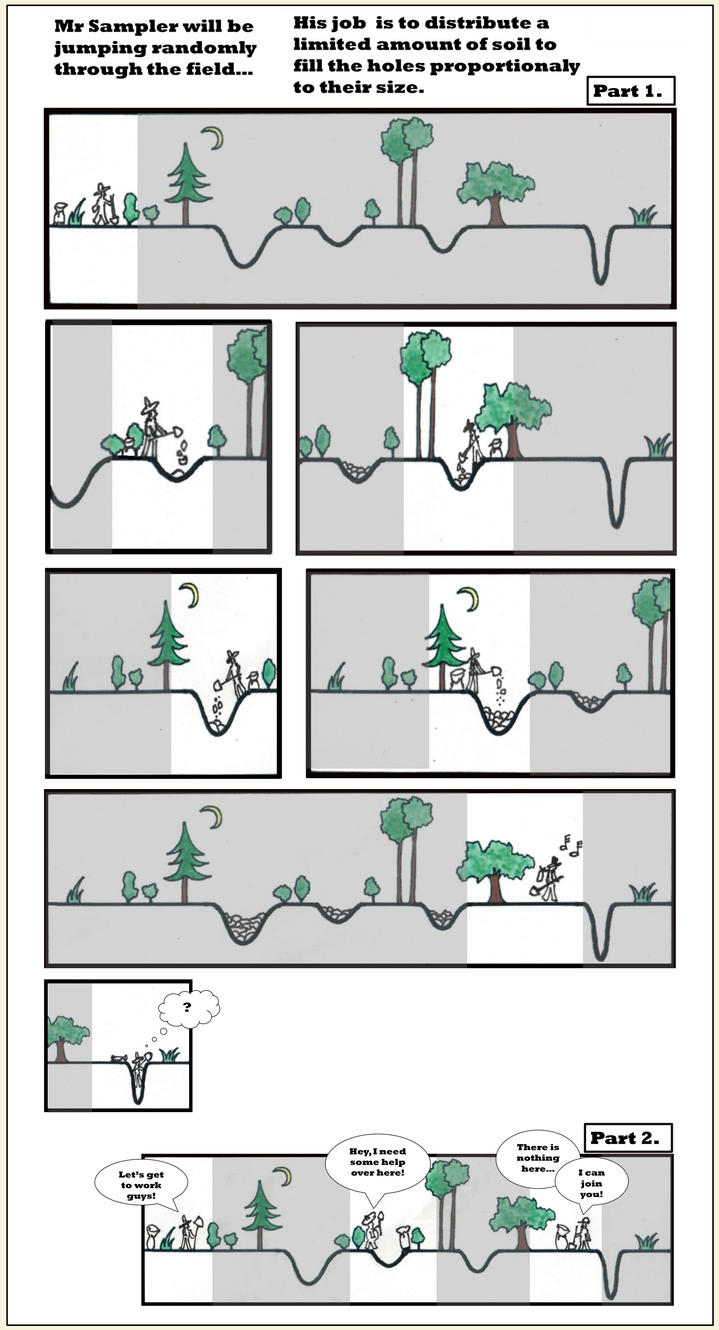
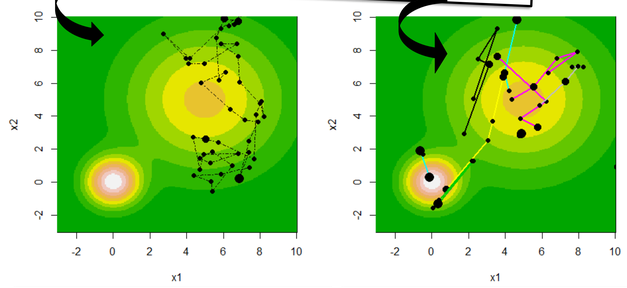
HOW TO JUMP in ensemble based MCMC samplers

Sanda Dejanic[1], Andreas Scheidegger[1], Jörg Rieckermann[1], Carlo Albert[1]

[1]Eawag, Swiss Federal Institute of Aquatic Science and Technology, Dübendorf, Zürich, Switzerland



Single chain vs. Ensemble



Advantages of ensemble samplers:

- walkers learn from each other
- parallelization is still possible

Disadvantage:

