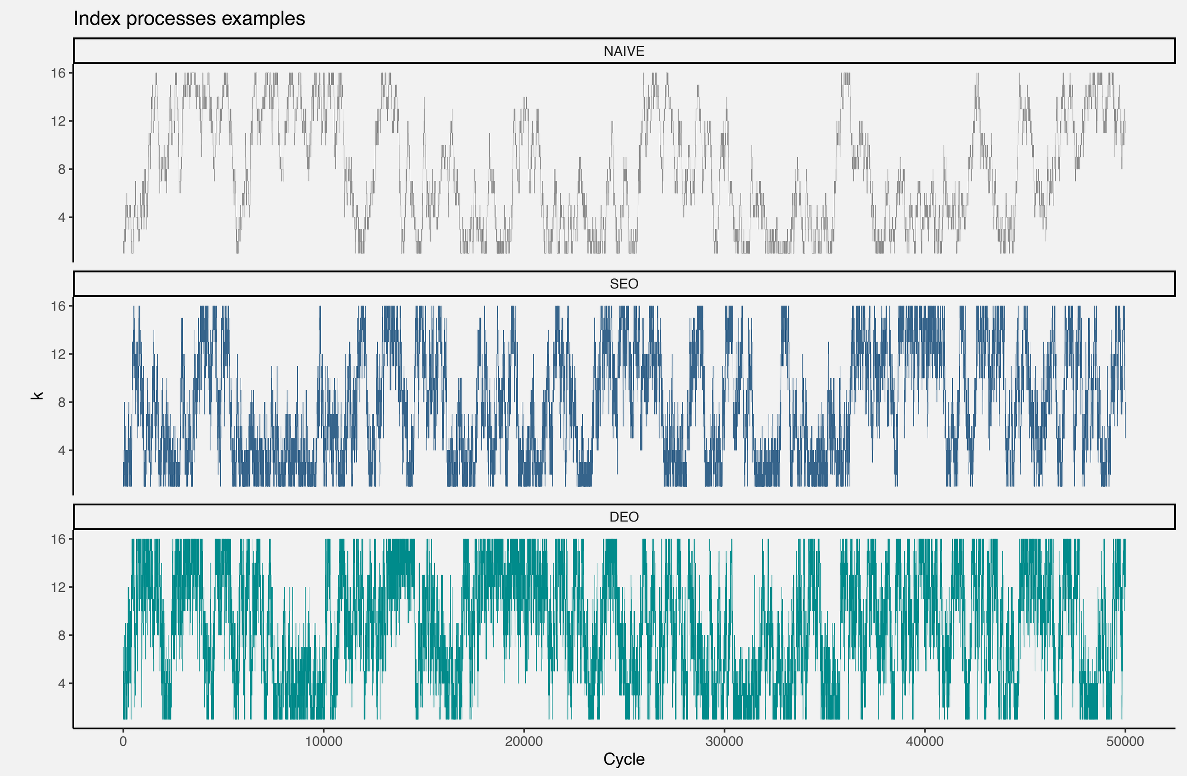


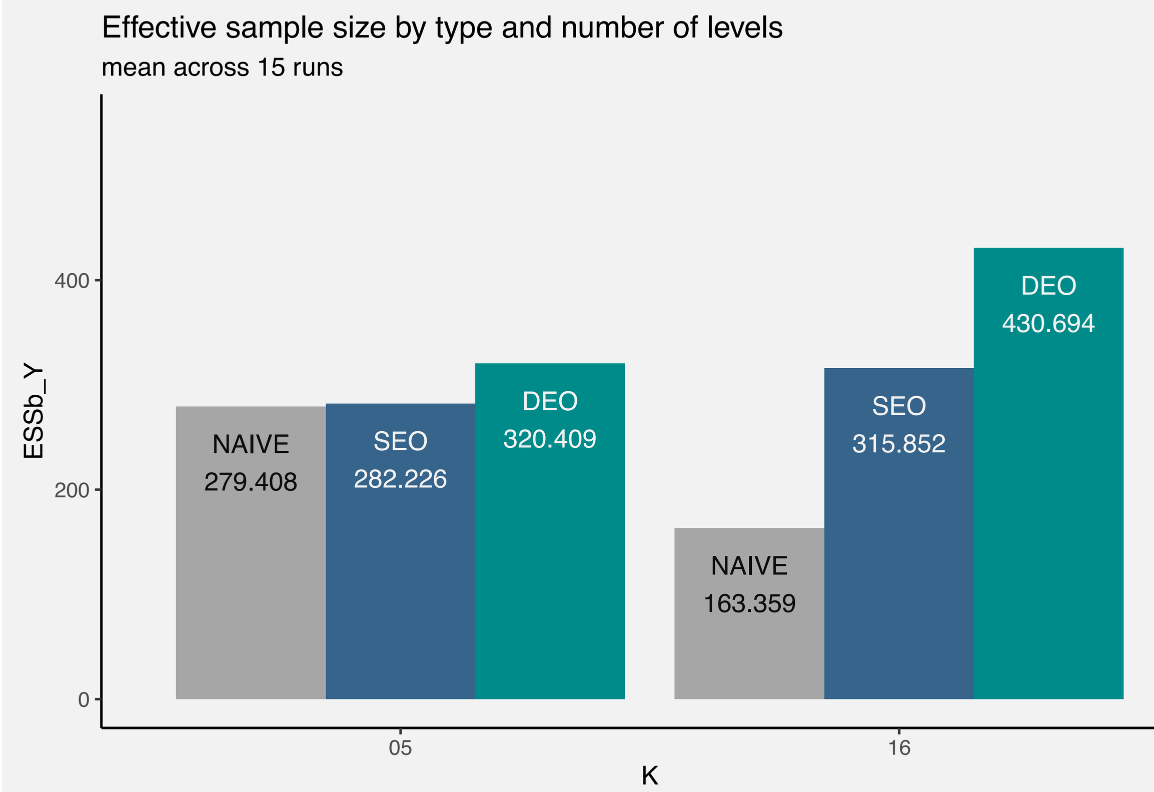
Non Rev. Swapping

- Consider 3 types of swapping:
  1. **Naive-** randomly select a *single* pair
  2. **SEO-** randomly *select block* of Even or Odd pairs
  3. **DEO-** *deterministically alternate* between Even and Odd pairs.

