Appendix A: Simulation Parameters

In this section we present the parameters of the algorithms used to solve two-dimensional and SK spin-glass problems with bimodal and Gaussian disorder, then explain the parameters used to calculate the TTSs of the DA and SA when solving spin-glass problems with different densities.

2D & SK Spin-Glass Problems

For the DA, SA, and PT (PT+ICM) and each problem class, we have performed a grid search in the parameter space to determine the best parameters, using a subset of problem instances. The subset of instances used for parameter tuning include instances of size 576, 784, and 1024. The number of instances solved to optimality, the success probability, and the residual energy have been used to select the best parameter combination for the benchmarking study.

In order to find the optimal TTS for a given problem size, we have varied the number of sweeps (iterations in the DA and the PTDA) and have used the procedure shown in Algorithm 4 to find the empirical distribution of the TTS for each number of sweeps. We have then found the number of sweeps for which the TTS distribution has the lowest mean and report the statistics of that distribution as the optimal TTS results. In the calculation of the TTS, we have excluded the initialization and post-processing times. The time spent on replica-exchange moves, currently performed via CPU, has been considered to be part of the PTDA's execution time. The time that it takes to execute one run of SA and PT (and PT+ICM) has been measured using *r4.8xlarge* Amazon EC2 instances, which consist of Intel Xeon E5-2686 v4 (Broadwell) processors.

To set the grid for the high and the low temperatures, we have simulated the distribution of the energy differences associated with proposed moves and search in the vicinity of the 5th (80th) to the 10th (85th) percentiles of this distribution to find the best-performing low (high) temperature value. In PT (PT+ICM), because we are able to measure different quantities during the simulation, we have further ensured that the highest temperature has been chosen such that the Monte Carlo acceptance probabilities are between 0.6 and 0.8. The parameter values used in each algorithm are outlined below.

The DA parameters

For all experiments, we have used 100 runs, each starting at a vector of zeros. The temperature schedule is linear in the inverse temperature, and the temperature has been adjusted after every iteration. The DA uses the Metropolis criterion to accept Monte Carlo moves. It is worth mentioning that our early experimentation has suggested better performance for the linear inverse temperature schedule than the exponential temperature schedule.

Our investigation of different parameter combinations has further shown that the performance of the DA on spin-glass problems is indifferent to the dynamic offset mechanism. Therefore, we turn this feature off for our final experimentation. The high (T_h) and the low (T_l) temperatures used for each problem class are shown in Table I.

TABLE I: High (T_h) and low (T_l) temperatures used in the DA, SA, and PT runs. Note that the row marked with "PT" also includes the PT+ICM parameters. The low and the high temperatures of SA for 2D-bimodal and SK-bimodal are selected according to Refs. [46, 47]. The temperature values used in our simulations are unitless.

	2D-bimodal		2D-Gaussian		SK-bimodal		SK-Gaussian	
	T_h	T_l	T_h	T_l	T_h	T_l	T_h	T_l
DA	2	0.66	$2 \cdot 10^4$	$4 \cdot 10^2$	40	4	$2 \cdot 10^5$	10^4
SA	10	0.33	10^6	10^4	$\sqrt{N}$	1	10^7	10^5
PT	2	0.33	10^6	$2 \cdot 10^4$	80	2	10^7	$5 \cdot 10^5$

SA parameters

Each instance has been solved 100 times using SA and the temperature schedule has been set to be linear in the inverse temperature, which is a typical choice for SA in the literature [46, 47]. The high (T_h) and the low (T_l) temperatures used for each problem class are shown in Table I.

PT (PT+ICM) parameters

Although the performance of replica-exchange algorithms is significantly dependent on the choice of temperature schedule, the temperatures at each replica have been set based on the commonly used geometric schedule [42]. After determining the low and the high temperatures, the number of replicas has been chosen such that the replica-exchange acceptance probabilities are above 0.2. The number of replicas has been set to 25, 60, 50, and 60, respectively, for the 2D-bimodal, SK-bimodal, 2D-Gaussian, and SK-Gaussian instances. In contrast to the runs of the DA and SA, the replicas in PT (PT+ICM) are not independent and to calculate the TTS, the whole PT (PT+ICM) algorithm has been repeated 30 times for each instance. This is time consuming, so each run (repeat) has been stopped immediately if the reference solution has been found. The high (T_{h}) and the low (T_{l}) temperatures used for each problem class are shown in Table I.

The PTDA parameters

The number of replicas has been set to 40 for both the SK-bimodal and SK-Gaussian problem instances, and the dynamic offset feature has been turned off. The high and the low temperatures and the temperature schedule are set internally by an automatic parameter-tuning strategy (see also Sec. II). As done for PT, since the replicas are dependent, the whole algorithm has been repeated 30 times to have enough observations to calculate the TTS for each instance.

Spin-Glass Problems with Different Densities

A grid-search approach on a subset of instances has been used to tune the parameters of the DA and SA for spin-glass problems, as explained in the previous section, separately for each density. The parameters of the DA and SA are the same as the ones used for 2D and SK spin-glass problems, except for the temperature values that are given below.

TABLE II: High (T_{h}) and low (T_{l}) temperatures used in the DA and SA runs of spin-glass problems with different densities. The temperature values used in our simulations are unitless.

	DA		SA	
d (%)	T_{h}	T_{l}	T_{h}	T_{l}
10	20	2.0	14	0.5
20	50	2.0	20	0.5
30	60	2.0	24	0.5
40	65	2.5	28	1.0
50	75	3.0	30	1.0
60	75	3.0	34	1.0
70	60	3.5	36	1.0
80	65	3.5	38	1.0
90	70	4.0	40	1.0