- all walkers suffer from burn in

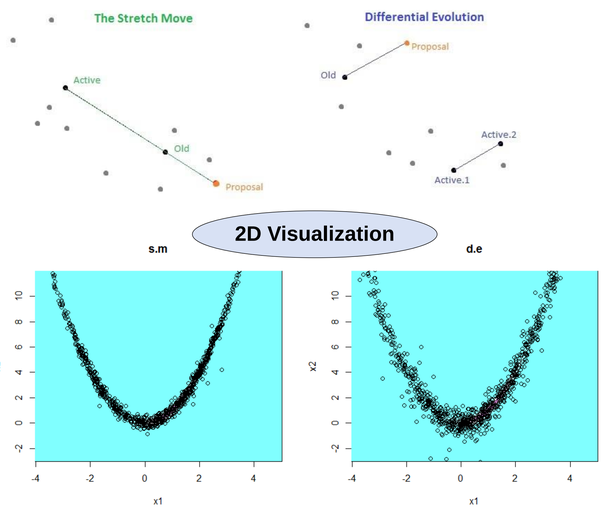
Comparing two algorithms

Result: Obtaining MCMC samples with s.m
for $t = 1, \ldots, T$ **do**
  for $k = 1, \ldots, L$ **do**
    choose x_j from $\mathbf{x}(t)$ at random;
    pick at random $z \in [\frac{1}{a}, a]$ where $g(z) \propto \frac{1}{\sqrt{z}}$;
    generate $Y = x_j + z(x_k(t) - x_j)$;
    if $z^{(n.dim-1)} \left(\frac{\pi_Y}{\pi_{x_k(t)}}\right) >$ random(1) **then**
      $x_k(t+1) = Y$;
    else
      $x_k(t+1) = x_k(t)$;
    end
  end
end

Algorithm 1: Stretch move algorithm.

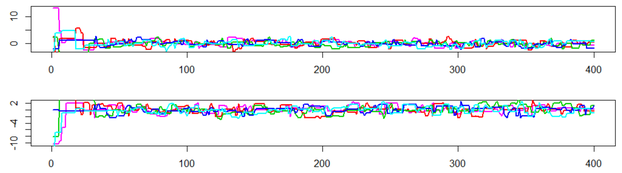
Result: Obtaining MCMC samples with d.e.
for $t = 1, \ldots, T$ **do**
  for $k = 1, \ldots, L$ **do**
    choose x_j and x_m from $\mathbf{x}(t)$ at random;
    $z = \frac{2.38}{\sqrt{2 \cdot n.dim}}$; **if** t%10 = 0 **then** z = 1;
    generate $Y = x_k(t) + z(x_l - x_m)$;
    if $\frac{\pi_Y}{\pi_{x_k(t)}} >$ random(1) **then**
      $x_k(t+1) = Y$;
    else
      $x_k(t+1) = x_k(t)$;
    end
  end
end

Algorithm 2: Differential evolution algorithm.



Performance measures

Convergence assessment



Chain diagnostics

- mean stabilization
- good or bad mixing

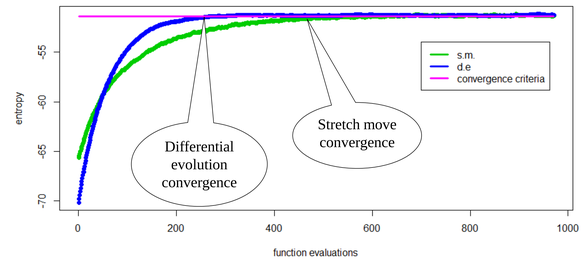
Entropy measure

- empirical entropy: mean function evaluation within the ensemble
- analytical entropy:

$$S = \int f(\mathbf{x}) \ln f(\mathbf{x}) d\mathbf{x}$$

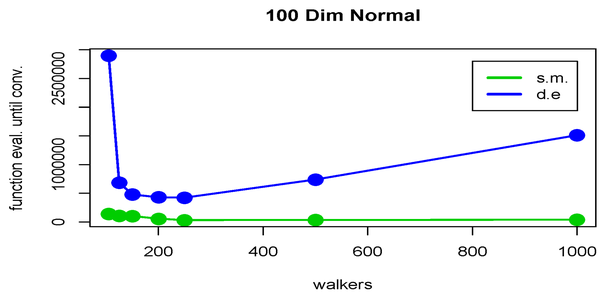
- the convergence criteria:

$S_{chains} \approx S$ (within one standard deviation)



Robustness

- In high dimensions, the sensitivity to the tuning parameters becomes prominent.



Conclusion

- We assessed the samplers by means of their convergence speed, robustness and effective sample sizes.
- For some cases of posteriors characterized by multimodality differential evolution has shown to be superior.
- For posteriors with strongly non-linear features as well as high dimensional normal distributions, we found that the stretch move outperformes the differential evolution move with respect to all three aspects.

Our R package

HOW TO JUMP
Poster and abstract

Aknowledgements: This research is part of the QUICS project funded from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 607000.

References:

Ensamble samples with affine invariance, Goodman, J. and Weare, J. Communications in Applied Mathematics and Computational Science Vol. 5 (2010), No. 1, 65–80. doi: 10.2140/camcos.2010.5.65

A Markov Chain Monte Carlo version of the genetic algorithm Differential Evolution: easy Bayesian computing for real parameter spaces, Braak, C.J.F.T. Stat Comput (2006) 16: 239. doi:10.1007/s11222-006-8769-1