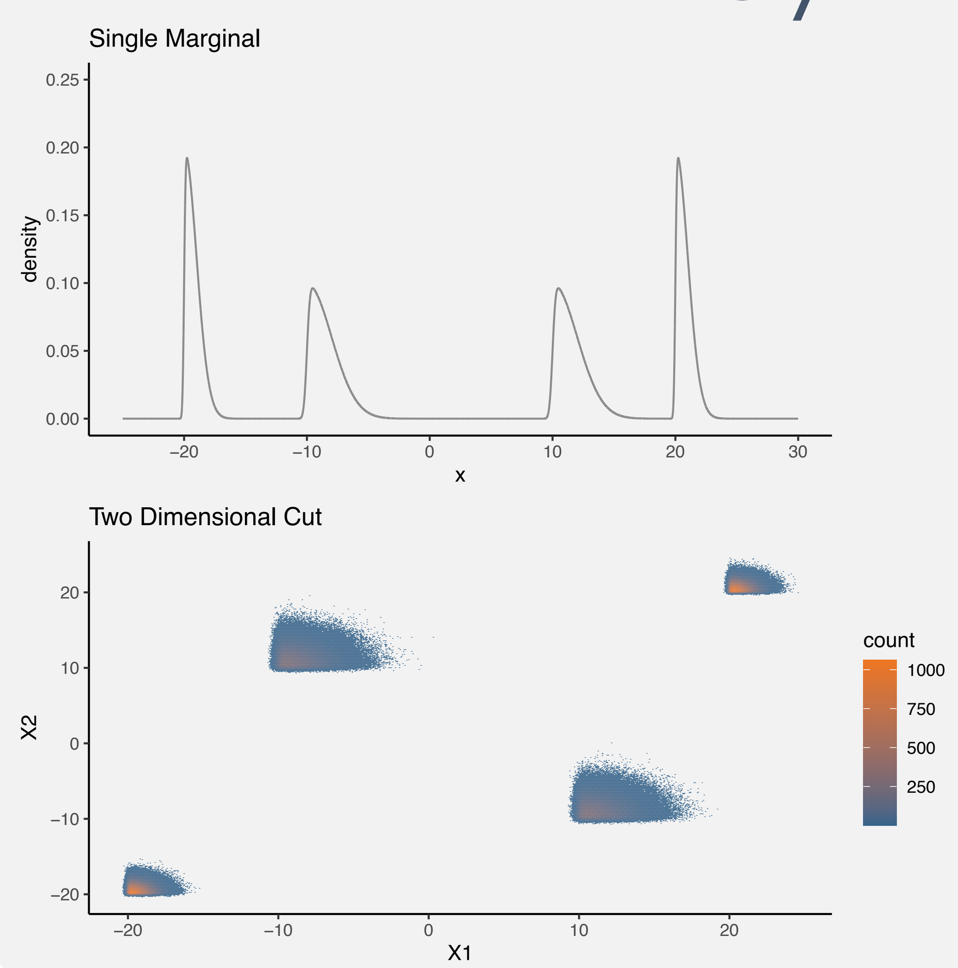
- Syed et al (2022) proved that in PT, non-reversible DEO swapping of Okabe et al (2001) dominates its reversible SEO version
- Both SEO and DEO take advantage of parallelization, but DEO *carries momentum* in the index process

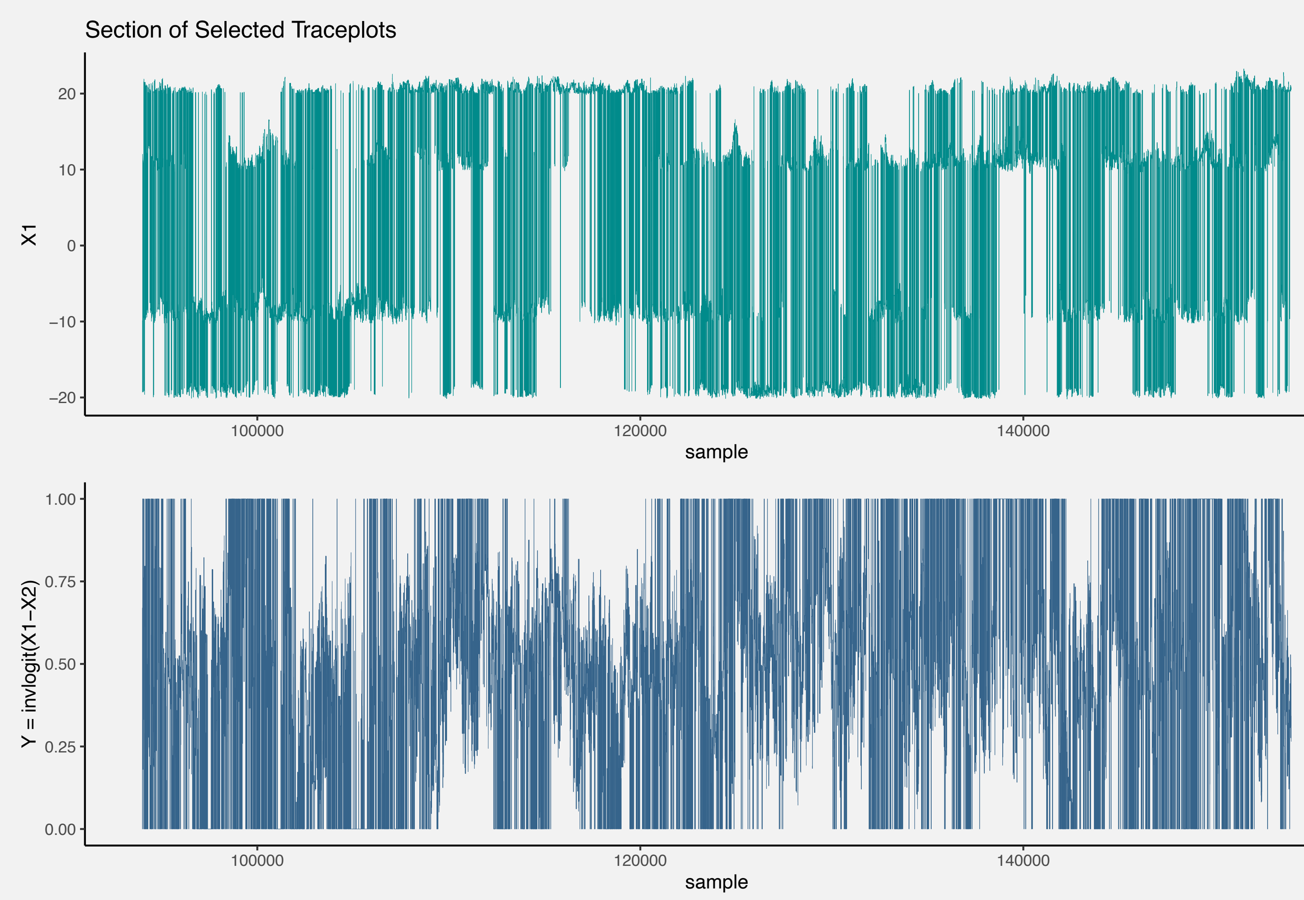
Results





A Toy Example for D-ALPS





[1] ALPS targets are not direct powered-up versions of the original target but use mode-point information to approximately preserve regional weights (full details in paper):

$$\pi_\beta(x) \propto \begin{cases} \pi(x)^\beta \pi(\mu(x,\beta))^{1-\beta}, & \mu(x,\beta) = \mu(x,1) \\ G(x,\beta), & \text{a.o.c.} \end{cases}$$

$$G(x,\beta) = \left(\frac{\tau}{\beta}\right)^{d/2} \pi(\mu(x,\beta))\, |\Sigma(x,\beta)|^{1/2}\, \phi\left(x \,\middle|\, \mu(x,\beta), \frac{\Sigma(x,\beta)}{\beta}\right)$$

References

- Geyer, Charles J. "Markov Chain Monte Carlo Maximum Likelihood." Interface Foundation of North America, 1991. http://conservancy.umn.edu/handle/11299/58440.
- Okabe, Tsuneyasu, Masaaki Kawata, Yuko Okamoto, and Masuhiro Mikami. "Replica-Exchange Monte Carlo Method for the Isobaric–Isothermal Ensemble." *Chemical Physics Letters* 335, no. 5 (March 2, 2001): 435–39. https://doi.org/10.1016/S0009-2614(01)00055-0.
- Syed, Saifuddin, Alexandre Bouchard-Côté, George Deligiannidis, and Arnaud Doucet. "Non-Reversible Parallel Tempering: A Scalable Highly Parallel MCMC Scheme." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 84, no. 2 (2022): 321–50. https://doi.org/10.1111/rssb.12464.
- Tawn, Nicholas G., Matthew T. Moores, and Gareth O. Roberts. "Annealed Leap-Point Sampler for Multimodal Target Distributions." *ArXiv:2112.12908 [Stat]*, December 23, 2021. http://arxiv.org/abs/2112.12908.
- Tawn, Nicholas G., and Gareth O. Roberts. "Accelerating Parallel Tempering: Quantile Tempering Algorithm (QuanTA)." *Advances in Applied Probability* 51, no. 3 (September 2019): 802–34. https://doi.org/10.1017/apr.2019.35.